

Personalized cancer diagnosis

1. Business Problem

1.1. Description

Source: <https://www.kaggle.com/c/msk-redefining-cancer-treatment/>

Data: Memorial Sloan Kettering Cancer Center (MSKCC)

Download training_variants.zip and training_text.zip from Kaggle.

Context:

Source: <https://www.kaggle.com/c/msk-redefining-cancer-treatment/discussion/35336#198462>

Problem statement :

Classify the given genetic variations/mutations based on evidence from text-based clinical literature.

1.2. Source/Useful Links

Some articles and reference blogs about the problem statement

1. <https://www.forbes.com/sites/matthewherper/2017/06/03/a-new-cancer-drug-helped-almost-everyone-who-took-it-almost-heres-what-it-teaches-us/#2a44ee2f6b25>
[\(https://www.forbes.com/sites/matthewherper/2017/06/03/a-new-cancer-drug-helped-almost-everyone-who-took-it-almost-heres-what-it-teaches-us/#2a44ee2f6b25\)](https://www.forbes.com/sites/matthewherper/2017/06/03/a-new-cancer-drug-helped-almost-everyone-who-took-it-almost-heres-what-it-teaches-us/#2a44ee2f6b25)
2. <https://www.youtube.com/watch?v=UwbuW7oK8rk> (<https://www.youtube.com/watch?v=UwbuW7oK8rk>)
3. <https://www.youtube.com/watch?v=qxXRKVompl8> (<https://www.youtube.com/watch?v=qxXRKVompl8>)

1.3. Real-world/Business objectives and constraints.

- No low-latency requirement.
- Interpretability is important.
- Errors can be very costly.
- Probability of a data-point belonging to each class is needed.

2. Machine Learning Problem Formulation

2.1. Data

2.1.1. Data Overview

- Source: <https://www.kaggle.com/c/msk-redefining-cancer-treatment/data> (<https://www.kaggle.com/c/msk-redefining-cancer-treatment/data>)
- We have two data files: one contains the information about the genetic mutations and the other contains the clinical evidence (text) that human experts/pathologists use to classify the genetic mutations.
- Both these data files have a common column called ID
- Data file's information:
 - training_variants (ID , Gene, Variations, Class)
 - training_text (ID, Text)

2.1.2. Example Data Point

training_variants

ID,Gene,Variation,Class

0,FAM58A,Truncating Mutations,1

1,CBL,W802*,2

2,CBL,Q249E,2

...

training_text

ID,Text

0||Cyclin-dependent kinases (CDKs) regulate a variety of fundamental cellular processes. CDK10 stands out as one of the last orphan CDKs for which no activating cyclin has been identified and no kinase activity revealed. Previous work has shown that CDK10 silencing increases ETS2 (v-ets erythroblastosis virus E26 oncogene homolog 2)-driven activation of the MAPK pathway, which confers tamoxifen resistance to breast cancer cells. The precise mechanisms by which CDK10 modulates ETS2 activity, and more generally the functions of CDK10, remain elusive. Here we demonstrate that CDK10 is a cyclin-dependent kinase by identifying cyclin M as an activating cyclin. Cyclin M, an orphan cyclin, is the product of FAM58A, whose mutations cause STAR syndrome, a human developmental anomaly whose features include toe syndactyly, telecanthus, and anogenital and renal malformations. We show that STAR syndrome-associated cyclin M mutants are unable to interact with CDK10. Cyclin M silencing phenocopies CDK10 silencing in increasing c-Raf and in conferring tamoxifen resistance to breast cancer cells. CDK10/cyclin M phosphorylates ETS2 in vitro, and in cells it positively controls ETS2 degradation by the proteasome. ETS2 protein levels are increased in cells derived from a STAR patient, and this increase is attributable to decreased cyclin M levels. Altogether, our results reveal an additional regulatory mechanism for ETS2, which plays key roles in cancer and development. They also shed light on the molecular mechanisms underlying STAR syndrome. Cyclin-dependent kinases (CDKs) play a pivotal role in the control of a number of fundamental cellular processes (1). The human genome contains 21 genes encoding proteins that can be considered as members of the CDK family owing to their sequence similarity with bona fide CDKs, those known to be activated by cyclins (2). Although discovered almost 20 y ago (3, 4), CDK10 remains one of the two CDKs without an identified cyclin partner. This knowledge gap has largely impeded the exploration of its biological functions. CDK10 can act as a positive cell cycle regulator in some cells (5, 6) or as a tumor suppressor in others (7, 8). CDK10 interacts with the ETS2 (v-ets erythroblastosis virus E26 oncogene homolog 2) transcription factor and inhibits its transcriptional activity through an unknown mechanism (9). CDK10 knockdown derepresses ETS2, which increases the expression of the c-Raf protein kinase, activates the MAPK pathway, and induces resistance of MCF7 cells to tamoxifen (6). ...

2.2. Mapping the real-world problem to an ML problem

2.2.1. Type of Machine Learning Problem

There are nine different classes a genetic mutation can be classified into => Multi class classification problem

2.2.2. Performance Metric

Source: <https://www.kaggle.com/c/msk-redefining-cancer-treatment#evaluation> (<https://www.kaggle.com/c/msk-redefining-cancer-treatment#evaluation>)

Metric(s):

- Multi class log-loss
- Confusion matrix

2.2.3. Machine Learning Objectives and Constraints

Objective: Predict the probability of each data-point belonging to each of the nine classes.

Constraints:

- Interpretability
- Class probabilities are needed.
- Penalize the errors in class probabilities => Metric is Log-loss.
- No Latency constraints.

2.3. Train, CV and Test Datasets

Split the dataset randomly into three parts train, cross validation and test with 64%, 16%, 20% of data respectively

3. Exploratory Data Analysis

```
In [9]: # https://machinelearningmastery.com/how-to-fix-futurewarning-messages-in-scikit-learn/
from warnings import simplefilter
# ignore all future warnings
simplefilter(action='ignore', category=FutureWarning)

import pandas as pd
import matplotlib.pyplot as plt
import re
import time
import warnings
import numpy as np
from nltk.corpus import stopwords
from sklearn.decomposition import TruncatedSVD
from sklearn.preprocessing import normalize
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.manifold import TSNE
import seaborn as sns
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix
from sklearn.metrics.classification import accuracy_score, log_loss
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import SGDClassifier
from imblearn.over_sampling import SMOTE
from collections import Counter
from scipy.sparse import hstack
from sklearn.multiclass import OneVsRestClassifier
from sklearn.svm import SVC
from sklearn.model_selection import StratifiedKFold
from collections import Counter, defaultdict
from sklearn.calibration import CalibratedClassifierCV
from sklearn.naive_bayes import MultinomialNB
from sklearn.naive_bayes import GaussianNB
from sklearn.model_selection import train_test_split
from sklearn.model_selection import GridSearchCV
import math
from sklearn.metrics import normalized_mutual_info_score
from sklearn.ensemble import RandomForestClassifier
warnings.filterwarnings("ignore")

from mlxtend.classifier import StackingClassifier

from sklearn import model_selection
from sklearn.linear_model import LogisticRegression
```

3.1. Reading Data

3.1.1. Reading Gene and Variation Data

```
In [2]: data = pd.read_csv('training/training_variants')
print('Number of data points : ', data.shape[0])
print('Number of features : ', data.shape[1])
print('Features : ', data.columns.values)
data.head()
```

Number of data points : 3321
 Number of features : 4
 Features : ['ID' 'Gene' 'Variation' 'Class']

Out[2]:

	ID	Gene	Variation	Class
0	0	FAM58A	Truncating Mutations	1
1	1	CBL	W802*	2
2	2	CBL	Q249E	2
3	3	CBL	N454D	3
4	4	CBL	L399V	4

training/training_variants is a comma separated file containing the description of the genetic mutations used for training.

Fields are

- **ID** : the id of the row used to link the mutation to the clinical evidence
- **Gene** : the gene where this genetic mutation is located
- **Variation** : the aminoacid change for this mutations
- **Class** : 1-9 the class this genetic mutation has been classified on

3.1.2. Reading Text Data

```
In [3]: # note the separator in this file
data_text = pd.read_csv("training/training_text", sep="\|\|", engine="python", names=["ID", "TEXT"], skiprows=1)
print('Number of data points : ', data_text.shape[0])
print('Number of features : ', data_text.shape[1])
print('Features : ', data_text.columns.values)
data_text.head()
```

Number of data points : 3321
 Number of features : 2
 Features : ['ID' 'TEXT']

Out[3]:

	ID	TEXT
0	0	Cyclin-dependent kinases (CDKs) regulate a var...
1	1	Abstract Background Non-small cell lung canc...
2	2	Abstract Background Non-small cell lung canc...
3	3	Recent evidence has demonstrated that acquired...
4	4	Oncogenic mutations in the monomeric Casitas B...

3.1.3. Preprocessing of text

```
In [4]: # Loading stop words from nltk library
stop_words = set(stopwords.words('english'))

def nlp_preprocessing(total_text, index, column):
    if type(total_text) is not int:
        string = ""
        # replace every special char with space
        total_text = re.sub('[^a-zA-Z0-9\n]', ' ', total_text)
        # replace multiple spaces with single space
        total_text = re.sub('\s+', ' ', total_text)
        # converting all the chars into lower-case.
        total_text = total_text.lower()

        for word in total_text.split():
            # if the word is a not a stop word then retain that word from the data
            if not word in stop_words:
                string += word + " "

        data_text[column][index] = string
```

```
In [5]: #text processing stage.
start_time = time.clock()
for index, row in data_text.iterrows():
    if type(row['TEXT']) is str:
        nlp_preprocessing(row['TEXT'], index, 'TEXT')
    else:
        print("there is no text description for id:",index)
print('Time took for preprocessing the text :',time.clock() - start_time, "seconds")
```

there is no text description for id: 1109
 there is no text description for id: 1277
 there is no text description for id: 1407
 there is no text description for id: 1639
 there is no text description for id: 2755
 Time took for preprocessing the text : 182.716755298 seconds

```
In [6]: #merging both gene_variations and text data based on ID
result = pd.merge(data, data_text, on='ID', how='left')
result.head()
```

Out[6]:

	ID	Gene	Variation	Class	TEXT
0	0	FAM58A	Truncating Mutations	1	cyclin dependent kinases cdks regulate variety...
1	1	CBL	W802*	2	abstract background non small cell lung cancer...
2	2	CBL	Q249E	2	abstract background non small cell lung cancer...
3	3	CBL	N454D	3	recent evidence demonstrated acquired uniparen...
4	4	CBL	L399V	4	oncogenic mutations monomeric casitas b lineage...

```
In [7]: result[result.isnull().any(axis=1)]
```

Out[7]:

	ID	Gene	Variation	Class	TEXT
1109	1109	FANCA	S1088F	1	NaN
1277	1277	ARID5B	Truncating Mutations	1	NaN
1407	1407	FGFR3	K508M	6	NaN
1639	1639	FLT1	Amplification	6	NaN
2755	2755	BRAF	G596C	7	NaN

```
In [8]: result.loc[result['TEXT'].isnull(),'TEXT'] = result['Gene'] + ' '+result['Variation']
```

```
In [9]: result[result['ID']==1109]
```

Out[9]:

	ID	Gene	Variation	Class	TEXT
1109	1109	FANCA	S1088F	1	FANCA S1088F

3.1.4. Test, Train and Cross Validation Split

3.1.4.1. Splitting data into train, test and cross validation (64:20:16)

```
In [10]: y_true = result['Class'].values
result.Gene      = result.Gene.str.replace('\s+', '_')
result.Variation = result.Variation.str.replace('\s+', '_')

# split the data into test and train by maintaining same distribution of output variable 'y_true' [stratify=y_true]
X_train, test_df, y_train, y_test = train_test_split(result, y_true, stratify=y_true, test_size=0.2)
# split the train data into train and cross validation by maintaining same distribution of output variable 'y_train' [stratify=y_train]
train_df, cv_df, y_train, y_cv = train_test_split(X_train, y_train, stratify=y_train, test_size=0.2)
```

In [9]: y_train

Out[9]: array([7, 2, 5, ..., 4, 7, 5], dtype=int64)

We split the data into train, test and cross validation data sets, preserving the ratio of class distribution in the original data set

```
In [11]: print('Number of data points in train data:', train_df.shape[0])
print('Number of data points in test data:', test_df.shape[0])
print('Number of data points in cross validation data:', cv_df.shape[0])
```

Number of data points in train data: 2124
Number of data points in test data: 665
Number of data points in cross validation data: 532

3.1.4.2. Distribution of y_i's in Train, Test and Cross Validation datasets

In [12]: pd.__version__

Out[12]: '0.24.2'

```
In [13]: # it returns a dict, keys as class labels and values as the number of data points in that class
train_class_distribution = train_df['Class'].value_counts().sort_index()
test_class_distribution = test_df['Class'].value_counts().sort_index()
cv_class_distribution = cv_df['Class'].value_counts().sort_index()

my_colors = ['r','g','b','k','y','m','c','m','m']
# my_colors = List(islice(cycle(['b', 'r', 'g', 'y', 'k', 'r', 'g', 'y', 'k']), None, Len(train_class_distribution)))
train_class_distribution.plot(kind='bar',color =my_colors)
plt.xlabel('Class')
plt.ylabel('Data points per Class')
plt.title('Distribution of yi in train data')
plt.grid()
plt.show()

# ref: argsort https://docs.scipy.org/doc/numpy/reference/generated/numpy.argsort.html
# -(train_class_distribution.values): the minus sign will give us in decreasing order
sorted_yi = np.argsort(-train_class_distribution.values)
for i in sorted_yi:
    print('Number of data points in class', i+1, ':',train_class_distribution.values[i], '(', np.round((train_class_distribution.values[i]/train_df.shape[0]*100), 3), '%)')

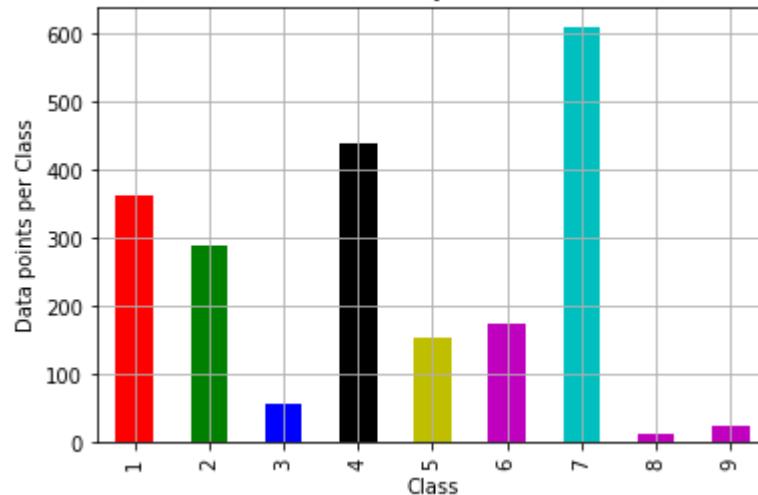
print('*'*80)
# my_colors = 'rgbkymc'
test_class_distribution.plot(kind='bar',color =my_colors)
plt.xlabel('Class')
plt.ylabel('Data points per Class')
plt.title('Distribution of yi in test data')
plt.grid()
plt.show()

# ref: argsort https://docs.scipy.org/doc/numpy/reference/generated/numpy.argsort.html
# -(train_class_distribution.values): the minus sign will give us in decreasing order
sorted_yi = np.argsort(-test_class_distribution.values)
for i in sorted_yi:
    print('Number of data points in class', i+1, ':',test_class_distribution.values[i], '(', np.round((test_class_distribution.values[i]/test_df.shape[0]*100), 3), '%)')

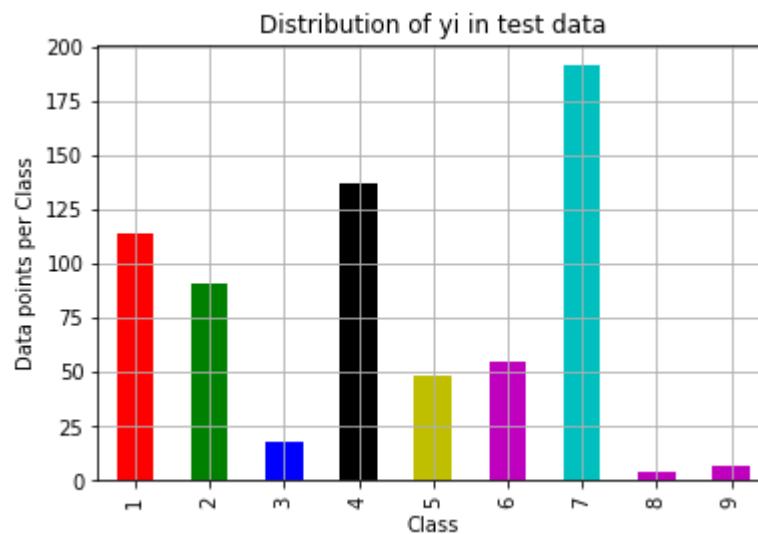
print('*'*80)
# my_colors = 'rgbkymc'
cv_class_distribution.plot(kind='bar',color =my_colors)
plt.xlabel('Class')
plt.ylabel('Data points per Class')
plt.title('Distribution of yi in cross validation data')
plt.grid()
plt.show()

# ref: argsort https://docs.scipy.org/doc/numpy/reference/generated/numpy.argsort.html
```

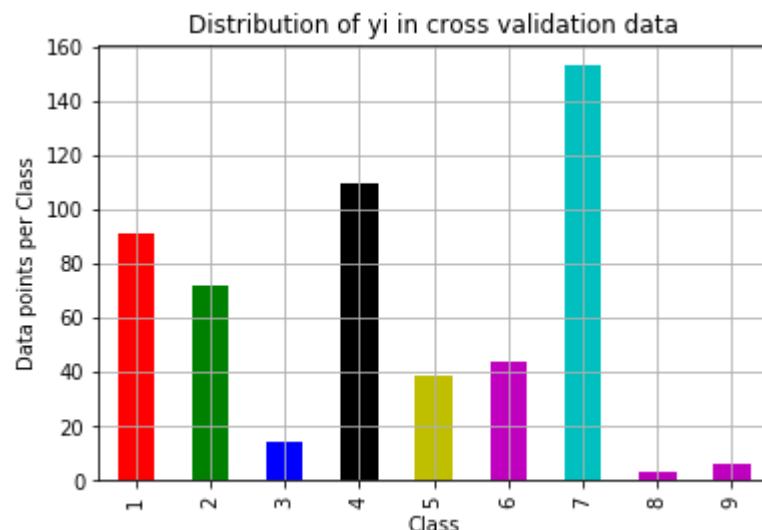
```
ort.html
# -(train_class_distribution.values): the minus sign will give us in decreasing order
sorted_yi = np.argsort(-train_class_distribution.values)
for i in sorted_yi:
    print('Number of data points in class', i+1, ':', cv_class_distribution.values[i], '(', np.round((cv_class_distribution.values[i]/cv_df.shape[0])*100), 3), '%')
```

Distribution of y_i in train data

Number of data points in class 7 : 609 (28.672 %)
 Number of data points in class 4 : 439 (20.669 %)
 Number of data points in class 1 : 363 (17.09 %)
 Number of data points in class 2 : 289 (13.606 %)
 Number of data points in class 6 : 176 (8.286 %)
 Number of data points in class 5 : 155 (7.298 %)
 Number of data points in class 3 : 57 (2.684 %)
 Number of data points in class 9 : 24 (1.13 %)
 Number of data points in class 8 : 12 (0.565 %)



Number of data points in class 7 : 191 (28.722 %)
 Number of data points in class 4 : 137 (20.602 %)
 Number of data points in class 1 : 114 (17.143 %)
 Number of data points in class 2 : 91 (13.684 %)
 Number of data points in class 6 : 55 (8.271 %)
 Number of data points in class 5 : 48 (7.218 %)
 Number of data points in class 3 : 18 (2.707 %)
 Number of data points in class 9 : 7 (1.053 %)
 Number of data points in class 8 : 4 (0.602 %)



Number of data points in class 7 : 153 (28.759 %)
 Number of data points in class 4 : 110 (20.677 %)
 Number of data points in class 1 : 91 (17.105 %)
 Number of data points in class 2 : 72 (13.534 %)
 Number of data points in class 6 : 44 (8.271 %)
 Number of data points in class 5 : 39 (7.331 %)
 Number of data points in class 3 : 14 (2.632 %)
 Number of data points in class 9 : 6 (1.128 %)
 Number of data points in class 8 : 3 (0.564 %)

In [14]: `train_class_distribution`

Out[14]:

1	363
2	289
3	57
4	439
5	155
6	176
7	609
8	12
9	24

Name: Class, dtype: int64

In [15]:

```
# X_train = X_train.drop(['Class'],axis=1)
# test_df = test_df.drop(['Class'],axis=1)
# cv_df = cv_df.drop(['Class'],axis=1)
```

3.2 Prediction using a 'Random' Model

In a 'Random' Model, we generate the NINE class probabilities randomly such that they sum to 1.

```
In [15]: # This function plots the confusion matrices given y_i, y_i_hat.

def plot_confusion_matrix(test_y, predict_y):
    C = confusion_matrix(test_y, predict_y)
    # C = 9,9 matrix, each cell (i,j) represents number of points of class i a
    re predicted class j

    A =(((C.T)/(C.sum(axis=1))).T)
    #divid each element of the confusion matrix with the sum of elements in th
    at column

    # C = [[1, 2],
    #      [3, 4]]
    # C.T = [[1, 3],
    #          [2, 4]]
    # C.sum(axis = 1) axis=0 corresonds to columns and axis=1 corresponds to
    rows in two dimensional array
    # C.sum(axis =1) = [[3, 7]]
    # ((C.T)/(C.sum(axis=1))) = [[1/3, 3/7]
    #                               [2/3, 4/7]]

    # ((C.T)/(C.sum(axis=1))).T = [[1/3, 2/3]
    #                               [3/7, 4/7]]
    # sum of row elements = 1

    B =(C/C.sum(axis=0))
    #divid each element of the confusion matrix with the sum of elements in th
    at row

    # C = [[1, 2],
    #      [3, 4]]
    # C.sum(axis = 0) axis=0 corresonds to columns and axis=1 corresponds to
    rows in two dimensional array
    # C.sum(axis =0) = [[4, 6]]
    # (C/C.sum(axis=0)) = [[1/4, 2/6],
    #                      [3/4, 4/6]]

    labels = [1,2,3,4,5,6,7,8,9]
    # representing A in heatmap format
    print("-"*20, "Confusion matrix", "*"-20)
    plt.figure(figsize=(20,7))
    sns.heatmap(C, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y
    ticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.show()

    print("-"*20, "Precision matrix (Column Sum=1)", "*"-20)
    plt.figure(figsize=(20,7))
    sns.heatmap(B, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y
    ticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.show()

    # representing B in heatmap format
    print("-"*20, "Recall matrix (Row sum=1)", "*"-20)
    plt.figure(figsize=(20,7))
```

```
    sns.heatmap(A, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y  
    ticklabels=labels)  
    plt.xlabel('Predicted Class')  
    plt.ylabel('Original Class')  
    plt.show()
```

```
In [16]: # we need to generate 9 numbers and the sum of numbers should be 1
# one solution is to generate 9 numbers and divide each of the numbers by their sum
# ref: https://stackoverflow.com/a/18662466/4084039
test_data_len = test_df.shape[0]
cv_data_len = cv_df.shape[0]

# we create a output array that has exactly same size as the CV data
cv_predicted_y = np.zeros((cv_data_len,9))
for i in range(cv_data_len):
    rand_probs = np.random.rand(1,9)
    cv_predicted_y[i] = ((rand_probs/sum(sum(rand_probs)))[0])
print("Log loss on Cross Validation Data using Random Model",log_loss(y_cv,cv_predicted_y, eps=1e-15))

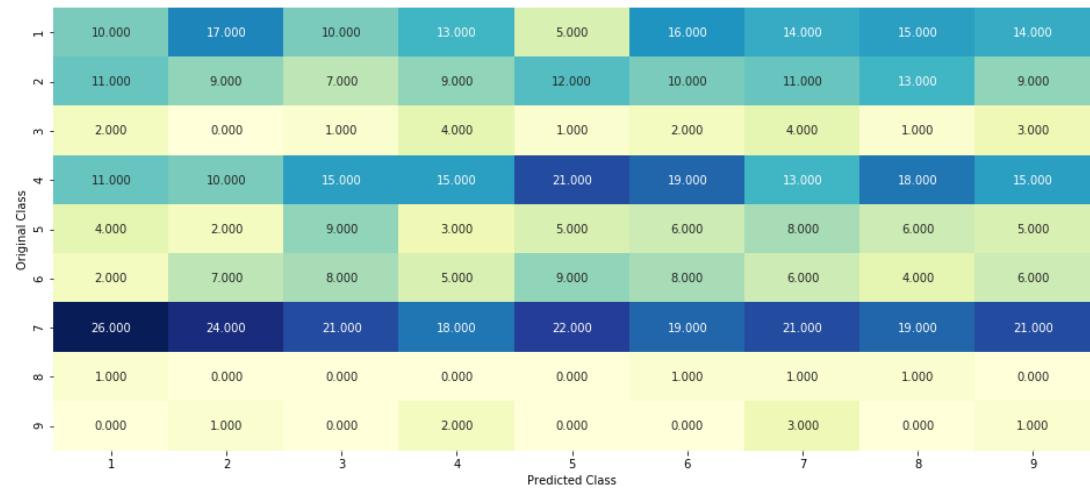
# Test-Set error.
#we create a output array that has exactly same as the test data
test_predicted_y = np.zeros((test_data_len,9))
for i in range(test_data_len):
    rand_probs = np.random.rand(1,9)
    test_predicted_y[i] = ((rand_probs/sum(sum(rand_probs)))[0])
print("Log loss on Test Data using Random Model",log_loss(y_test,test_predicted_y, eps=1e-15))

predicted_y =np.argmax(test_predicted_y, axis=1)
plot_confusion_matrix(y_test, predicted_y+1)
```

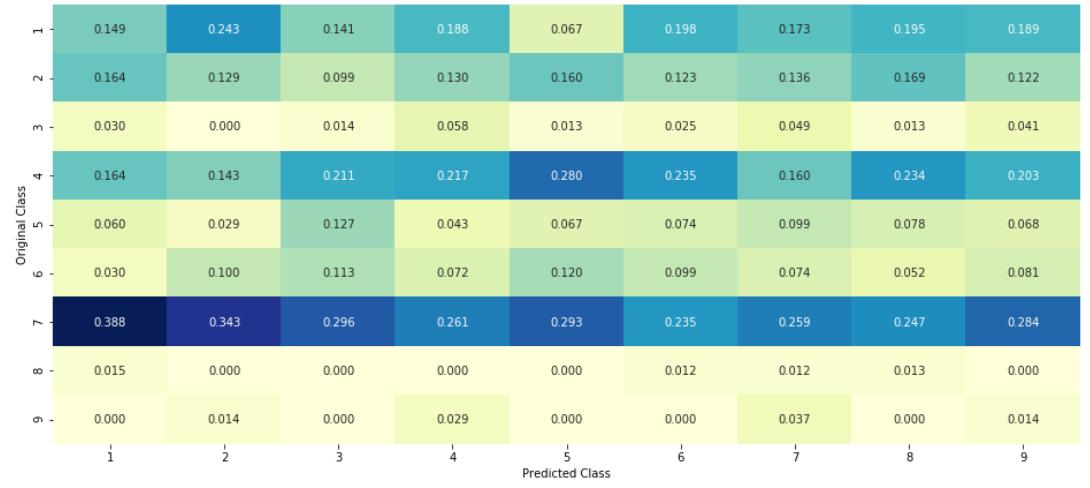
Log loss on Cross Validation Data using Random Model 2.561855138805523

Log loss on Test Data using Random Model 2.483427680683312

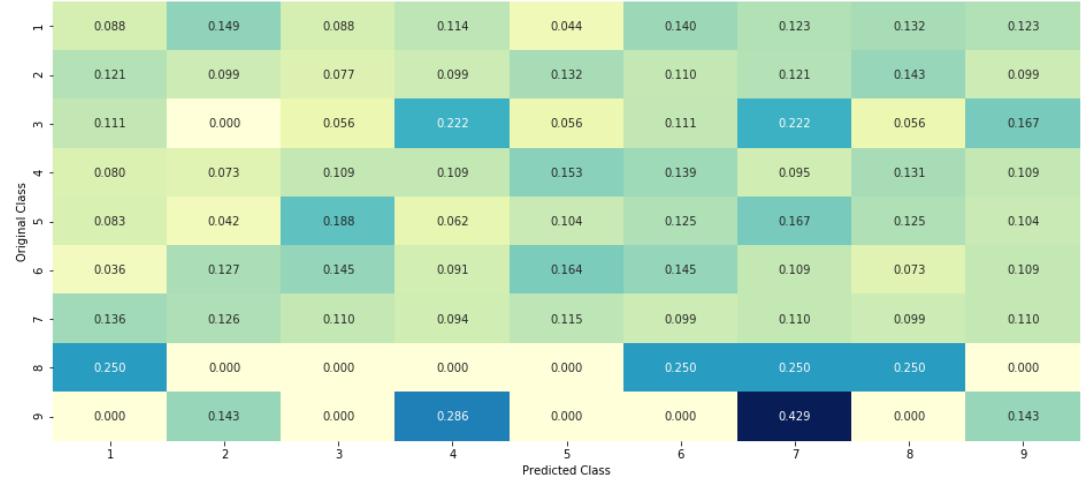
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



```
In [17]: rand_probs
```

```
Out[17]: array([[0.69932817, 0.12257837, 0.01296844, 0.29520686, 0.96352183,
 0.66346015, 0.19657565, 0.3315067 , 0.39964286]])
```

```
In [18]: sum(rand_probs)
```

```
Out[18]: array([0.69932817, 0.12257837, 0.01296844, 0.29520686, 0.96352183,
 0.66346015, 0.19657565, 0.3315067 , 0.39964286])
```

```
In [19]: sum(sum(rand_probs))
```

```
Out[19]: 3.6847890298728894
```

```
In [20]: rand_probs/sum(sum(rand_probs))
```

```
Out[20]: array([[0.18978785, 0.03326605, 0.00351945, 0.080115 , 0.26148629,
 0.18005377, 0.05334787, 0.08996626, 0.10845746]])
```

```
In [21]: (rand_probs/sum(sum(rand_probs)))[0]
```

```
Out[21]: array([0.18978785, 0.03326605, 0.00351945, 0.080115 , 0.26148629,
 0.18005377, 0.05334787, 0.08996626, 0.10845746])
```

```
In [22]: test_predicted_y[23].sum()
```

```
Out[22]: 1.0
```

3.3 Univariate Analysis

```
In [23]: # code for response coding with Laplace smoothing.
# alpha : used for Laplace smoothing
# feature: ['gene', 'variation']
# df: ['train_df', 'test_df', 'cv_df']
# algorithm
# -----
# Consider all unique values and the number of occurrences of given feature in
# train data dataframe
# build a vector (1*9) , the first element = (number of times it occurred in cl
ass1 + 10*alpha / number of time it occurred in total data+90*alpha)
# gv_dict is like a look up table, for every gene it store a (1*9) representat
ion of it
# for a value of feature in df:
# if it is in train data:
# we add the vector that was stored in 'gv_dict' Look up table to 'gv_fea'
# if it is not there is train:
# we add [1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9] to 'gv_fea'
# return 'gv_fea'
# -----
# get_gv_fea_dict: Get Gene variation Feature Dict
def get_gv_fea_dict(alpha, feature, df):
    # value_count: it contains a dict like
    # print(train_df['Gene'].value_counts())
    # output:
    #     {BRCA1      174
    #      TP53      106
    #      EGFR      86
    #      BRCA2      75
    #      PTEN      69
    #      KIT       61
    #      BRAF      60
    #      ERBB2      47
    #      PDGFRA     46
    #      ...}
    # print(train_df['Variation'].value_counts())
    # output:
    # {
    # Truncating_Mutations      63
    # Deletion                  43
    # Amplification              43
    # Fusions                   22
    # Overexpression              3
    # E17K                      3
    # Q61L                      3
    # S222D                     2
    # P130S                     2
    # ...
    # }
    value_count = train_df[feature].value_counts()

    # gv_dict : Gene Variation Dict, which contains the probability array for
    # each gene/variation
    gv_dict = dict()

    # denominator will contain the number of time that particular feature occu
```

```

red in whole data
for i, denominator in value_count.items():
    # vec will contain ( $p(y_i=1/G_i)$ ) probability of gene/variation belongs
    to particular class
    # vec is 9 dimensional vector
    vec = []
    for k in range(1,10):
        # print(train_df.loc[(train_df['Class']==1) & (train_df['Gene']=
        = 'BRCA1')])]

        #           ID   Gene          Variation Class
        # 2470  2470  BRCA1      S1715C     1
        # 2486  2486  BRCA1      S1841R     1
        # 2614  2614  BRCA1      M1R       1
        # 2432  2432  BRCA1      L1657P     1
        # 2567  2567  BRCA1      T1685A     1
        # 2583  2583  BRCA1      E1660G     1
        # 2634  2634  BRCA1      W1718L     1
        # cls_cnt.shape[0] will return the number of rows

    cls_cnt = train_df.loc[(train_df['Class']==k) & (train_df[feature]==i)]

    # cls_cnt.shape[0](numerator) will contain the number of time that
    # particular feature occurred in whole data
    vec.append((cls_cnt.shape[0] + alpha*10)/ (denominator + 90*alpha))
)

# we are adding the gene/variation to the dict as key and vec as value
gv_dict[i]=vec
return gv_dict

# Get Gene variation feature
def get_gv_feature(alpha, feature, df):
    # print(gv_dict)
    # {'BRCA1': [0.200757575757575, 0.03787878787878788, 0.068181818181818177, 0.13636363636363635, 0.25, 0.1931818181818181, 0.03787878787878788, 0.03787878787878788, 0.03787878787878788], 'TP53': [0.32142857142857145, 0.061224489795918366, 0.061224489795918366, 0.27040816326530615, 0.061224489795918366, 0.066326530612244902, 0.051020408163265307, 0.051020408163265307, 0.056122448979591837], 'EGFR': [0.0568181818181816, 0.21590909090909091, 0.0625, 0.06818181818177, 0.06818181818177, 0.0625, 0.34659090909090912, 0.0625, 0.0568181818181816], 'BRCA2': [0.1333333333333333, 0.060606060606060608, 0.060606060606060608, 0.0787878787878782, 0.1393939393939394, 0.34545454545454546, 0.060606060606060608, 0.060606060606060608, 0.060606060606060608], 'PTEN': [0.06918238937106917, 0.062893081761006289, 0.069182389937106917, 0.46540880503144655, 0.075471698113207544, 0.062893081761006289, 0.069182389937106917, 0.062893081761006289, 0.062893081761006289], 'KIT': [0.066225165562913912, 0.25165562913907286, 0.072847682119205295, 0.072847682119205295, 0.066225165562913912, 0.066225165562913912, 0.066225165562913912, 0.27152317880794702, 0.066225165562913912, 0.066225165562913912], 'BRAF': [0.066666666666666666, 0.1799999999999999, 0.073333333333333334, 0.07333333333333334, 0.09333333333333338, 0.08000000000000002, 0.2999999999999999, 0.066666666666666666, 0.066666666666666666], '# ...'
# }

```

```

gv_dict = get_gv_fea_dict(alpha, feature, df)
# value_count is similar in get_gv_fea_dict
value_count = train_df[feature].value_counts()

# gv_fea: Gene_variation feature, it will contain the feature for each feature value in the data
gv_fea = []
# for every feature values in the given data frame we will check if it is there in the train data then we will add the feature to gv_fea
# if not we will add [1/9,1/9,1/9,1/9,1/9,1/9,1/9,1/9] to gv_fea
for index, row in df.iterrows():
    if row[feature] in dict(value_count).keys():
        gv_fea.append(gv_dict[row[feature]])
    else:
        gv_fea.append([1/9,1/9,1/9,1/9,1/9,1/9,1/9,1/9])
#
return gv_fea

```

when we calculate the probability of a feature belongs to any particular class, we apply laplace smoothing

- $(\text{numerator} + 10^{*\alpha}) / (\text{denominator} + 90^{*\alpha})$

3.2.1 Univariate Analysis on Gene Feature

Q1. Gene, What type of feature it is ?

Ans. Gene is a categorical variable

Q2. How many categories are there and How they are distributed?

```
In [24]: unique_genes = train_df['Gene'].value_counts()
print('Number of Unique Genes :', unique_genes.shape[0])
# the top 10 genes that occurred most
print(unique_genes.head(10))
```

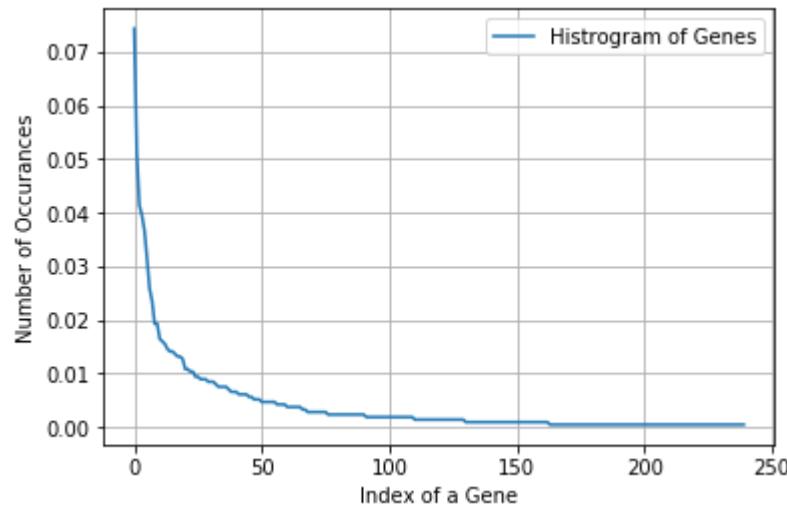
Gene	Count
BRCA1	158
TP53	109
EGFR	88
PTEN	84
BRCA2	78
BRAF	67
KIT	55
ERBB2	50
PIK3CA	41
ALK	41

Name: Gene, dtype: int64

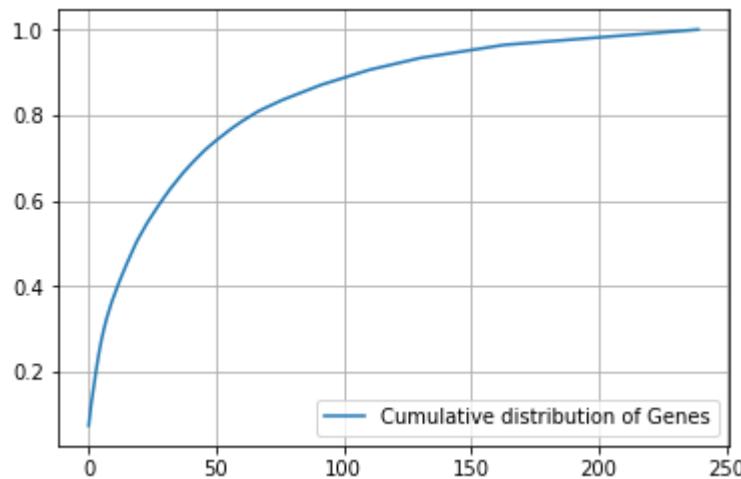
In [25]: `print("Ans: There are", unique_genes.shape[0] , "different categories of genes in the train data, and they are distributed as follows",)`

Ans: There are 240 different categories of genes in the train data, and they are distributed as follows

In [26]: `s = sum(unique_genes.values);
h = unique_genes.values/s;
plt.plot(h, label="Histrogram of Genes")
plt.xlabel('Index of a Gene')
plt.ylabel('Number of Occurrences')
plt.legend()
plt.grid()
plt.show()`



In [27]: `c = np.cumsum(h)
plt.plot(c,label='Cumulative distribution of Genes')
plt.grid()
plt.legend()
plt.show()`



Q3. How to featurize this Gene feature ?

Ans.there are two ways we can featurize this variable check out this video:

<https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/handling-categorical-and-numerical-features/>

1. One hot Encoding
2. Response coding

We will choose the appropriate featurization based on the ML model we use. For this problem of multi-class classification with categorical features, one-hot encoding is better for Logistic regression while response coding is better for Random Forests.

```
In [28]: #response-coding of the Gene feature
# alpha is used for Laplace smoothing
alpha = 1
# train gene feature
train_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", train_df))
# test gene feature
test_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", test_df))
# cross validation gene feature
cv_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", cv_df))
```

```
In [29]: print("train_gene_feature_responseCoding is converted feature using response coding method. The shape of gene feature:", train_gene_feature_responseCoding.shape)
```

train_gene_feature_responseCoding is converted feature using response coding method. The shape of gene feature: (2124, 9)

```
In [30]: # one-hot encoding of Gene feature.
gene_vectorizer = CountVectorizer()
train_gene_feature_onehotCoding = gene_vectorizer.fit_transform(train_df[ 'Gene' ])
test_gene_feature_onehotCoding = gene_vectorizer.transform(test_df[ 'Gene' ])
cv_gene_feature_onehotCoding = gene_vectorizer.transform(cv_df[ 'Gene' ])
```

```
In [31]: train_df[ 'Gene' ].head()
```

```
Out[31]: 2005    MAP2K1
2022    MAP2K1
2702    BRAF
1077    FOXA1
1554    ALK
Name: Gene, dtype: object
```

```
In [32]: gene_vectorizer.get_feature_names()
```

Out[32]: ['abl1',
'acvr1',
'ago2',
'akt1',
'akt2',
'akt3',
'alk',
'apc',
'ar',
'araf',
'arid1b',
'arid2',
'arid5b',
'asxl2',
'atm',
'atr',
'atrx',
'aurka',
'axin1',
'axl',
'b2m',
'bap1',
'bard1',
'bcl10',
'bcl2',
'bcl2l11',
'bcor',
'braf',
'brca1',
'brca2',
'brd4',
'brip1',
'btk',
'card11',
'carm1',
'casp8',
'cbl',
'ccnd1',
'ccnd3',
'ccne1',
'cdh1',
'cdk12',
'cdk6',
'cdk8',
'cdkn1a',
'cdkn1b',
'cdkn2a',
'cdkn2b',
'cebp',
'chek2',
'cic',
'crebbp',
'ctcf',
'ctla4',
'ctnnb1',
'ddr2',
'dicer1',

'dnmt3a',
'dnmt3b',
'dusp4',
'egfr',
'eif1ax',
'elf3',
'ep300',
'epas1',
'epcam',
'erbb2',
'erbb3',
'erbb4',
'ercc2',
'ercc4',
'erg',
'errfi1',
'esr1',
'etv1',
'etv6',
'ews1',
'ezh2',
'fam58a',
'fanca',
'fat1',
'fbxw7',
'fgf19',
'fgf3',
'fgf4',
'fgfr1',
'fgfr2',
'fgfr3',
'fgfr4',
'flt1',
'flt3',
'foxa1',
'foxl2',
'foxp1',
'gata3',
'gli1',
'gna11',
'gnaq',
'gnas',
'h3f3a',
'hist1h1c',
'hla',
'hnf1a',
'hras',
'idh1',
'idh2',
'igf1r',
'ikzf1',
'inpp4b',
'jak1',
'jak2',
'jun',
'kdm5a',
'kdm6a',

'kdr',
'keap1',
'kit',
'klf4',
'kmt2a',
'kmt2b',
'kmt2c',
'kmt2d',
'knstrn',
'kras',
'lats2',
'map2k1',
'map2k2',
'map2k4',
'map3k1',
'mapk1',
'mdm2',
'mdm4',
'med12',
'mef2b',
'men1',
'met',
'mlh1',
'mpl',
'msh2',
'msh6',
'mtor',
'myc',
'mycn',
'myd88',
'nf1',
'nf2',
'nfe212',
'nfbia',
'nkx2',
'notch1',
'notch2',
'npm1',
'nras',
'nsd1',
'ntrk1',
'ntrk2',
'ntrk3',
'nup93',
'pak1',
'pbrm1',
'pdgfra',
'pdgfrb',
'pik3ca',
'pik3cb',
'pik3cd',
'pik3r1',
'pik3r2',
'pik3r3',
'pim1',
'pms2',
'pole',

'ppm1d',
'ppp2r1a',
'ppp6c',
'prdm1',
'ptch1',
'pten',
'ptpn11',
'ptprd',
'ptprt',
'rab35',
'rac1',
'rad21',
'rad50',
'rad51b',
'rad51c',
'rad51d',
'rad541',
'raf1',
'rara',
'rasa1',
'rb1',
'rbm10',
'ret',
'rheb',
'rhoa',
'rictor',
'rit1',
'rnf43',
'ros1',
'runx1',
'rxra',
'rybp',
'sdhb',
'sdhc',
'setd2',
'sf3b1',
'shq1',
'smad2',
'smad3',
'smad4',
'smarca4',
'smarcb1',
'smo',
'sos1',
'sox9',
'spop',
'src',
'srsf2',
'stag2',
'stat3',
'stk11',
'tcf7l2',
'tert',
'tet1',
'tet2',
'tgfb1',
'tgfb2',

```
'tmpRSS2',
'tp53',
'tp53bp1',
'tsc1',
'tsc2',
'u2af1',
'vegfa',
'vhl',
'xpo1',
'xrcc2',
'yap1']
```

In [33]: `print("train_gene_feature_onehotCoding is converted feature using one-hot encoding method. The shape of gene feature:", train_gene_feature_onehotCoding.shape)`

```
train_gene_feature_onehotCoding is converted feature using one-hot encoding m
ethod. The shape of gene feature: (2124, 239)
```

Q4. How good is this gene feature in predicting y_i?

There are many ways to estimate how good a feature is, in predicting y_i . One of the good methods is to build a proper ML model using just this feature. In this case, we will build a logistic regression model using only Gene feature (one hot encoded) to predict y_i .

In [34]: alpha = [10 ** x **for** x **in** range(-5, 1)] # hyperparam for SGD classifier.

```
# read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.Linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='l2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link:
#-----
```



```
cv_log_error_array=[]
for i in alpha:
    clf = SGDClassifier(alpha=i, penalty='l2', loss='log', random_state=42)
    clf.fit(train_gene_feature_onehotCoding, y_train)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_gene_feature_onehotCoding, y_train)
    predict_y = sig_clf.predict_proba(cv_gene_feature_onehotCoding)
    cv_log_error_array.append(log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
    print('For values of alpha = ', i, "The log loss is:", log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
    ax.annotate((alpha[i],np.round(txt,3)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()
```



```
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
clf.fit(train_gene_feature_onehotCoding, y_train)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_gene_feature_onehotCoding, y_train)

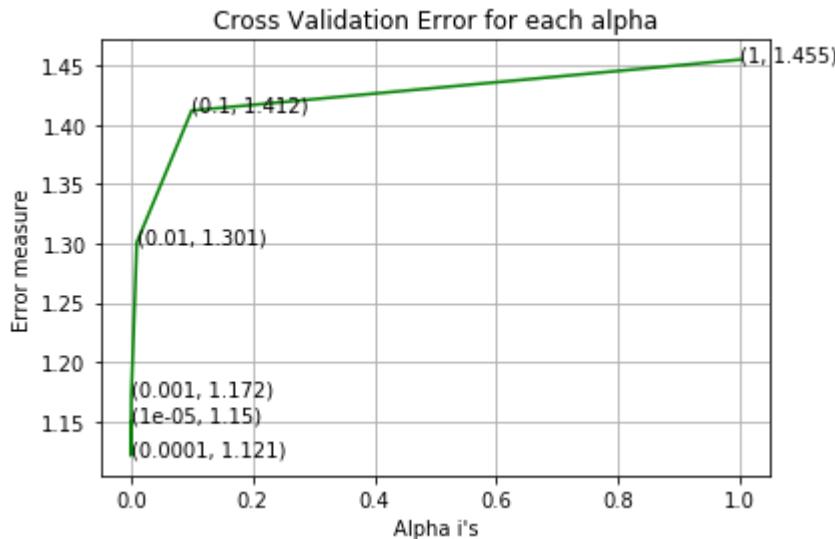
predict_y = sig_clf.predict_proba(train_gene_feature_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:", log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_gene_feature_onehotCoding)
```

```

print('For values of best alpha = ', alpha[best_alpha], "The cross validation
log loss is:", log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_gene_feature_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss i
s:", log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))

```

For values of alpha = 1e-05 The log loss is: 1.1496800237350675
 For values of alpha = 0.0001 The log loss is: 1.1208623007339227
 For values of alpha = 0.001 The log loss is: 1.171716229402704
 For values of alpha = 0.01 The log loss is: 1.3008678766794974
 For values of alpha = 0.1 The log loss is: 1.4120956379495808
 For values of alpha = 1 The log loss is: 1.4550767125563346



For values of best alpha = 0.0001 The train log loss is: 1.0120370228036482
 For values of best alpha = 0.0001 The cross validation log loss is: 1.120862
 3007339227
 For values of best alpha = 0.0001 The test log loss is: 1.2034754307980773

Q5. Is the Gene feature stable across all the data sets (Test, Train, Cross validation)?

Ans. Yes, it is. Otherwise, the CV and Test errors would be significantly more than train error.

```
In [35]: print("Q6. How many data points in Test and CV datasets are covered by the ", unique_genes.shape[0], " genes in train dataset?")

test_coverage=test_df[test_df['Gene'].isin(list(set(train_df['Gene'])))].shape[0]
cv_coverage=cv_df[cv_df['Gene'].isin(list(set(train_df['Gene'])))].shape[0]

print('Ans\n1. In test data',test_coverage, 'out of',test_df.shape[0], ":" ,(test_coverage/test_df.shape[0])*100)
print('2. In cross validation data',cv_coverage, 'out of ',cv_df.shape[0], ":" ,(cv_coverage/cv_df.shape[0])*100)
```

Q6. How many data points in Test and CV datasets are covered by the 240 genes in train dataset?

Ans

1. In test data 646 out of 665 : 97.14285714285714
2. In cross validation data 519 out of 532 : 97.55639097744361

3.2.2 Univariate Analysis on Variation Feature

Q7. Variation, What type of feature is it ?

Ans. Variation is a categorical variable

Q8. How many categories are there?

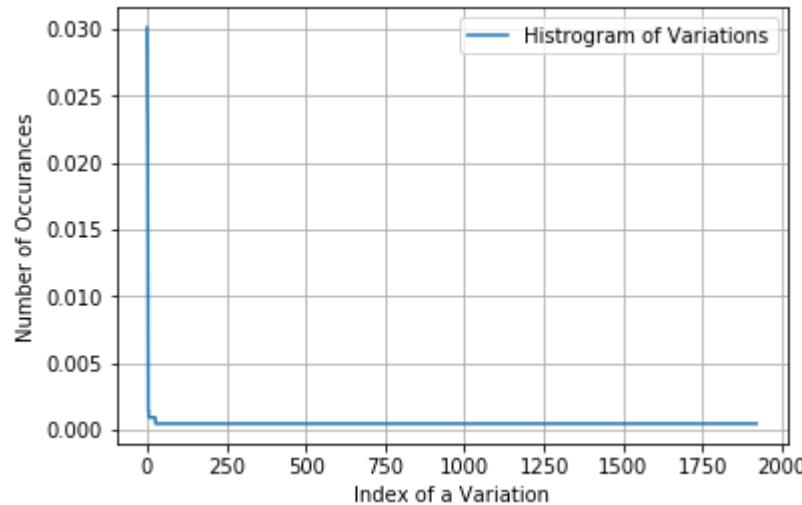
```
In [36]: unique_variations = train_df['Variation'].value_counts()
print('Number of Unique Variations :', unique_variations.shape[0])
# the top 10 variations that occurred most
print(unique_variations.head(10))
```

Number of Unique Variations :	1921
Truncating_Mutations	64
Amplification	52
Deletion	49
Fusions	17
T58I	3
Overexpression	3
G12C	2
A146T	2
T286A	2
Q209L	2
Name: Variation, dtype:	int64

```
In [37]: print("Ans: There are", unique_variations.shape[0] , "different categories of variations in the train data, and they are distributed as follows",)
```

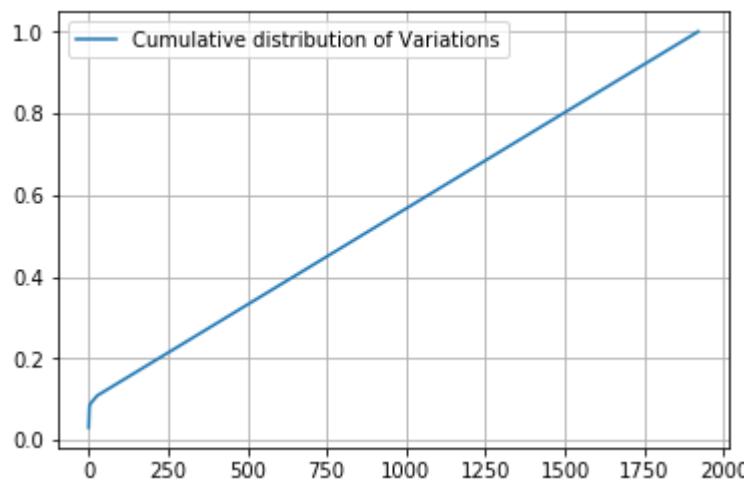
Ans: There are 1921 different categories of variations in the train data, and they are distributed as follows

```
In [38]: s = sum(unique_variations.values);
h = unique_variations.values/s;
plt.plot(h, label="Histogram of Variations")
plt.xlabel('Index of a Variation')
plt.ylabel('Number of Occurrences')
plt.legend()
plt.grid()
plt.show()
```



```
In [39]: c = np.cumsum(h)
print(c)
plt.plot(c,label='Cumulative distribution of Variations')
plt.grid()
plt.legend()
plt.show()
```

```
[0.03013183 0.05461394 0.07768362 ... 0.99905838 0.99952919 1.] ]
```



Q9. How to featurize this Variation feature ?

Ans. There are two ways we can featurize this variable check out this video:

<https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/handling-categorical-and-numerical-features/>

1. One hot Encoding
2. Response coding

We will be using both these methods to featurize the Variation Feature

```
In [40]: # alpha is used for Laplace smoothing
alpha = 1
# train gene feature
train_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", train_df))
# test gene feature
test_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", test_df))
# cross validation gene feature
cv_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", cv_df))
```

```
In [41]: print("train_variation_feature_responseCoding is a converted feature using the
response coding method. The shape of Variation feature:", train_variation_feature_responseCoding.shape)
```

train_variation_feature_responseCoding is a converted feature using the response coding method. The shape of Variation feature: (2124, 9)

```
In [42]: # one-hot encoding of variation feature.
variation_vectorizer = CountVectorizer()
train_variation_feature_onehotCoding = variation_vectorizer.fit_transform(train_df['Variation'])
test_variation_feature_onehotCoding = variation_vectorizer.transform(test_df['Variation'])
cv_variation_feature_onehotCoding = variation_vectorizer.transform(cv_df['Variation'])
```

```
In [43]: print("train_variation_feature_onehotEncoded is converted feature using the
one-hot encoding method. The shape of Variation feature:", train_variation_feature_onehotCoding.shape)
```

train_variation_feature_onehotEncoded is converted feature using the one-hot encoding method. The shape of Variation feature: (2124, 1955)

Q10. How good is this Variation feature in predicting y_i?

Let's build a model just like the earlier!

```
In [44]: alpha = [10 ** x for x in range(-5, 1)]

# read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.Linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='l2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link:
#-----


cv_log_error_array=[]
for i in alpha:
    clf = SGDClassifier(alpha=i, penalty='l2', loss='log', random_state=42)
    clf.fit(train_variation_feature_onehotCoding, y_train)

    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_variation_feature_onehotCoding, y_train)
    predict_y = sig_clf.predict_proba(cv_variation_feature_onehotCoding)

    cv_log_error_array.append(log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
    print('For values of alpha = ', i, "The log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))

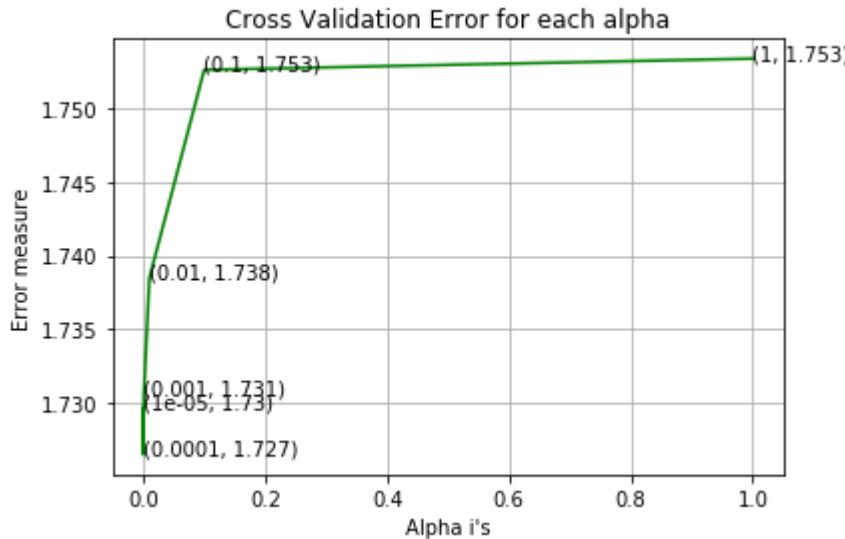
fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
    ax.annotate((alpha[i],np.round(txt,3)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
clf.fit(train_variation_feature_onehotCoding, y_train)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_variation_feature_onehotCoding, y_train)

predict_y = sig_clf.predict_proba(train_variation_feature_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is", log_loss(y_train, predict_y))
```

```
s:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_variation_feature_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation
    log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_variation_feature_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss i
s:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

For values of alpha = 1e-05 The log loss is: 1.7296891861140264
 For values of alpha = 0.0001 The log loss is: 1.726508239164853
 For values of alpha = 0.001 The log loss is: 1.7305767389405462
 For values of alpha = 0.01 The log loss is: 1.7384772694258812
 For values of alpha = 0.1 The log loss is: 1.752594443850392
 For values of alpha = 1 The log loss is: 1.753358780697446



For values of best alpha = 0.0001 The train log loss is: 0.716214807089609
 For values of best alpha = 0.0001 The cross validation log loss is: 1.726508239164853
 For values of best alpha = 0.0001 The test log loss is: 1.6956866669203974

Q11. Is the Variation feature stable across all the data sets (Test, Train, Cross validation)?

Ans. Not sure! But lets be very sure using the below analysis.

```
In [45]: print("Q12. How many data points are covered by total ", unique_variations.shape[0], " genes in test and cross validation data sets?")
test_coverage=test_df[test_df['Variation'].isin(list(set(train_df['Variation'])))].shape[0]
cv_coverage=cv_df[cv_df['Variation'].isin(list(set(train_df['Variation'])))].shape[0]
print('Ans\n1. In test data',test_coverage, 'out of',test_df.shape[0],":", (test_coverage/test_df.shape[0])*100)
print('2. In cross validation data',cv_coverage, 'out of ',cv_df.shape[0],":", (cv_coverage/cv_df.shape[0])*100)
```

Q12. How many data points are covered by total 1921 genes in test and cross validation data sets?

Ans

1. In test data 72 out of 665 : 10.827067669172932
2. In cross validation data 46 out of 532 : 8.646616541353383

```
In [46]: alpha
```

```
Out[46]: [1e-05, 0.0001, 0.001, 0.01, 0.1, 1]
```

3.2.3 Univariate Analysis on Text Feature

1. How many unique words are present in train data?
2. How are word frequencies distributed?
3. How to featurize text field?
4. Is the text feature useful in predicting y_i?
5. Is the text feature stable across train, test and CV datasets?

```
In [47]: # cls_text is a data frame
# for every row in data fram consider the 'TEXT'
# split the words by space
# make a dict with those words
# increment its count whenever we see that word

def extract_dictionary_paddle(cls_text):
    dictionary = defaultdict(int)
    for index, row in cls_text.iterrows():
        for word in row['TEXT'].split():
            dictionary[word] +=1
    return dictionary
```

```
In [48]: import math
#https://stackoverflow.com/a/1602964
def get_text_responseCoding(df):
    text_feature_responseCoding = np.zeros((df.shape[0],9))
    for i in range(0,9):
        row_index = 0
        for index, row in df.iterrows():
            sum_prob = 0
            for word in row['TEXT'].split():
                sum_prob += math.log((dict_list[i].get(word,0)+10)/(total_dict.get(word,0)+90))
            text_feature_responseCoding[row_index][i] = math.exp(sum_prob/len(row['TEXT'].split()))
            row_index += 1
    return text_feature_responseCoding
```

```
In [49]: # building a CountVectorizer with all the words that occurred minimum 3 times in train data
text_vectorizer = CountVectorizer(min_df=3)
train_text_feature_onehotCoding = text_vectorizer.fit_transform(train_df['TEXT'])
# getting all the feature names (words)
train_text_features= text_vectorizer.get_feature_names()

# train_text_feature_onehotCoding.sum(axis=0).A1 will sum every row and returns (1*number of features) vector
train_text_fea_counts = train_text_feature_onehotCoding.sum(axis=0).A1

# zip(list(text_features),text_fea_counts) will zip a word with its number of times it occurred
text_fea_dict = dict(zip(list(train_text_features),train_text_fea_counts))

print("Total number of unique words in train data :", len(train_text_features))
```

Total number of unique words in train data : 53213

```
In [ ]: #####
```

```
In [2]: train_text_features= text_vectorizer.vocabulary_.keys()
```

In [5]: train_text_features

Out[5]: dict_keys(['recent', 'results', 'clinical', 'trials', 'braf', 'inhibitors', 'gsk2118436', 'dabrafenib', 'plx4032', 'vemurafenib', 'shown', 'encouraging', 'response', 'rates', 'however', 'duration', 'limited', 'identify', 'determinants', 'acquired', 'resistance', 'strategies', 'overcome', 'isolated', 'drug', 'resistant', 'clones', 'a375', 'brafv600e', 'brafv600k', 'melanoma', 'cell', 'lines', 'also', 'showed', 'reduced', 'sensitivity', 'allosteric', 'mitogen', 'activated', 'protein', 'extracellular', 'signal', 'regulated', 'kinase', 'mek', 'inhibitor', 'gsk1120212', 'trametinib', 'genetic', 'characterization', 'identified', 'frame', 'deletion', 'mek1', 'nras', 'mutation', 'nrasq61k', 'without', 'background', 'stable', 'knockdown', 'short', 'hairpin', 'rna', 'partially', 'restored', 'mutant', 'whereas', 'expression', 'parental', 'cells', 'decreased', 'similarly', 'combination', 'effectively', 'inhibited', 'growth', 'erk', 'phosphorylation', 'cyclin', 'd1', 'increased', 'p27kip1', 'moreover', 'phosphoinositide', 'mtor', 'gsk2126458', 'enhanced', 'inhibition', 's6', 'ribosomal', 'show', 'mutations', 'contribute', 'vitro', 'overcomes', 'addition', 'respond', 'ongoing', 'planned', 'test', 'combinations', 'mol', 'cancer', 'ther', '11', '909', '20', '2012', 'aacr', 'article', 'featured', 'highlights', 'issue', '793', 'introduction', 'highly', 'aggressive', 'form', 'skin', 'effective', 'treatment', 'identification', 'frequent', '50', 'valine', 'position', '600', 'isoform', 'raf', 'brafv600', 'held', 'great', 'promise', 'therapeutic', 'intervention', 'agents', 'target', 'impressive', 'responses', 'patients', 'food', 'administration', 'approved', 'positive', 'inoperable', 'metastatic', 'develops', 'within', 'months', 'following', 'initiation', 'observations', 'make', 'imperative', 'determine', 'factors', 'promote', 'potential', 'therapies', 'development', 'tumor', 'single', 'targeted', 'inevitable', 'gatekeeper', 'one', 'mechanism', 'proposed', 'additional', 'yet', 'observe', 'clinic', '10', 'instead', 'genes', 'seems', 'reestablish', 'ras', 'signaling', 'tumors', 'exposed', 'activation', 'platelet', 'derived', 'factor', 'pdgf', 'receptors', 'pdgfr', 'ref', 'igf1r', 'akt3', '12', 'cot1', '13', 'either', 'preclinical', 'models', 'conferring', 'plx4720', 'sb590885', 'potent', 'atp', 'competitive', 'wild', 'type', 'craf', 'enzymes', '14', 'selective', 'mek2', '15', 'inhibits', 'king', 'colleagues', 'submitted', 'compounds', 'phase', 'iii', 'nct01245062', 'lead', 'sustained', 'presence', 'compound', 'may', 'enhance', 'deter', 'outgrowth', 'inhibiting', 'pathway', 'separate', 'points', 'tested', 'ii', '16', 'members', 'kinases', 'pi3k', 'elevated', '17', 'reduces', 'model', 'systems', '18', 'pi3', 'family', 'well', '19', 'current', 'trial', 'solid', 'transduction', 'feedback', 'mediate', 'contributes', 'activity', 'although', 'data', 'associated', 'emerged', 'exist', 'efforts', 'evaluate', 'samples', 'collected', 'progression', 'upon', 'parallel', 'established', 'proliferate', 'findings', 'system', 'guide', 'evaluation', 'obtained', 'provide', 'scientific', 'insights', 'effectiveness', 'materials', 'methods', 'chemicals', 'gsk1120212b', '21', 'synthesized', 'glaxosmithkline', 'research', 'labs', 'structures', 'fig', '1a', 'figure', 'download', 'figureopen', 'new', 'tabdownload', 'powerpoint', 'relative', 'day', 'western', 'blot', 'analysis', 'selected', '24', 'hour', 'gapdh', 'glyceraldehyde', 'phosphate', 'dehydrogenase', 'clonal', 'isolate', 'line', 'american', 'culture', 'collection', '96', 'similar', 'genome', 'project', 'reference', 'wellcome', 'trust', 'sanger', 'institute', 'nucleotide', 'polymorphism', 'snp', '22', 'yale', 'dermatology', 'facility', 'characterized', 'exome', 'sequencing', 'cultured', 'less', 'gsk', 'referred', 'henceforth', 'increasing', 'concentrations', 'maintained', 'final', 'concentration', 'limiting', 'dilution', 'populations', '95', 'determined', 'chip', 'representative', 'shared', '100', 'snps', 'grown', 'rpmi', '1640', 'medium', 'containing', 'fbs', 'least', 'passage', 'experiment', 'dna', 'extraction', 'described', 'supplementary', 'proteins', 'extracted', 'lysis', 'buffer', 'fifty', 'micrograms', 'soluble', 'analyzed', 'using', 'indicated', 'antibody', 'appropriate', 'secondary', 'listed', 'odyssey', 'infrared', 'imaging', 'li', 'cor', 'biosciences', 'apoptosis', 'assays', 'estim'])

ated', 'celltiter', 'glo', 'promega', 'previously', '23', 'caspase', 'via', 'assay', 'hours', 'transfected', 'transduced', 'interfering', 'sirna', 'shrn a', 'constructs', 'longer', 'term', 'proliferation', 'plated', 'treated', 'rm pi', 'days', 'treatments', 'replaced', 'stained', 'methylene', 'blue', 'ethan ol', 'images', 'captured', 'flatbed', 'scanner', 'construction', 'modeled', 'crystal', 'structure', 'affymetrix', 'analyses', 'comparison', 'alone', 'don e', 'microarray', 'deposited', 'ncbi', 'gene', 'omnibus', 'geo', 'http', 'ww w', 'nlm', 'nih', 'gov', 'accessible', 'series', 'accession', 'number', 'derivation', 'establish', 'grow', 'multiple', 'ic50', '1b', 'displayed', 'greate r', 'fold', 'compared', 'table', '38', '002', '005', 's1', 'view', 'inlinevie w', 'popup', 'heterozygous', 'q61', 'q61k', 'a146', 'a146t', 'concomitant', 'remaining', 'contained', 'k59', 'confirmed', 'retention', 'homozygous', 'kra s', 'hras', 'araf', 'pten', 'pik3ca', 'retained', 'synonymous', 'codon', 'lik ewise', 'map', 'four', '1c', 'absence', 'high', 'levels', 's6p', 'low', 'phosphorylated', 'akt', 't308', 'potently', '75', '65', 'effects', 'attenuated', '30', '70', '25', 'respectively', 'unlike', 'relatively', 'unchanged', 'subse t', 'next', 'evaluated', 'effect', 'coprecipitation', 's2', 'reduction', '2 a', 's3', 'ability', 'decrease', '2b', 'partial', 'restoration', 'control', 'targeting', 'clone', 'antibodies', 'express', 'flag', '01', '212', 'arrow', 'expected', 'confirm', 'plasmids', 'encoding', 'neomycin', 'selection', 'g41 8', 'disulfate', 'selects', 'resulted', 'expressing', 'mock', 'transfection s', 'failed', 'colonies', 'insensitive', '2c', 's4', 'reduce', '2d', 'conde r', 'manner', 'whether', 'found', 'activate', 'amino', 'acid', 'residue', 'located', 'adjacent', 'negative', 'regulatory', 'helix', 'preceding', 'core', 'domain', 'hypothesized', 'stabilize', 'inactive', 'conformation', 'binds', '3a', 'q56', 'agent', 'azd6244', 'largely', 'disordered', 'terminus', 't386', 'involved', 'regulation', 'front', 'top', 'views', 'showing', 'positions', 'yellow', 'adp', 'viruses', 'vector', 'gfp', 'wt', 'lysates', 'mediated', 'contain', '3b', 'inhibit', '126', 'controls', '024', '3c', 'consistent', 'highe r', '3d', 'induction', 'minimal', 'profound', 'still', 'basal', 'used', 'knock', 'sensitive', 'downregulation', '60', 'suggesting', 'dependence', '3e', 's 5', 'combined', 'would', 'beneficial', 'constant', 'molar', 'ratio', 'irrespective', 'index', 'ci', 'could', 'calculated', 'lack', 'excess', 'highest', 'parameter', 'increase', 'benefit', 'stably', 'expressed', 's7', 's8', 'micromolar', 's9', 'extended', '90', '4a', 'pronounced', 'required', '03', 'observe', 'slightly', 'synergistic', 'nearly', 'additive', 'modest', 'potency', 'staining', 'presented', 'lanes', 'heatmap', 'normalized', 'probe', 'sets', 'representing', 'differentially', 'vehicle', 'characterize', 'examined', '4b', 'sufficient', 'little', 'extent', 'varying', 'degrees', 'global', 'altered', 'select', 'related', 'survival', 'transcription', '4c', 'alter', 'transcripts', 'downregulated', 'ccnd1', 'cdc25a', 'pcna', 'myc', 'mcl1', 'bik', 'casp1', 'upregulated', 'dependent', 'dusp4', 'dusp5', 'etv1', 'etv4', 'foxc2', 'spry2', 'alternative', 'pathways', 'rhob', 'epithelial', 'mesenchymal', 'transition', 'snai2', 'profile', 'seen', 'mapk', 'enhances', 'profoundly', 'completely', 'induced', '5a', 'cleaved', 'parp', 'indicators', 'harboring', 'regardless', 'long', 'enhancement', '5b', 'general', 'appreciable', 'antiproliferative', 'benefits', 'moderately', 'approximately', 'discussion', 'suggest', 'confers', 'thus', 'mutational', 'maintain', 'even', 'suppresses', 'thereby', 'reducing', 'responsive', 'mutated', 'patient', '26', 'common', 'substitution', 'maintains', 'active', 'state', 'impairing', 'gtp', 'hydrolysis', 'reported', 'lymphoblastic', 'leukemia', '27', 'activates', 'promotes', 'transformation', '3t3', 'lesser', '28', 'reports', 'indicate', 'heterodimerization', 'activating', '29', '31', 'complex', 'leading', '32', 'co', 'occurring', 'sample', 'furthermore', 'support', 'hypothesis', 'decreasing', 'unlikely', 'site', 'know n', 'affect', 'binding', 'directly', 'destabilize', 'confirmation', 'instance', 'unphosphorylated', 'nmol', 'dramatically', 'shifted', 'cardio', 'facio', 'cutaneous', 'syndrome', 'result', 'constitutively', '33', '35', 'mutagenesi

s', 'screen', 'among', 'others', 'conferred', 'lung', 'adenocarcinoma', '36', 'progressing', 'mek1c121s', 'outside', 'cause', 'mek1f1291', 'interaction', 'intrinsic', 'circumventing', '37', 'studies', 'necessary', 'frequency', 'mek2q60p', 'pretreatment', 'disease', 'hyperactive', 'insufficiently', 'led', 'a rrest', 'transcriptional', 'output', 'included', 'designated', 'dusp6', 'etv5', '39', 'clinically', 'relevant', 'parent', 'tolerated', 'including', 'progressed', 'incidence', 'rash', 'hyperproliferative', 'lesions', '40', 'responded', 'albeit', 'combining', 'represent', 'alternatives', 'bez235', '41', 'evaluating', 'progress', 'nct01072175', 'naive', 'utilizing', 'yeast', 'saccharomyces', 'cerevisiae', 'novel', 'autoactivation', 'region', 'mammalian', 'specific', 'pd', '184352', 'possible', 'due', 'homology', 'components', 'pheromone', 'eukaryotic', 'cascade', 'fus1', 'his3', 'reporter', 'functional', 'readout', 'reconstituted', 'randomly', 'mutagenized', 'variants', 'seven', 'base', 'change', 'five', 'mapped', 'subdomains', 'iv', 'leucine', 'proline', '115', 'leu115pro', 'inhibitory', 'mutants', 'strikingly', 'overexpression', 'hek293t', 'finding', 'stimulated', 'phospho', 'cyclic', 'amp', '1cdk', 'template', 'generated', 'residues', 'clustered', 'together', 'forming', 'hydrophobic', 'pocket', 'additionally', 'allowed', 'interact', 'directed', 'supported', 'involvement', 'comprised', 'evolutionarily', 'conserved', 'module', 'regulates', 'differentiation', 'movement', 'stimulation', 'reviewed', 'key', 'component', 'dual', 'specificity', 'phosphorylates', 'threonine', 'tyrosine', 'induces', 'host', 'downstream', 'cellular', 'subject', 'two', 'serine', 'ser218', 'ser222', 'lie', 'loop', 'vii', 'viii', 'negatively', 'charged', 'acids', 'aspartate', 'glutamate', 'mimics', 'modification', 'presumably', 'stabilization', 'allowing', 'enzyme', 'retain', 'another', 'feature', 'rich', 'carboxy', 'terminal', 'appears', 'critical', 'association', 'play', 'important', 'role', 'efficiently', 'polyproline', 'contains', 'sites', 'regulate', 'proximal', 'catalytic', '67', 'regulating', 'putative', 'docking', 'first', 'particular', 'sequences', 'positively', 'extreme', 'lys3', 'mkk1', 'mkk7', 'considered', 'essential', 'termini', 'recognition', 'sequence', 'anthrax', 'lethal', 'proteolytic', 'bacillus', 'anthracis', 'virulence', 'cleaves', 'impairs', 'enzymatic', 'vivo', 'finally', '44', '51', 'since', '80', 'coupled', 'substitutions', 'ser218asp', 'ser222asp', '640', 'small', 'molecule', 'capable', 'human', 'murine', 'colon', 'carcinomas', 'mice', 'noncompetitive', 'block', 'mutationally', 'present', 'study', 'order', 'gain', 'insight', 'molecular', 'devised', 'approach', 'budding', 'report', 'identity', 'structural', 'motif', 'interacts', 'describe', 'overall', 'go', 'vectors', 'prs314', 'trp1', 'cen6', 'arsh4', 'prs425', 'leu2', 'modified', 'insertion', 'kb', 'fragment', 'gal1', 'promoter', 'pbj247', 'genbank', 'nm002880', 'rendered', '302', 'pcr', 'injected', 'bamhi', 'ssti', 'generating', 'prs314galraf', 'nt', 'cdna', 'l11284', 'underlined', 'kozak', 'italicized', 'changes', 'trna', 'bias', 'boldface', 'oligonucleotide', 'ggatcccacacataaataaaatgccaaagaaggcaacgc当地', 'corresponding', 'epitope', 'followed', 'added', 'aagctttacttgtcatcgccgttagtc当地gacgccagcagcatgg', 'product', 'prs425galmek', 'recombinant', 'bacteria', 'subcloned', 'nde1', 'hindiii', 'pet21b', 'plasmid', 'novagen', 'cn', 'quikchange', 'kit', 'stratagene', 'inc', 'generate', 'various', 'point', 'coding', 'transfection', 'pcmv4a', 'standard', 'techniques', '3acc', 'generation', 'ste11', 'ste7', 'ste5', 'strain', 'haploid', 'sy2002', 'matahis3', 'mfa2', 'lacZ', 'ade1', '112', 'dh1', 'ura3', '52', 'generous', 'gift', 'sprague', 'jr', 'protocols', 'disrupt', 'mating', 'unless', 'otherwise', 'noted', 'procedure', 'media', 'sherman', 'et', 'al', 'disruption', 'loci', 'nutrient', 'deficient', 'histidine', 'auxotrophy', 'complementation', 'disrupted', 'transforming', 'restoring', 'propagate', 'lacking', 'random', 'escherichia', 'coli', 'x1', 'red', 'deficiency', 'three', 'primary', 'repair', 'muts', 'mutd', 'mutt', 'rate', '000', 'prs425galmek1', 'transformed', 'propagated', 'luria', 'bertani', 'agar', 'ampicillin', 'ml', 'accumulate', 'scraped', 'plates', 'transferred', 'liters', 'liquid', 'expand', 'harvested', 'centrifugation', 'prepared'

ed', 'reagents', 'qiagen', 'library', '300', 'lithium', 'acetate', 'method', 'ito', 'onto', 'synthetic', 'tryptophan', 'sd', 'trp', 'leu', 'nitrogen', 'yn b', 'supplemented', 'galactose', 'induce', 'propagating', 'prototrophs', 'population', '275', 'restreaked', 'plus', 'bio', '101', 'rescued', 'kc8', 'clont ech', 'laboratories', 'automated', 'sequencer', 'perkin', 'elmer', 'biosystems', 'purification', 'his6', 'tagged', 'fusion', 'b121', 'de3', 'bacterial', 'iptg', 'isopropyl', 'thiogalactopyranoside', '200', 'purified', 'imac', 'mass', 'sodium', 'dodecyl', 'sulfate', 'polyacrylamide', 'gel', 'electrophoresis', 'sds', 'page', 'coomassie', 'activities', 'glutathione', 'transferase', 'gst', 'erk1', 'gsterk1k71r', 'substrate', 'ng', 'incubated', 'mm', 'hepes', 'ph', 'mgcl2', 'egta', '32p', 'total', 'volume', 'upstate', 'biotechnology', 'reaction', 'initiated', 'gsterk1', 'k71r', 'working', 'linear', 'range', 'min', 'quenched', 'laemmli', 'resolved', 'phosphoproteins', 'visualized', 'autoradiography', 'radiolabel', 'incorporated', 'excision', 'band', 'quantitate d', 'scintillation', 'counter', 'beckman', 'instruments', '3h', 'proximity', 'spa', 'utilized', 'direct', 'custom', 'amersham', 'mmol', 'k2hpo4', 'kh2po 4', 'nacl', 'chaps', 'cholamidopropyl', 'dimethylammonio', 'propanesulfonate', 'glycerol', 'flat', 'bottom', 'tissue', 'shaken', 'room', 'temperature', 'ysi', 'copper', 'tag', 'beads', 'mg', 'bovine', 'immunoglobulin', 'igg', 'sigma', '45', 'amount', 'bound', 'bead', 'wallac', 'microbeta', 'plate', 'kbab', 'dulbecco', 'eagle', 'newborn', 'calf', 'serum', 'glutamax', 'life', 'technologies', 'deprived', 'glycerophosphate', 'triton', 'aldrich', 'phenylmethy lsulfonyl', 'fluoride', 'vanadate', 'leupeptin', 'clarified', 'bca', 'pierce', 'perbio', 'extracts', 'subjected', 'nitrocellulose', 'probed', 'anti', 'england', 'biolabs', 'corp', 'p44mapk1', 'p42mapk2', 'santa', 'cruz', 'immunoreactive', 'ecl', 'international', 'plc', 'fetal', 'transient', 'performed', 'lipofectamine', 'posttransfection', 'blotting', 'm2', 'modeling', 'constructed', 'package', 'look', 'applications', 'group', 'insightii', 'msi', 'prior', 'internal', '268', '308', 'pdb', 'based', 'alignment', 'minimized', 'discover', 'procheck', 'profile3d', 'modules', 'occurs', 'blocking', 'assayed', 'phosphorylate', 'gstmekk97r', 'measurement', '218', '222', 'alanine', 'detectible', 'unable', 'whole', 'recognizing', 'latter', 'recognizes', 'unexpectedly', 'abolished', 'level', 'actually', 'dose', 'caused', 'threefold', 'agonist', 'time', 'hyperphosphorylation', 'merely', 'artifact', 'colon26', 'h61', 'fibroblast', 'explanation', 'existence', 'loops', 'whereby', 'leads', 'hyperactivation', 'upstream', 'dependency', 'accumulation', 'distinct', 'clear', 'exerts', 'unexpected', 'interacting', 'allows', 'homologues', 'scaffolding', 'integrated', 'locus', 'auxotrophic', 'intact', 'member', 'renders', 'incapable', 'growing', 'counterparts', 'specifically', 'integrity', 'prototrophic', 'accomplished', 'placing', 'inducible', 'blocked', 'toxic', 'able', 'demonstrates', 'homologue', 'grew', 'normally', 'isolation', 'defective', 'major', 'mechanisms', 'subsequent', 'pool', 'forms', 'premature', 'stop', 'codons', 'removed', 'inability', 'complement', 'return', 'prototrophy', 'transformants', 'screened', 'different', 'alleles', 'pair', 'summarized', 'schematically', 'interestingly', 'spanning', 'confirming', 'spontaneous', 'illustration', 'indicating', 'native', 'marked', 'asterisk', 'biochemical', 'properties', 'recreated', 'construct', 'immobilized', 'metal', 'affinity', 'chromatography', 'surprisingly', 'exception', 'phe53ser', 'glu203lys', 'values', 'table11', 'variant', 'substantially', 'ranged', '528', '895', 'pmol', 'much', 'times', 'significantly', 'likely', 'escape', 'despite', 'exhibited', 'represented', 'hatched', 'bars', '98059', 'disclosed', 'previous', 'work', 's218', 's222', 'render', 'several', 'except', 'phe53', 'glu203', 'cluster', 'part', 'interface', 'stranded', 'sheet', 'lobe', 'rear', 'face', 'suggests', 'accommodate', 'size', 'lys97', 'glu114', 'hypothetical', 'salt', 'bridge', 'wall', 'separating', 'formed', 'clustering', 'apparent', 'complete', 'side', 'chain', 'leu115', 'hence', 'altering', 'prevent', 'substituted', 'arginine', 'chains', 'conservative', 'leu115ala', 'quite', '68', 'versus', '067', 'resulting', '197',

'478', '882', '581', '739', '1115r', '1115p', '1115w', '1115e', '1115a', '121', 'illustrating', 'ribbon', 'diagram', 'viewed', 'upper', 'suggested', 'migt', 'isoleucine111', 'ile111', 'leucine118', 'leu118', 'isoleucine139', 'ile139', 'appeared', 'comprise', 'subsequently', 'table2', 'somewhat', 'exceede d', 'lining', 'affects', 'reflective', 'perturbations', 'examine', 'figure7a7a', 'labeled', 'correlates', 'detected', 'ile111pro', 'abrogated', '7b', 'tables', 'tables11', 'and2', 'contrast', 'roughly', 'conditions', 'argument', 'assess', '500', 'nm', 'endogenous', 'possessed', 'cloned', 'carboxyl', 'transiently', 'purposes', 'ser218glu', 'ser222glu', '2e', 'starved', 'comparable', 'exceeding', 'cotransfected', 'rasv12', 'ile103asn', 'agreement', 'plays', 'revealed', 'hindered', 'compete', 'interesting', 'harbors', 'lysine', 'phosphoryl', 'transfer', 'thought', 'stabilizing', 'compromise', 'interfere', 'bind', 'fact', 'substantial', 'falls', 'usual', 'significant', 'reasonably', 'effort', 'close', 'comprising', 'ahn', 'acting', 'markedly', 'clearly', 'exerted', 'greatest', 'inside', 'favored', 'helices', 'amide', 'hydrogen', 'bond', 'donor', 'steric', 'constraints', 'imposed', 'helical', 'lend', 'localized', 'ray', 'involves', 'conformational', 'lip', 'helps', 'second', 'lobes', 'rotate', 'respect', 'cleft', 'bringing', 'correct', 'catalysis', 'subtle', 'dynam ic', 'motion', 'involving', 'deuterium', 'exchange', 'flexibility', 'undergo', 'segment', 'believe', 'nature', 'implications', 'introduced', 'loss', 'ordering', 'possibly', 'autophosphorylation', 'indeed', 'preliminary', 'experiments', 'conversely', 'imposition', 'impact', 'lock', 'preventing', 'phosphatas es', 'explaining', 'u0126', 'proven', 'exceptionally', 'mkk', 'mkk3', 'mkk4', 'mkk6', 'closely', 'recently', 'weaker', 'mkk5', 'exhibits', 'closest', 'ca', '46', 'alignments', 'reveal', 'chosen', 'identical', 'diverge', 'exists', 'enough', 'recognized', 'correlate', 'overexpressed', 'hyperphosphorylate d', 'tempting', 'speculate', 'investigation', 'unknown', 'fully', 'understan d', 'gaining', 'importance', 'wide', 'cancers', 'extensively', 'documented', 'bcr', 'abl', 'oncogene', 'imatinib', 'leukaemia', 'mechanistic', 'exploited', 'drugs', 'override', 'created', 'rationale', 'multi', 'cocktails', 'antagonists', 'antitumour', 'combine', 'risk', 'emergence', 'glance', 'superfamily', 'aspects', 'neoplasia', 'tumour', 'migration', 'survival1', 'deregu late', 'alterations', 'early', 'tumorigenesis', 'remains', 'aspect', 'physiology', 'throughout', 'progression2', 'therefore', 'attractive', 'strategy', 'treatment3', 'promises', 'improved', 'efficacy', 'anticancer', 'inactivation', 'often', 'pursued', 'substrates', 'creates', 'pressures', 'analogous', 'ordinary', 'natural', 'evolve', 'widespread', 'studied', 'formation', 'exposure', 'antimicrobial', 'agents6', 'take', 'variety', 'routes', 'include', 'producing', 'substituting', 'function', 'upregulating', 'alternate', 'transporters', 'efflux', 'particularly', 'therapy', 'reviews', 'dealt', 'topic8', 'review', 'discuss', 'countermeasures', 'address', 'emerging', 'problem', 'besets', 'lessons', 'prototype', 'phenylaminopyrimidine', 'mesylate', 'gleevec', 'glive c', 'sti571', 'novartis', 'developed', 'receptor', 'demonstrated', 'remarkable', 'malignancies', 'chronic', 'myeloid', 'cml', 'cases', 'pathogenesis', 'co nsequence', 'defined', 'translocation', 'event', 'philadelphia', 'chromosome', 'breakpoint', 'abelson', 'oncogene14', 'constitutive', 'encoded', 'crucial', 'malignant', 'represents', 'intervention16', 'chromosomal', 'adult', 'acute', 'lymphoma', 'patients20', 'durable', 'haematological', 'cytogenetic', 'free', 'advanced', 'states', 'blast', 'crisis', 'refractory', 'relapsed', 'continued', 'therapy21', 'fuelled', 'years', 'explore', 'causative', 'underlying', 'importantly', 'class', 'inherent', 'basis', 'attributed', 'delivery', 'sequestration', 'alpha1', 'glycoprotein', 'plasma', 'transport', 'processes', 'reaching', 'intracellular', 'abl23', 'unfortunately', 'text', 'require', 'assistance', 'access', 'image', 'obtain', 'description', 'please', 'contact', 'npg', 'com', 'majority', 'reactivation', 'predominantly', 'implicated', 'amplification', 'compensatory', 'switch', 'signalling', 'compensate', 'leukaemic', 'myelogenous', 'depending', 'quiescent', 'stem', 'intrinsically', 'pharma

cokinetic', 'actively', 'transported', 'pump', 'keep', 'availability', 'resolution', 'legend', 'figures', 'compensation', 'src', 'lyn', 'resistance26', 'relevance', 'mechanism26', 'alternatively', 'actual', 'whose', 'amplification28', 'halting', 'stages', 'seem', 'eradicate', 'mature', 'differentiated', 'progenitors', 'tolerate', 'forced', 'imatinib29', 'reappearance', 'cessation', 'account', 'per', 'se', 'nevertheless', 'necessitates', 'continuation', 'was', 'domain28', '34', 'numerous', 'far', 'many', 'verified', 'focus', 'desensitizing', 'prevalent', 'frequently', 'encode', 'features', 'location', 'hotspots', 'comprehensive', 'lists', 'compiled', 'articles9', 'ndash', 'green', 't hr315', 'highlighted', 'tyr253', 'glu255', 'affected', 'harbouring', 'his396', 'orange', 'met351', 'm351t', 'pink', 'closer', 'surface', 'representation', 'perspective', 'cover', 'regions', 'chemical', 'hydroxyl', 'called', '315', 'sterically', 't315i', 'interferes', 'large', 'isoleucine', 'position28', 'case', 'depend', 'accessibility', 'cavity', 'incorporation', 'bulky', 'individual', 'requirements', 'mean', 'promoting', 'preserved', 'sensitivities', 'targets', 'diverse', 'chemotypes', 'determinant', 'discussed', 'detail', 'sections37', 'distinguished', 'requirement', 'selectively', 'distorted', 'fit', 'mechanism9', 'remarkably', 'towards', 'implies', 'equilibrium', 'trap', 'oncoprotein', 'transits', 'mode', 'potentially', 'achilles', 'heel', 'catalytically', 'occur', 'binding33', 'mentioned', 'poor', 'prognosis', 'y253f', 'e255k39', 'strongly', 'h396r', 'closed', 'counteract', 'remote', 'falling', 'accounts', 'resistance10', 'precise', 'positioning', 'imatinib9', 'difficult', 'predict', 'rather', 'focused', 'restricted', 'saturating', 'full', 'length', 'screening', 'procedure', 'selecting', 'driven', 'weaken', 'resistance40', 'probably', 'escaped', 'detection', 'earlier', 'mainly', 'generates', 'abnormality', 'linked', 'myeloproliferative', 'rearrangements', 'exclusive', 'create', 'pdgfralpha', 'pdgfrbeta', 'rare', 'haematologic', 'diseases', 'idiopathic', 'hypereosinophilic', 'myelomonocytic', 'respectively41', '42', 'excessive', 'eosinophils', 'chimeric', 'fip1', 'like', 'fip111', 'apparently', 'arises', 'pdgfra', 'interstitial', 'deletion42', 'haematopoietic', 'lower', 'suppress', 'proliferation42', 'cools', 'detectable', 'experienced', 'remission', 'eosinophil', 'count', 'normal', 'values42', 'eventually', 'equivalent', 'alpha', 'surrounding', 'e255k', 'context', 'already', 'possess', 'phenylalanine', 'stretch', 'includes', 'parts', 'asp816', 'desensitizes', 'tyr823', 'conservation', 'monotherapy', 'gastrointestinal', 'stromal', 'tumours', 'gist', 'habour', 'juxtamembrane', 'kinase43', 't670i', 'domain44', 'homologous', 'confined', 'lesion', 'responding', 'treatment44', 'exclusively', 'genomic', 'instability', 'drawback', 'abl19', 'according', 'note', 'tyr393', 'aspartate45', 'destabilizing', 'loop36', 'mimic', 'relapse', 'd816v', 'proto', 'rtk', 'mastocytosis46', 'adopting', 'modes', 'acquisition', 'oncogenic', 'screens', 'experimental', 'abl40', 'anticipate', 'become', 'reality', 'prompt', 'testing', 'back', 'provides', 'evidence', 'reconfirms', 'viable', 'distort', 'circumvented', 'developing', 'configuration', 'resembles', 'originally', '47', 'belonging', 'pyrido', 'pyrimidine', 'considerable', 'isoforms', 'inhibitors48', '49', 'pd180970', 'ap23464', 'bms', '354825', 'efficient', 'pyrimidines', 'h396p', 'tenfold', 'higher48', 'glycine', 'requires', 'fivefold', 'non', 'clinic39', 'inactivated', 'abl48', 'half', 'maximal', 'amenable', 'escalation', 'patient51', 'promising', 'interest', 'structurally', '53', 'abilities', 'allele', 'mouse', 'cml52', 'doses', 'animal', 'none', 'indicates', 'stringent', 'characteristic', 'selectivity', 'state36', 'adverse', 'favourable', 'toxicity', 'profiles', 'neither', 'concern', 'considering', 'classes', 'hotspot', 'formation37', '54', 'example', 'larger', 'methionine', 'p38', 'epidermal', 'egfr', 'types', 'sb203580', 'pp1', 'pd153035', 'respectively55', '56', '57', 'aromatic', 'substituents', 'accommodated', 'illustrated', 'erlotinib', 'thr766', 'projection', 'resistance58', '5c', 'erlotinib58', 'antagonist', 'belongs', 'gefitinib', 'alphac', 'orientation', 'presentation', 'hinge', 'clash', 'moiety', '766', 'abrogate', 'kinase35', 'iressa', 'zd1839', 'astrazeneca', 'nsclc

s', 'chemotherapy', 'regimens59', 't351i', 'replacing', 'residue37', 'light', 'strong', 'correlation', 'objective', 'nsclc', 'patients61', '62', '63', '64', 'perhaps', 'way', 'develop', 'element', 't674i', 'staurosporine', 'analog ue', 'pkc412', 'disease65', 't681i', 'indolinone', 'su6668', 'desensitizatio n', 'imatinib37', 'scaffold', 'fgfr1', 'position37', '66', 'relate', 'desig n', 'composition', 'egfr54', 'kinome', 'oncology', 'insulin', 'flt3', 'exten d', 'obviously', 'pressure', 'survive', 'basic', 'principle', 'true', 'exempl ified', 'derives', 'alteration', 'available', 'use', 'guided', 'diagnostic', 'tools', 'polymerase', 'aso', 'mutations67', 'late', 'stage', 'subpopulatio n', 'treatment67', '69', 'approaches', 'find', 'tailoring', 'mutate', 'hittin g', 'cells70', 'simultaneous', 'endothelial', 'neovascularization', 'aims', 'angiogenic', 'therapy71', 'achieved', 'medicines', 'promiscuous', 'act', 'se t', 'proteins72', 'share', 'concept', 'poly', 'pharmacology', 'shift', 'su541 6', 'monospecific', 'vascular', 'vegfr', 'angiogenesis73', 'follow', 'pharmac ological', 'simultaneously', 'rtks', 'vegfr2', 'fgfr174', 'fgfr', 'hardly', 'seemed', 'angiogenesis', 'inhibitor74', 'attain', 'balanced', 'certainly', 'relates', 'respective', 'profiles76', 'su11248', 'illustrates', 'possesses', 'flt377', '78', '79', 'shows', 'translate', 'neoplasias', 'renal', 'carcinoma 71', 'imatinib71', 'indolinones', 'perivascular', 'vascularization', 'blood', 'supply', 'tumour80', 'clarify', 'efficacious', 'terms', 'thwarting', 'runs', 'becoming', 'unselective', 'minimum', 'must', 'ideally', 'ineffective', 'chal lenge', 'thorough', 'candidates', 'undertaken', 'throughput', 'collections', 'parallel81', '82', 'addressed', 'proteomic', 'spectrometry', 'identification 83', '84', '85', 'carried', 'belonged', 'groups', 'fairly', 'unrelated', 'se r', 'thr', 'roles', 'cycle', 'apoptotic', 'daub', 'unpublished', 'counterintu itive', 'hit', 'carcinogenesis', 'notably', 'physiological', 'outcome', 'nece ssarily', 'reflect', 'sum', 'contributions', 'signalling86', 'her2', 'regress ion', 'aneuploidy', 'restores', 'phenotype', 'ovarian', 'carcinoma', 'ablatio n', 'alone87', 'certain', 'undergoing', 'evaluation88', 'conventional', 'over coming', 'combinatorial', 'offer', 'fundamentally', 'biological', 'action', 'administered', 'ratios', 'variable', 'schedules', 'optimal', 'intensive', 'traditional', 'yielded', 'cytotoxic', 'disorder89', 'idea', 'farnesyltransferase', 'heat', 'shock', 'domain90', '91', 'elements', 'future', 'overriding', 'conclusions', 'poses', 'successful', 'process', 'confirms', 'validity', 'uti lization', 'consequently', 'permit', 'generally', 'susceptible', 'definitio n', 'phenotypes', 'conceivable', 'complementary', 'desensitized', 'ensures', 'emerge', 'scenario', 'always', 'therapeutically', 'reminiscent', 'viral', 'h iv', 'infection', 'minimize', 'resistance92', 'hopefully', 'implementation', 'thwart', 'balance', 'eradication', 'last', 'hyperactivated', 'drivers', 'prompted', 'search', 'tractable', 'spectacular', 'driving', 'discovery', 'progra ms', 'aimed', 'pathologies', 'humans', 'cytokine', 'hormone', 'third', 'multi tude', 'cytosolic', 'nuclear', 'senescence', 'magnitude', 'behavior', 'carefu lly', 'ensure', 'appropriately', 'changing', 'environmental', 'outcomes', 'so matic', '43', 'thyroid', 'colorectal', 'glutamic', 'v600e', 'stimulate', 'fib roblasts', 'melanocytes', 'v600ebraf', 'molecules', 'depletion', 'interferenc e', 'validate', 'raised', 'deal', 'deadly', 'mortality', 'diagnosed', 'cure d', 'surgery', 'median', 'year', 'sought', 'urgency', 'enter', 'sorafenib', 'designed', 'dfg', 'proved', 'hepatocellular', 'demonstrating', 'valuable', 'highlight', 'pressing', 'problems', 'prove', 'mediates', 'antitumor', 'unexp lained', 'uncertainty', 'confusion', 'failure', 'suggestions', 'issues', 'out lined', 'aim', 'goal', 'raf265', 'panel', 'breast', '129', 'average', 'effici ency', 'increases', 'equally', '1d', '81', 'profiled', '1e', 'ck1', 'lck', 'nek2a', 'hipk2', 'aurora', 'erk8', 'res', 'open', 'tab', 'gi50', 'colo829', 'a375m', 'wm266', 'ht29', 'uacc62', 'skmel128', 'colo205', 'd04', 'skmel2', 'hct 116', 'wm1361', 'skmel23', 'chl', 'km12', 'd35', 'bt474', 'd24', 'percent', 'color', 'bar', 'lyte', 'delfia', '001', '1f', '043', '155', 'published', '00 02', '038', '140', 'value', 'considerably', '013', 'commercial', 'technolog

y', 'invitrogen', '134', '1g', 'intended', 'tease', 'apart', 'adenosine', 'triphosphate', 'thr529', 't529', 't529m', 't529i', 'asparagine', 't529n', 'retains', '87', '77', '480', 'equate', '420', '370', '48', 'assessed', '650', '250', '2f', 'display', 'differential', 'cos7', 'measured', 'immunoprecipitation', 'mbp', 'sequential', 'km', 'conducted', 'graphs', 'independent', 'duplicate', 'error', 'sem', 't529nbraf', 't529mbraf', 't529ibraf', 'embryonic', 'kidney', '293', 'hek293', 'pd184352', 'demonstrate', 'quantitative', 'fluorescent', 'loading', 'engineered', 'mbraf', 'pperk', 'erk2', 'star', 'nonspecific', 'blotted', 'almost', 'unaffected', 'hindrance', 'differences', 'difference', 'consequences', 'elaborate', 'channel', 'displacement', 'walls', 'right', 'hand', 'elaborates', 'silico', 'van', 'der', 'waals', 'forces', 'spheres', 'arrows', 'introduce', 'conformations', 'means', 'program', 'gold', 'faithfully', 'reproduced', 'analogs', 'relaxing', 'rotamers', 'newly', 'docked', 'adopt', 'rotamer', 'run', 'dock', 'superimpositions', 'pose', 'clashes', 'ba', 'f3', 'pro', 'absolutely', 'interleukin', 'il', 'suitable', 'cannot', '5d', 'empty', 'ppmek', 'counting', 'synthesis', '6a', '6b', '6c', '6d', 'thymidine', 'nonlinear', 'nude', 'bearing', 'daily', 'intraperitoneal', 'dosing', 'dmso', 'water', 'kg', 'sections', 'nuclei', 'counterstained', 'methyl', 'scale', 'scored', 'percentage', 'displaying', 'intensity', 'subcutaneous', 'volumes', 'every', 'six', 'animals', 'allografts', 'immunohistochemistry', 'delayed', 'failing', 'suppressed', 'investigated', 'intermediate', 'unclear', 'format', 'validating', 'exquisitely', 'making', 'interpret', 'past', 'decade', 'witnessed', 'enormous', 'dedicated', 'molecularly', 'hold', 'carry', 'formally', 'good', 'extensive', 'supporting', 'validation', 'confounds', 'judgment', 'arise', 'naturally', 'agc', 'tcc', 'gcc', 'disruptive', 'ccc', 'amounts', '1858regfr', 'weakly', 'causes', 'binders', 'accommodates', 'better', 'flexible', 'pack', 'tightly', 'gatekeepers', 'notion', 'difficulty', 'translating', 'critically', 'enabled', 'us', 'independently', 'robustly', 'dispel', 'suitability', 'relating', 'achieve', 'x1281', 'need', 'heterodimers', 'homodimers', 'drive', 'highlighting', 'investigate', 'questions', 'comparing', 'given', 'obvious', 'advantage', 'microenvironment', 'providing', 'tool', 'assessment', 'excellent', 'formal', 'proof', 'offers', 'powerful', 'acquire', 'bd', 'abcam', 'chir265', 'purchased', 'symansis', 'dissolved', 'dimethyl', 'sulfoxide', 'baculoviruses', 'baculovirus', 'infected', 'sf9', '106', '72', 'ni', 'nta', 'chromatography', 'measure', 'graphpad', 'prism', 'software', 'g12vras', 'dideoxy', 'blots', 'applied', 'gels', 'transfecting', 'nucleofector', 'amaxa', 'amplified', 'myelin', 'measuring', 'tritium', '104', 'seeded', 'wells', 'desired', '08', 'ge', 'healthcare', 'multiscrreen', 'glass', 'fiber', 'millipore', 'washed', 'twice', 'buffered', 'saline', 'methanol', 'vacuum', 'manifold', 'microscint', 'topcount', 'nxt', 'quadruplicate', 'sulforhodamine', 'bank', 'code', '1uwj', '3c4c', '2fb8', 'removing', 'ligand', 'moe', '2006', 'computing', 'release', 'strains', 'energy', 'minimization', 'radius', 'amber99', 'force', 'field', 'explorer', 'utility', 'version', 'starting', 'uwj', 'torsion', 'angles', 'charges', 'ligands', 'charge', 'corina', 'tsar', 'geometries', 'optimized', 'cosmic', 'calculations', 'terminated', 'gradient', 'smaller', 'ten', 'solutions', 'best', 'stored', 'around', 'crystallography', 'injected', 'subcutaneously', 'v600eba', '107', 'matched', 'allocated', 'caliper', 'measurements', 'width', 'depth', 'formula', '5236', 'accordance', 'uk', 'home', 'office', 'regulations', 'coordinating', 'committee', 'guidelines', 'biomarker', 'reach', 'mm3', 'consecutive', 'fixed', 'neutral', 'formalin', 'overnight', 'quantify', 'recorded', 'section', 'score', 'blind', 'forkhead', 'box', 'a1', 'foxa1', 'modulates', 'transactivation', 'steroid', 'influence', 'responsiveness', 'prostate', 'parameters', 'antigen', 'psa', 'cohort', 'differ', 'benign', 'glands', 'peripheral', 'zone', 'metastases', 'castration', 'neoplastic', 'transitional', 'tissues', 'correlate', 'pt', 'gleason', 'scores', 'androgen', 'ar', 'estrogen', 'faster', 'tumorogenicity', 'ars', 'substantiating', 'cooperation', 'conclusion', 'dedifferentiation'

tiation', 'poorer', 'inducing', 'men', 'countries', 'incompletely', 'understood', 'central', 'initially', 'crpc', 'tends', 'rapidly', 'manifestations', 'c hemotherapeutic', 'acceptance', 'oncological', 'practice', 'proportion', 'biology', 'independence', 'convert', 'paramount', 'prognostic', 'markers', 'herald', 'propensities', 'needed', 'array', 'transcript', 'rat', 'liver', 'hepatoma', 'hepg2', 'named', 'hepatocyte', 'hnf', 'drosophila', 'became', 'vertebrates', 'winged', 'dimensional', 'linker', 'histones', 'enables', 'fox', 'compacted', 'chromatin12', 'facilitate', 'researchers', 'give', 'name', 'pioneer', 'functionally', 'er', 'organs', 'respiratory', 'pancreas', 'urogenital', 'tract', 'bladder', 'promoters', 'metabolic', 'embryogenesis', 'organogenesis', 'organism', 'centrally', 'glucose', 'homeostasis', 'knockout', 'die', 'soon', 'birth', 'hypoglycemia', 'broad', 'spectrum', 'functions', 'nonetheless', 'body', 'literature', 'esophagus', 'barrett', 'adenocarcinomas', 'favorable', 'l uminal', 'date', 'significance', 'correlations', 'complemented', 'unravel', 'cancerous', '324', 'origin', 'continuous', 'shorter', 'lncap', 'initial', '5 29', 'underwent', 'radical', 'prostatectomy', 'transurethral', 'resection', 'organ', 'metastasis', '1993', 'department', 'urology', 'university', 'hospit al', 'zurich', 'prostatic', 'hyperplasia', 'bph', 'evaluable', '288', 'distribution', 'specimen', 'follows', 'specimens', '207', 'lymph', 'node', 'distan t', 'age', 'preoperative', '357', 'distributed', '119', 'categories', 'pt2', 'pt3', 'margins', 'r0', '131', 'r1', '92', 'rising', 'nadir', 'tma', 'resected', 'palliation', 'along', 'bone', 'trephine', 'archives', 'surgical', 'pathology', 'ethics', 'approval', 'stv', '2007', 'immunohistochemical', 'immunofluorescence', 'blocks', 'freshly', 'cut', 'thick', 'mounted', 'superfrost', 'slides', 'menzel', 'gl', 'braunschweig', 'germany', 'benchmark', 'ventana', 'medical', 'tucson', 'az', 'entire', 'hnf3', '400', 'diluent', 'biotechnologies', 'f39', 'biogenex', 'san', 'ramon', 'sp1', 'prediluted', 'albumin', 'tris', 'abd', 'serotec', 'oxford', '4000', 'dako', 'glostrup', 'denmark', 'ki', 'mib', 'sp4', 'thermo', 'fremont', 'p27', 'ultraview', 'dab', 'retrieval', 'hematoxylin', 'dehydrated', 'double', 'stainings', 'tmas', 'pathologists', 'microscope', 'intensities', 'cytoplasmic', 'tiered', 'weak', 'moderate', 'examples', 'see', 'supplemental', 'ajp', 'org', 'quantified', 'pc', 'du', '145', 'atcc', 'manassas', 'va', 'german', 'microorganisms', 'cultures', 'dsmz', 'cultivated', 'ham', 'kaighn', 'rwpe', 'keratinocyte', 'mixed', 'pituitary', 'extract', 'supplements', 'carlsbad', 'co2', 'humidity', 'real', 'rt', 'rneasy', 'hilden', 'applying', 'capacity', 'reverse', 'foster', 'city', 'qpcr', '7900ht', 'fast', 'taqman', 'master', 'mix', 'universal', 'roche', 'science', 'mannheim', 'recommendations', 'forward', 'primer', 'anterior', 'microsynth', 'switzerland', 'lysed', 'octyl', 'st', 'louis', 'mo', 'protease', 'mini', 'edta', 'cocktail', 'tablets', 'cleared', 'separated', 'polyvinylidene', 'membrane', 'rad', 'hercules', 'pbs', 'Tween', 'membranes', '13g', 'actin', 'billerica', 'horseradish', 'peroxidase', 'conjugated', 'rockford', 'incubation', 'supersignal', 'west', 'dura', 'immediately', 'seeding', 'reagent', 'protocol', 'provided', '2000', 'changed', 'opti', 'mem', 'phenol', 'prepare', 'qrt', 'bromodeoxyuridine', 'brdu', 'strands', 'elisa', 'briefly', 'labeling', 'solution', 'stopped', 'adding', 'optical', 'density', '450', '595', 'wavelength', 'transmigration', 'transwell', 'chambers', 'corning', 'sciences', 'lowell', 'coated', 'fibronectin', 'chamber', 'filled', 'cotton', 'swab', 'acetic', 'dapi', 'nine', 'locations', 'imagej', 'bethesda', 'md', 'determining', 'area', 'coved', 'scratch', 'wound', 'cm', 'confluence', 'sirnas', 'monolayer', 'pipette', 'tip', 'detached', 'propidium', 'iodide', 'rnase', '05', 'flow', 'cytometry', 'g1', 'g2', 'phases', 'dean', 'jett', 'flowjo', 'tree', 'ashland', 'oregon', 'charcoal', 'stripped', 'dihydrotestosterone', 'dht', 'statistical', 'clinicopathological', 'descriptive', 'statistics', 'cross', 'tailed', 'fisher', 'exact', 'spearman', 'univariate', 'kaplan', 'meier', 'log', 'rank', 'spss', 'chicago', 'repeated', 'paired', 'unpaired', 'welch', 'correction', 'immunoreactivity', 'mild', 'degree', 'pz', '204', '146', '71', 'secretory', 'inconsis

tently', 'hyperplastic', 'tz', 'positivity', 'chromogenic', 'original', 'magnification', 'options', 'differentiate', 'zones', '114', '55', '59', 'coexisting', 'wilcoxon', '016', 'summary', 'saw', 'tendency', 'stemming', 'subgroup s', 'elucidate', 'relationship', 'relation', 'striking', 'epithelium', 'stroma', 'steeply', 'toward', '534', '213', '444', '354', '769', '89', '109', '123', '198', '018', '094', '74', '128', '205', '017', '774', '102', '105', 'conduct', 'concordance', 'marker', 'cox', 'margin', 'status', 'stratified', '86', 'discerned', 'subgroup', 'multivariate', 'curves', 'immortalized', 'undetectable', 'mrna', 'strongest', 'discrepancy', 'post', 't75', 'flasks', 'nontransfected', 'main', 'characteristics', 'uncontrolled', 'counted', 'propagation', '0008', 'taken', 'dishes', 'reseeded', 'od', 'question', 'reflected', 'declined', 'accord', 'proportions', 'compare', 'histograms', 'peak', 'replicates', 'analyze', 'reanalyzed', '183', '02', '295', 'prominent', 'elucidated', 'porous', 'accounted', 'uncoated', 'reliably', 'applicable', 'tend', 'dense', 'monolayers', 'migrated', 'wounding', '10c', 'diminished', 'motility', 'altogether', 'triplicate', 'later', 'wounds', 'inducibility', 'asked', 'mediating', '11a', '11b', 'etoh', '11c', '11d', 'mount', 'intraepithelial', 'analyzing', 'knowledge', 'largest', 'endorse', 'revealing', 'variability', 'turned', 'resolves', 'current', 'mostly', 'noticed', 'emphasizes', 'accurately', 'defining', 'origins', 'comparative', 'commonly', 'profiling', 'distinctly', 'misleading', 'inflammation', 'atrophy', 'preneoplastic', 'explains', 'popularity', 'morphologically', 'appearing', 'extends', 'detailed', 'especially', 'ie', 'lost', 'influences', 'slowed', 'slowing', 'substantiate', 'mitotic', 'regulator', 'maintenance', 'promoted', 'trastuzumab', 'herceptin', 'stimuli', 'conflicting', 'details', 'interplay', 'remain', 'explained', 'understanding', 'setting', 'contradicts', 'assumption', 'unfavorable', 'aggressiveness', 'fourth', 'hgpin', 'dealing', 'theory', 'guy', 'tmpRSS2', 'erg', 'agonists', '3beta', 'diol', 'repesses', 'tgf', 'hypoxia', 'cohorts', 'coexpression', 'intriguing', 'encountered', 'crpcs', 'fits', 'observation', 'repressing', 'jain', 'authors', 'speculated', 'recruitment', 'cis', 'physical', 'possibility', 'androgens', 'substantiated', 'synchronous', 'invariably', 'subtype', 'assumed', 'eg', 'hormonal', 'amplify', 'contributing', 'candidate', 'anaplastic', 'atc', 'responsible', 'deaths', 'diagnosis', 'radiation', 'prolonging', 'systemic', 'urgently', 'copy', 'silencing', 'reversing', 'gains', '14q21', 'follicular', 'classified', 'papillary', 'undifferentiated', 'indolent', 'invading', 'neck', 'spreading', 'existing', 'prolong', 'dysregulated', 'pax8', 'tf', 'neoplasms', 'endodermal', 'primordia', 'metabolism', 'hepatic', 'enhancers', 'enabling', 'chromatin', 'esophageal', 'subfamilies', 'foxo', 'indirect', 'effectors', 'endoderm', 'lamin', 'hills', 'ribose', 'p21', 'pharmingen', 'e1', 'bcpap', 'tpc', 'kindly', 'dr', 'california', 'los', 'angeles', 'cytokeratins', 'akiyama', 'foundation', 'squamous', 'ishii', 'center', 'medicine', 'jichi', 'school', 'mcf', 'mda', '231', 'dmem', 'fcs', 'streptomycin', 'fugene', 'reproducibility', 'plko', 'lentiviral', 'rnai', 'consortium', 'sh', 'infections', '293t', 'packaging', 'gag', 'pol', 'vsv', 'progeny', 'virus', 'released', 'filtered', 'filters', 'infect', 'polybrene', 'puromycin', 'brown', 'primers', 'p21waf1', 'amplicon', 'systemically', 'dissociation', 'curve', 'fractionation', 'composed', 'dtt', 'cold', 'left', 'ice', 'centrifuged', 'lysate', '800', 'rpm', 'supernatant', 'fraction', 'resuspended', 'pellet', 'mixture', 'immunoblot', 'bands', 'densitometry', '103', 'brdurd', 'trypsinized', 'pelleted', 'bsa', 'denatured', 'hcl', 'borate', 'fitc', 'monoclonal', 'dark', 'facscalibur', 'beeton', 'dickinson', 'immunocytometry', 'exciting', '488', 'fluorescence', '514', 'bandpass', 'filter', 'pass', 'overlap', 'sub', 'paraffin', 'embedded', 'microarrays', 'criteria', '177', 'adenomas', '58', 'nodular', 'lymphocytic', 'thyroiditis', 'graves', 'classic', 'anonymized', 'archival', 'files', 'laboratory', 'james', 'dublin', 'insular', 'thyroidectomy', 'males', 'females', '93', 'thyroglobulin', 'calcitonin', 'carcinoembryonic', 'medullary', 'institutional', 'board', 'processing', 'harvesting', 'gently', 'sp

un', 'aspirated', 'paraformaldehyde', 'fix', 'wash', 'aspiration', 'richard', 'allan', 'dropwise', 'vortexed', 'cooled', 'pellets', 'processed', 'processo r', 'permeabilized', 'washes', 'diluted', 'rinsed', 'texas', 'diamidino', 'ph enylindole', 'vectashield', 'mounting', 'imaged', 'plan', 'apochromat', 'oil', 'immersion', 'lens', 'axioplan', 'epifluorescence', 'zeiss', 'scoring', 'successive', 'microwave', 'cycles', '6000', 'autostainer', 'steps', 'biotin', 'avidin', 'diaminobenzidine', 'counterstaining', 'semiquantitatively', 'quantity', 'averaged', 'across', 'replicate', 'cores', 'lymphocytes', 'slide', 'preparation', 'baked', 'soaked', 'xylene', 'air', 'dried', 'boiled', 'ssc', 'placed', 'thermobrite', 'digest', 'digestions', 'situ', 'hybridization', 'nick', 'translation', 'abbott', 'vysis', 'directions', 'precipitated', 'cot', 'stock', 'formamide', 'dextran', 'probes', 'hybridizations', 'darkened', 'olympus', 'bx51', 'running', 'cytovision', 'genus', 'rp11', 'spectrumgreen', 'spectrumred', 'dutp', 'pac', 'resource', 'children', 'oakland', 'signals', 'polysomic', 'student', 'pearson', 'coefficient', 'microsoft', 'excel', 'redundant', 'contaminated', 'arising', 'reason', 'authenticity', 'cytoplasm', 'bright', 'mb', 'immunostaining', 'confluence', 'shrnas', 'depleted', 'infecting', 'lh', 'giant', 'spindle', 'squamous', 'eccentric', 'conspicuous', 'nucleoli', 'diffusely', 'focal', 'pattern', 'atypical', 'proliferative', 'mitoses', 'sia', 'foci', 'metaplasia', 'fish', 'stains', 'morphology', 'intense', 'polysomy', 'interphase', 'cohybridized', 'areas', 's1b', 'barely', 'threshold', 's1c', 'suboptimal', 'fixation', 'precluded', 'homogeneously', 'copies', 'occurred', 'disomy', 'monosomy', 'disomic', 's2a', 's2b', 's2c', 'examining', 'poorly', 'exhibit', 'imbalance', 'abnormal', 'rapid', 'answer', 'week', 'slight', 'uptake', 'redistribution', '04', 'silenced', 'controlled', 'cyclins', 'monitored', 's3a', '3f', 's3b', 'inverse', 'excluded', 'technical', 'divided', '547', 'eleven', '61', 'absent', 'residual', 'link', 'pathologic', 'unbalanced', 'chromosomes', 'ploidy', 'events', 'clinicopathologic', 'harbor', 'complexity', '14q', 'rule', 'simply', 'machinery', 'propose', 'reexpression', 'usually', 'affecting', 'brca1', 'widely', 'tumoral', 'controversial', 'predictor', 'fkhr11', 'fkhr', 'death', 'deregulate', 'trigger', 'lineages', 'translates', 'epigenetic', 'signatures', 'establishing', 'lineage', 'identifying', 'networks', 'transcriptome', 'losses', 'ets', 'fusions', 'castrate', 'mutations2', 'sequenced', 'exomes', 'heavily', 'pretreated', 'autopsy', 'na', 'grade', '00', 'integrating', 'disruptions', 'chd1', 'define', 'fusionnegative', 'ets2', 'deleted', 'recurrent', 'histone', 'modifying', 'mll2', 'mll', 'collaborating', '147', 'untreated', 'physically', 'utx', 'asxl1', 'landscape', 'prioritize', 'curable', 'annually', 'deprivation', 'varied', 'in1', 'resequencing', 'cancer3', 'genomes', 'berger', 'kumar', 'xenografts', 'cases2', 'wa2', 'autopsy8', 't1', 't97', 'coverage', 'aberrations', 'cgh', 'acgh', 'uploaded', 'oncomine', 'visualization', 'exploration', 'exceptions', 'false', 'recurrently', 'tp53', 'zfhx3', 'rb1', 'apc', 'or511', 'cdk12', 'encodes', 'h3k4', 'methyltransferase', 'serous', 'carcinoma10', 'wnt', 'network', 'driver', 'damage', 'checkpoint', 'classical', 'suppressors', 'oncogenes', 'deleterious', 'prkdc', 'i1137fs', 'e640', 'subunit', 'strand', 'break', 'recombination', 't96', 'extremely', 'peaks', 'intrigued', 'chr', '5q21', 'remodeling', 'splice', 'side d', '13b', 'deregulation', 'totaling', '835', '954', '0001', 'identifies', 'rearranged', 'explored', 'dataset', 'harbored', 'etv3', 'p327fs', 'wa56', 'w38', 'wa26', 'r437c', 'wa30', 'erg11', 'telomeric', 'head', 'deletions', 'extending', 'wa31', 'outlier', 'contacting', 'factors11', 'intervening', 'ets212', 'invasion', 'vcap', 'opposite', 'suppressive', 'besides', 'figs', 'infrequent', 'concentrated', 'remodelers', 'immunoprecipitated', 'ash2l', 'cofactor15', 'ezh2', 'h3k27', 'interactions', 's16a', 'modifiers', '16b', 'r1881', 'klk3', 'ar16', 'regulators', 'bp', 's453fs', 't12', '340fs', 'p358fs', 'indels', 'lapc', 'signaling17', 'transactivating', '4d', 'soft', 'colony', 'xenograft', '4e', 'notable', 'pre', 'lastly', 'integrative', 'genomics', 'useful', 'ultimately', 'survey', 'exons', '577', '8p22', '10q23', '13q13', '13q14', 'm

apping', 'brca2', 'amplifications', '5p13', 'p12', '7q22', 'xq12', 'skp2', 'pairs', 'chd5', 'divergent', 'progenitor', 'divergence', 'deep', 'male', 'malignancy', 'united', '2009', '192', '280', '360', 'jemal', 'decreases', 'primarily', 'estimating', 'society', 'remained', '050', '2008', 'incredible', 'make up', 'uncover', 'eight', 'died', 'expose', 'replicated', '8p', '8q', 'associate', 'arrays', '4200', 'spaced', 'liu', 'million', '94', 'contrary', 'mirna', 'taylor', '2010', 'interrogated', 'ncoa2', 'shed', 'gross', 'klf6', 'ephb2', 'chek2', 'formerly', '1992', 'gottlieb', '2004', 'sun', '2005', 'dong', 'zhen g', 'wong', 'describes', 'plexin', 'ngs', 'discovered', 'implicating', 'searched', 'uncovered', 'suppressor', 'consists', 'cellularity', 'confound', 'quality', 'frozen', 'content', 'platform', 'built', 'unique', 'log2', 'plotted', 'agilent', 'analytics', 'passed', 'relied', 'adm', 'algorithm', 'calling', 'lipson', 'neighboring', 'list', 'descriptions', 'massive', 'plex', 'droplet', '4002', 'annotated', 'exon', 'amplicons', 'sjoblom', 'bases', 'junctions', 'pipeline', 'homer', 'sa', 'sternberg', 'jd', 'carpten', 'dw', 'craig', 'prep', 'gatk', 'mckenna', 'insertions', 'metrics', 'annotation', 'missense', 'nonsense', 'filtering', 'visual', 'inspection', 'taking', 'positives', 'empirically', 'validated', 'help', 'nonsynonymous', 'sift', 'predicts', 'likelihood', 'henikoff', '2001', '2003', '83', 'damaging', 'calls', 'ceu', '1000', 'pilot', '1000genomes', 'formatted', 'input', 'submission', 'datasets', 'predicted', 'supports', 'boundaries', 'h874y', '21q', 'easily', 'maher', 'encompasses', 'mipol1', 'foxa2', 'gao', 'yu', 'vital', 'circos', 'plot', 'gray', 'inner', 'circle', 'deviating', 'neutrality', '176', 'rapamycin', 'catalog', 'database', 'maps', 'evolution', 'treating', '850', '5p', 'skp1', 'ubiquitinatin', 'degradation', 'cdkn1b', 'bai', '1996', 'hershko', 'yang', '2002', 'wang', 'encompassing', 'middle', 'defects', '13q12', 'q13', 'moro', 'dramatic', 'germline', 'brca', 'fong', 'interrogation', 'substantive', '279', 'glioma', 'p53', 'astrocytomas', 'ring', 'e3', 'ubiquitin', 'ligases', 'hu', '1997', 'cooperates', 'siah', 'ebi', 'adaptor', 'beta', 'catenin', 'matsuzawa', 'reed', 'somatically', 'snf2', 'rad54', 'helicase', 'swi', 'snf', 'atpase', 'chromodomain', 'thompson', 'neuroblastoma', 'gastric', 'gorringe', 'p19', 'arf', 'methylation', 'hypermethylation', 'probable', 'rho', 'gef', 'larg', 'fused', 'believed', 'rhoa', 'gtpase', 'b1', 'invasive', 'perturbed', 'comes', 'ancestral', 'stated', '10p', 'arm', '10q', '18q', '1q21', 'q22', '9q34', '17p11', '21q21', 'oldest', 'diverged', 'younger', 'evidenced', 'came', 'reconcile', 'completed', 'art', 'facilitated', 'detecting', 'heterogeneous', 'gleaned', 'glimpse', 'describing', 'opportunities', 'cost', 'poised', 'compendium', 'clues', 'treat', 'essentially', 'worldwide', 'alk', 'cancer2', 'akt1', 'erbb2', 'involve', 'ros1', 'ret', 'successful6', 'typically', 'sqcc', 'sqccs', 'sox2', 'whsc11', 'cdkn2a9', 'nfe212', 'keap1', 'bai3', 'fbxw7', 'grm8', 'muc16', 'runx1t1', 'stk11', 'erbb4', 'refs', 'ddr2', 'nominated', 'targets13', 'atlas', 'tcga', 'twin', 'epigenomic', 'exonic', 'messenger', '178', 'histopathologically', 'wgs', 'microrna', '159', 'demographic', 'downloaded', 'portal', 'https', 'nci', 'docs', 'publications', 'lusc', 'dataabstract', 'classification', 'cdk n2a', 'references', 'acknowledgements', 'author', 'information', 'commentstumour', 'histologically', '137', 'written', 'informed', 'consent', 'local', 'boards', 'alive', 'updated', 'november', '2011', 'ninety', 'cent', 'history', 'tobacco', 'north', 'sqcc16', 'assessments', 'experts', 'histological', 'alterationsabstract', 'commentsthe', 'analysed', '323', 'segments', '165', 'platforms', 'projects', 'february', '477', 'glioblastoma', 'multiforme', '282', '156', 'gave', 'rise', 'scnas', 'scna', 'similarities', '3q', 'reported9', 'gistic', 'cdk6', 'mdm2', 'bcl2l1', 'eys', 'foxp1', 'nf1', '193', '863', '690', 'silent', '228', 'megabase', 'cancer17', 'multiforme18', 'cpg', 'transitions', 'transversions', 'mutsig', 'fdr', 'hla', 'notch1', 'robust', 'reads', 'kilobase', 'rpkm', 'purity', 'truncating', 'sqccs22', 'vertically', 'columns', 'arranged', 'emphasize', 'mutual', 'exclusivity', 'syn', '462k', 'indexne xtto', 'power', 'detect', 'fam123b', 'wtx', 'smarca4', 'smad4', 'tsc1', 'tnfa

ip3', 'crebbp', 'differed', 'leu858arg', 'inframe', 'leu861gln', 'performing', 'hybrid', 'recapture', '76', '289', 'satisfactory', 'verify', '283', '235', 'computed', '108', 'myeloma', 'pathwaysabstract', 'commentsmany', 'oxidative', 'stress', 'cul3', 'chemo', 'radiotherapy', 'types28', 'motifs', 'dlg', 'etge', 'mutually', 'paradigm', 'shift30', 'exert', 'frequencies', '458k', 'previousfigures', 'indexnextwe', 'tp63', 'notch2', 'ascl4', 'cancers31', 'sqccs32', 'ascl1', 'cells33', 'modulators', 'overlapping', 'classificationabstract', 'commentswhole', '502', 'indices', 'signatures34', 'yielding', 'subtypes', 'primitive', '3q26', 'dominant', 'deltan', 'p40', 'lacks', 'tap63', 'oncogene35', 'enriched', 'anticorrelated', 'grouped', 'kruskal', 'wallis', 'tests', 'categorical', 'usage', 'cin', 'absolute', 'numbers', 'gistic23', 'iclust', 'assignments', 'standardized', 'expr', 'mut', '135', '516k', 'indexnextindependent', 'associations', 'locusabstract', 'commentsintegrated', 'multiplatform', 'sqcc38', 'p16ink4a', 'p14arf', 'inactivating', 'skipping', 'faceted', 'schematic', 'denoted', 'black', 'circles', 'axis', 'colour', 'coded', 'delta', 'het', 'homoz', 'kiaa1797', 'p16ink4', 'orf', 'reading', 'binary', '490k', 'indexnextanalysis', 'patterns', 'splicing', 'intergenic', 'intra', 'intrachromosomal', 'rearrangement', 'arid1a', 'ctnna2', 'vhl', 'ink4a', 'decay', 'targetsabstract', 'commentsmolecularly', 'sqccs13', 'instances', 'gefitinib39', '171', 'phosphatidylinositol', 'oh', 'pi', 'subunits', 'proteases', 'assessor', 'score40', 'fda', 'targetable', 'families', 'erythroblastic', 'erbb3', 'fgfrs', 'janus', 'jaks', 'amplified41', 'spectra', 'centred', 'vulnerabilities', 'restriction', 'curated', '7d', 'depicted', 'memo43', '496k', 'indexdiscussionabstract', 'commentslung', 'carcinoma17', 'dysfunction', 'hallmark', 'avoiding', 'immune', 'destruction44', 'programmed', 'pd1', 'lymphocyte', 'ctla4', 'cancer46', 'genotypic', 'immunotherapies', 'successfully', 'minimally', 'papilloma', 'hras22', 'organize', 'analyse', 'actionable', 'personalized', 'methodsabstract', 'commentsall', 'allprep', 'capture', 'shotgun', 'expectation', 'circular', 'segmented', 'consensus', 'approaches20', 'oncogenesis', 'genesis', 'abundant', 'translocations', 'interdependent', 'phenomenon', 'chromoplexy', 'dysregulation', 'coordinately', 'derangement', 'punctuated', 'characterizing', 'hierarchy', 'path', 'though', 'public', 'health', 'needs', 'deeper', 'dysregulate', 'baca', 'garraway', 'inactivate', 'tsgs', 'shen', 'abate', 'spop', 'kdm6a', 'implicate', 'barbieri', 'grasso', 'overexpress', 'tomlins', 'magi2', 'broken', 'ends', 'shuffled', 'ligated', 'shuffling', 'prevalence', 'operate', 'computational', 'inferred', 'disarray', 'typical', 'handful', 'discrete', 'employ', 'restructuring', 'greek', 'meaning', 'constraining', 'coordinated', 'fashion', 'neuroendocrine', 'grades', 'pathological', '1013', 'aligned', 'uniquely', 'hg19', 'depths', '356', '136', 'end', 'orientations', 'confidence', '172', 's3c', '113', 'singleton', 'sense', 'preserving', 'joined', 'nrf1', 'pr', 'crkl', 'mapk1', '1084', 'leaving', 'domains', 'gsk3b', 'foxo1', 'breakpoints', 'near', 'distributions', 'appear', 'aberrant', 'systematically', 'reveals', 'concurrent', 'material', 'evident', 'overlaid', 'bridges', 'span', 'informative', 'partners', 'creating', 'simple', 'predispersed', 'burrow', 'de', 'probabilistic', 'probability', 'distance', 'simulated', 'guanine', 'cytosine', 'replication', 'timing', 'scrambled', 'breakage', 'deviation', 'nearer', 'chance', 'correctly', 'raising', 'employs', 'statistically', 'graph', 'deviate', 'arisen', 'forty', 's5a', 'termed', 'produces', '88', 's5b', 'participated', 'genotypic', 'subsets', 'produced', 'inter', 'maximum', '009', 'arose', 'intricate', 'injury', 'hubs', 'occupied', 'localize', 's4a', 'vary', 'resembled', 'chromothripsis', 'rausch', 'stephens', 'shattering', 's4b', 'stability', 'huang', 's4c', 'replicating', 'heterochromatin', '199', 'postulated', 'architecture', 'gaspar', 'maia', 'raise', 'distinctive', 'dysregulates', 'kegg', 'kanehisa', 's5c', 'inclusion', 'nkx3', 'conceivably', 'disrupting', 'underlie', 'geographically', 'ding', 'etv6', '12p', '1q', 'pik3r1', '5q', '17p', '22q', 'advantages', 'incipient', 'accompanying', 'chen', 'paths', 'carter', 'nik', 'zainal', 'genotype', 'estimates', 'r2',

'99', 'admixed', 'clonality', 'reasoning', 'subclonal', 'strictly', 'produce', 'pernuer', 'earliest', 'pin', 'precursor', '1995', 'gives', 'originate', 'directionality', 'relationships', 'beginning', 'bhatia', '1999', 'thereafter', 'evasion', 'gating', 'imply', 'favors', 'fearon', 'vogelstein', '1990', 'continue', 's7a', 'sustain', 'rounds', 'subclones', 'histology', 'indicative', 'occupy', 'overlapped', 'mermel', 'gs', '0059', '6e', 'beyond', 'burden', 'forment', 'underpinnings', 'experimentally', 'systematic', 'suppressing', 'safeguards', 'individually', 'theoretically', 'generations', 'occurrence', 'depends', 's7b', 'fails', 'precisely', 'fuse', 'immensely', 's7c', 'exclude', 'plausible', 'unifying', 'intrachromosomal', 'gc', 'heterochromatic', 'compartments', 'drier', '2013', 'alternating', 'aggregate', 'ways', 'dispersed', 'neoplasm', 'fewer', 'catastrophic', 'lin', 'awaits', 'localizing', 'testosterone', 'align', 'overexpressing', 'brenner', 'chronology', 'founder', 'expands', 'begins', 'proceeds', 'continuum', 'classically', 'gradual', 'challenged', 'incurred', 'reassembly', 'expanded', 'apply', 'passenger', 'gradually', 'divisions', 'extremes', 'lies', 'drawing', 'analogy', 'species', 'periods', 'gould', '1977', 'framework', 'accords', 'ranging', 'initiating', 'quickly', 'diffuse', 'territory', 'disadvantage', 'regime', 'incidental', 'prediction', 'raises', 'portend', 'missed', 'panels', 'prevention', 'consented', 'weill', 'cornell', 'college', 'centers', 'australia', 'conjunction', 'pty', 'lt d', 'perth', 'nodules', 'homogenization', 'nucleotides', 'illumina', 'gaiix', 'instrument', 'bwa', 'picard', 'pipelines', 'sourceforge', 'net', 'mutect', 'indelocator', 'broadinstitute', 'cga', 'basepair', 'participating', 'inspected', 'manually', 'clara', 'implemented', 'considers', 'nodes', 'connected', 'edges', 'deviates', 'null', 'evaluates', 'adjacently', 'assign', 'searching', 'plausibly', 'correspond', 'nearest', 'neighbor', 'distances', 'breaks', '1mb', 'windows', 'window', 'consideration', 'observing', 'rejected', 'benjamin', 'connecting', 'scenarios', 'assigned', 'wise', 'finalized', 'assigning', 'step', 'maximizes', 'admixture', 'infer', 'estimate', 'allelic', 'fractions', 'hemizygously', 'lowest', 'simulation', 'corrected', 'subclonality', 'progesterone', 'triple', 'phosphatase', 'tensin', 'bioinformatics', 'grb7', 'ppm1d', 'hdac2', 'rasa1', 'anchorage', '511', 'abnormalities', 'technological', 'advances', 'enable', 'comprehensively', 'differently', 'definitions', 'improving', 'routinely', 'tamoxifen', 'premenopausal', 'women', 'aromatase', 'postmenopausal', 'application', 'hierarchical', 'duct', 'lobular', 'unit', 'milk', 'secreting', 'surrounded', 'outer', 'layer', 'contractile', 'myoepithelial', 'cadherin', 'biologically', 'dubbed', 'varies', 'methodologic', 'variation', 'meta', 'resemble', 'indications', 'made', 'logistic', 'classifiers', 'esr1', 'pgr', 'pathologist', 'training', '143', '142', 'insufficient', 'neve', 'surrogates', 'differs', 'exhibiting', 'trend', 'excluding', 'thresholds', 'signature', 'quantifying', 'reflects', 'worse', '500k', '11q13', '6q', '17q21', '17q23', '15p', 'asterisks', 'regard', 'intervals', 'consistent', 'merged', 'experimentation', 'hypotheses', 'gained', 'veer', 'meets', 'dotted', 'log10', 'reflecting', 'passing', 'fulfills', 'acetyltransferase', 'hbo1', 'meeting', 'rps6kb1', 'p70', 'recurrence', 'oraov1', 'fadd', '14q13', 'amplitude', 'relies', '116', '5q13', 'p120gap', 'drives', 'mcf7', 'skbr3', 'row', 'column', 'plots', 'causing', '73', 'enrichment', 'cal', 'imperfect', 'retinoblastoma', 'rb', '006', 'reproducible', 'adoption', 'mrnas', 'reached', 'reproducibly', 'protects', 'contributed', 'caveat', '435', 'intestine', 'requiring', 'repeat', 'adequate', 'pinpoint', 'links', 'sometimes', 'meet', 'prioritization', 'mir', 'bona', 'fide', 'acts', 'heterozygosity', 'shifting', 'gdp', 'maintaining', 'checkpoints', 'exploit', 'continuing', 'fresh', 'logic', 'accompanied', 'bt', '474', '483', '549', 'hcc1143', 'hcc1395', 'hcc1419', 'hcc1428', 'hcc1569', 'hcc1806', 'hcc1954', 'hcc2218', 'hcc38', 'hs578t', 'kpl', '361', '415', '436', '453', '468', 'sk', 'br', '47d', '812', '893', 'zr', '120', '148', 'efm', 'hcc1937', 'p1', 'jimt', '223', 'deutsche', 'sammlung', 'von', 'mikroorganismen', 'und', 'zellkulturen', 'gmbh', 'glutamine', 'su

bconfluent', 'dnaeasy', 'kits', 'hg', 'u133', 'suite', '100k', 'dchip', 'centered', 'converted', 'glad', 'r5', 'spotfire', 'tibco', 'national', 'repository', 'nos', 'assignment', 'supplied', 'smallest', 'variance', 'train', 'intensity', 'centroid', 'nonoverlapping', '100kb', '820', 'therein', '862', 'hg17', 'ucsc', 'browser', 'mass5', 'scaled', '556', 'interval', '714', 'multiplyin', 'blat', 'overlaps', 'refseq', 'proportional', 'hazards', 'adjustments', 'prioritizing', 'big', 'dye', 'terminator', 'abi', '3730x1', 'machines', 'surveyor', 'softgenetics', 'manual', 'trace', 'round', 'nontumor', 'wider', 'impair', 'record', 'dbsnp', 'polymorphisms', 'insignificant', 'rest', 'frameshifts', 'equal', 'nupage', 'bis', 'difluoride', 'criterion', 'chemiluminescence', 'scanned', 'pcmv', 'xl5', 'origene', 'dharmafect', 'dharmacon', 'thousand', 'hundred', 'agarose', 'wk', 'calcein', 'imagexpress', 'micro', 'devices', 'serious', 'complication', 'rectal', 'period', '1981', '1986', 'ward', 'ferrara', 'postoperatively', 'operations', 'dukes', 'recurred', 'locally', 'gradin', 'distal', 'histopathologic', 'anastomosis', 'indicator', '110', 'g3', 'examination', 'histologic', 'interpreted', 'staging', 'pioneering', 'discoveries', 'disparate', 'emt', 'opposing', 'orthotopic', 'concordant', 'contradictory', 'wherein', 'progresses', 'drastically', 'illustrate', 'developmentally', 'conditional', 'severely', 'ductal', 'immature', 'abnormally', 'layers', 'coincidentally', 'opens', 'compact', 'occupies', 'pancreatic', 'pca', 'bad', 'enhancing', 'modulated', 'attempted', '111', 'attenuating', 'manipulation', 'requested', 'northwestern', 'specialized', 'excellence', 'spore', 'operating', 'snap', 'removal', 'entry', 'pcr8', 'gw', 'topo', 'adenoviral', 'recombinant', 'pad', 'cmv', 'v5', 'lr', 'clonase', 'plenti', 'puro', 'dest', 'addgene', 'pgipz', 'slug', 'shrnmir', 'luc', 'luciferase', 'monitor', 'living', 'antisense', 'gl2', 'duplex', 'cad', 'hrp', 'vimentin', 'sc', 'frodo', 'wi', 'mit', 'edu', 'primer3', 'sybr', 'rnas', 'trizol', 'bioanalyzer', '2100', 'humanht', 'beadchip', 'preprocessed', 'genomestudio', 'quantile', 'beadarray', 'bioconductor', 'cutoff', 'ontology', 'gorilla', 'cbl', 'cs', 'ac', 'wst', 'instruction', '5000', '20000', 'ul', 'incubator', 'oc', 'absorbance', 'microplate', 'reader', 'invasions', 'handling', 'institutes', 'care', 'welfare', 'athymic', 'nu', 'weeks', 'old', 'jackson', 'implantations', 'suspension', '3m', 'weekly', 'ivis', 'hopkinton', '405', '516', 'division', 'cofactor', 'mediator', 'un', 'synergistically', 'aforementioned', 'namely', 'preclude', 'adaptive', 'deplete', '22rv1', 's1d', 'seq', 's1f', '50kb', 's1g', 'affirm', 'ectopic', 'manifested', 'greatly', 'tried', 'du145', 'decided', 'sensitize', 'blockade', 'attenuate', 'bicalutamide', '2g', 'collectively', 'matrigel', 'boyden', 'healing', 'drastic', 'invaded', 'course', 's3d', 'opposes', 'repressed', 'deliniate', 'vim', 'morphologic', 'dys', 'astrocyte', 'fusiform', 'fibroblastic', 'shaped', 'appearance', 'rounded', 'reversal', 'clusters', '4f', 'compares', 'intragenic', 'flank', 'enhancer', '5e', 'steadily', '5f', '5g', 'stimulates', '5h', 'direction', 's5d', 's5f', 'invasiveness', '5i', 'chose', 'ideal', 'studying', 'implantation', 'live', 'antibiotic', 'inoculated', 'regional', 'merge', 'regionally', 'inoculation', 'euthanized', '150', 'marginally', 's6a', 'corroborate', 's6b', 'dominates', 'functionalities', 'lentivirus', 'carrying', 'wildtype', '7a', 'mechanistically', 'localization', 'dominate', 'firstly', 'nicely', 'associating', 'dissect', 'tetracycline', 'story', 'indirectly', 'activator', 'begun', 'repressor', 'recruited', 'hdac7', 'repression', 'warranted', 'works', 'synergy', 'dangerous', 'reporting', 'impaired', 'designing', 'restore', 'avenue', 'endeavors', 'pinpointing', 'therapeutics', '230', 'eighteen', 'rit1', 'mga', 'female', 'rbm10', 'met', 'investigations', 'smoking', 'proportionally', 'never', 'smokers', 'cigarettes', 'oncogenessuch', 'egfr1', 'translocated', 'orros1', 'investigational', 'mostlung', 'identifiable', 'cdkn2a8', 'challenging', 'genome11', 'attention', 'pathobiology', 'lepidic', 'acinar', 'micropapillary', 'mucinous', 'colloid', 'unclassifiable', '163', 'eighty', 'pastor', 'summarizes', 'demographics', 'previously15', 'cellularity16', 'wes', 'germ', '97', 'previously17', 'mutsig2cv18', '182', 'adenoca

rcinomas12', '412', 'mutsig2cv', 'luad', '2014', 'setd2', 'u2af1', 'max', 'pathway19', 'frameshift', 'hitherto', 'unappreciated', 'indel', 'negatives', 'adenine', 'wide12', 'transversion', 'inversely', 'thymine', 'th', '269', 'tl', '144', 'reports13', 'harboured', 'highcohort', 'scrc', 'gender', 'adenocarcinoma22', 'chromosome23', '003', 'adenocarcinoma24', 'nkx2', 'tert', 'ccne1', 'terc', 'mecom', 'described24', '8q24', 'ccnd3', 'chromothriplisis25', 'catalogue', 'aberration', 'seventy', 'analyses15', 'stabilized', 'protein26', 'identified27', 'y1003', '8a', 's34f', 'adenocarcinoma12', 'contribution', 'u2af1s34f', 'selection28', 'cassette', 'ctnnb1', 'protooncogene', '8b', 'delineation', 'trans', 'governing', 'sixty', 'others30', 'uncharacterized', '9a', 'depi ction', 'differing', 'map2k1', 'comparisons', 'nominate', 'unrecognized', '9b', 'gistic31', 'adenocarcinoma11', 'q79', 'transform', 'nih3t3', 'signalling32', 'sequelae', 'phosphoproteomic', 'phosphoprotein', 'undetected', 'p70s6kinase', 'bp1', 'ampk', 'phosphorylation33', 'explain', 'unsupervised', 'nomenc lature', 'disease34', 'coordinate', 'naming', 'histopathological38', 'anatomy', 'classifications', 'tru', 'bronchioid', 'inflammatory', 'pp', 'magnoid', 'outcomes34', 'integration', 'analyte', 'histopathology', 'membership', 'pro gnostically', 'previously34', '7c', 'rppa', 'heterogeneity', 'unbiased', 'island', 'methylator', 'cimp', 'igh', 'ow', 'occupying', 'report40', 'gata2', 'gata4', 'gata5', 'hic1', 'hoxa9', 'hoxd13', 'rassf1', 'sfrp1', 'sox17', 'wif1', '0015', 'remodelling', 'clustering41', '15q', 'comparatively', 'distingui shing', '19p', 'contexts', 'setd2andderegulated', 'tonic', 'diversity', 'seem ingly', 'undescribed', 'underscores', 'implicates', 'modifications', 'methylo me', 'emphasizing', 'coincident', 'illuminating', 'refinement', 'personalizat ion', 'abstract', 'fragments', 'p16', 'landmark', 'scanning', 'technique', 'spans', 'tags', 'southern', 'regular', 'sharp', 'decades', 'metaplastic', 'columnar', 'epithelia', 'replace', 'gastroesophageal', 'reflux', 'dysplasia', '20q', 'cysteine', 'cathepsin', 'gata', 'rlgs', 'sts', '19q12', 'noti', 'vicinity', 'est', 'regeneration', 'gut', 'rodent', 'pathogenic', 'obtaining', 'mich igan', 'cardia', 'mucosa', 'pulmonary', 'portion', 'oct', 'miles', 'remainde r', 'weight', 'portions', 'brief', 'digested', 'ecrv', 'beverly', 'dctp', 'dgtp', 'nen', 'products', 'boston', 'disk', 'hinf1', 'cloning', 'dimension', 'autoradiographic', 'dynamics', 'sunnyvale', 'digital', 'imagequant', 'v1', 'standards', 'preparative', 'spots', 'pieces', 'electrophoresed', 'deae', 'schleicher', 'schuell', 'keene', 'nh', 'eluted', 'distilled', 'pbc', 'la', 'jolla', 'compatible', 'seakem', 'gtg', 'fmc', 'bioproducts', 'rockland', 'ligation', 'boehringer', 'indianapolis', 't4', 'ligase', 'chloroform', 'x11', 'competent', 'electroporation', 'pulser', 'preps', 'alkaline', 'sambrook', 'elect ronic', 'similarity', 'matches', 'databases', 'templates', 'perform', 'qg', 'melting', 'tm', 'fluorometer', 'hoefer', 'francisco', 'polynucleotide', 'ta q', 'madison', 'denaturing', 'ts', 'ns', 'loaded', 'hybridized', '459', '1485', 'untranslated', 'avoid', 'downers', 'grove', 'pepsin', 'gaithersburg', 'touch', 'prelabeled', 'bacpac', 'resources', 'concert', 'maxiprep', 'a260', 'd nase', 'transcribed', 'transcriptase', 'primed', 'dt', 'hexamers', 'coamplified', 'label', 'vic', 'quencher', 'pe', 'flanking', '7700', 'detector', 'secti oned', 'acetone', 'horse', 'vectastain', 'burlingame', 'spin', 'valencia', 'c rna', 'genechip', 'enlarged', 'thumb', 'visible', 'digestion', 'match', 'cont ig', 'linkage', 'whitehead', 'stanford', 'housekeeping', '12p13', 'centromeric', '14q11', 'reactions', 'dosage', 'dnas', '14q32', '254', 'resides', 'coamp lification', 'nkx', 'nfkb1a', 'mad', 'syntenic', 'distinguish', 'possessing', 'diagrammatic', 'drawn', 'verification', 'discriminate', 'constitutes', 'homo genously', 'minute', 'paralleled', 'ests', 'lungs', 'c66', 'h01', '6f', 'rece iving', '6g', 'sporadic', '6h', 'metaphase', 'manuscript', 'virtual', 'northe rn', 'serial', 'sage', 'polypeptide', '473', 'turn', 'fork', 'stomach', 'inte stines', 'secreted', 'sonic', 'hedgehog', 'shh', 'foxo4', 'foxo1a', 'pediatri c', 'alveolar', 'rhabdomyosarcoma', 'pax3', 'chicken', 'embryo', 'qin', 'fibr osarcomas', 'chickens', 'necessity', 'era', 'trump', 'morphological', 'diagno

stics', 'prognostics', 'choice', 'billions', 'possibilities', 'goes', 'mont h', 'clever', 'tricks', 'childhood', 'manageable', 'paraphrase', 'audrey', 'e vans', 'leader', 'ies', 'brodeur', 'mycn', 'propensity', 'familial', 'phox2 b', 'homeobox', 'moss', 'returned', 'fashioned', 'trees', 'pedigrees', 'autos omal', 'inheritance', 'scan', 'investigators', 'r1275q', 'y1604', 'bordering', 'janoueix', 'lerosey', 'george', 'minority', 'f1174l', 'endogenously', 'cells6', 'people', 'individuals', 'neuroblastomas', 'f1174', 'difference8', 'stratification', 'mixtures', 'multigene', 'sensitivity8', 'studies4', 'optimize', 'paper', 'sasaki', 'myoblastic', 'crizotinib', 'ranbp2', 'person', 'activity10', 'launch', 'eml4', 'cancers11', 'effectively12', 'helpful', 'emerges13', 'coincides', 'outcome8', 'pleiotrophin', 'midkine', 'alk14', 'ergistically', 'season', 'frontier', '1596', 'minor', 'allow', 'embryonal', 'sympathetic', 'nervous', 'cheung', 'dyer', 'maris', 'improvements', 'cure', 'lag', 'behind', 'smith', 'rely', 'immunotherapeutic', 'disialoganglioside', 'gd2', 'escalating', 'consolidation', 'improves', 'survivors', 'hobbie', 'focuses', 'neural', 'neonatal', 'brain', 'iwahara', 'morris', 'hereditary', 'tkd', 'hallberg', 'palmer', 'imparts', 'soda', 'lymphomas', '1994', 'motivating', 'kwak', 'shaw', 'engelman', 'expedited', 'nct00939770', 'underpin', 'predicting', 'assembled', 'collaboration', 'cog', 'cohorta', 'subjects', 'r1060', 'transmembrane', 'variations', 'r1231q', 'i1250t', 'd1349h', 'forbes', 'r1275', 'f1245', 'i1170', 'i1171', 'incidences', 'constitutional', 'knudson', '1972', 'r1060h', 'i1183t', '11204f', 'alongside', 'biased', 'skewed', 'brouwer', 'efs', '195', 'rarer', 'megabases', 'concurrently', 'bagci', 'accordingly', 'inferior', 'os', 'g1286', 'biomarkers', 'older', 'univariable', 'consisting', 'young', 'unresectable', 'multivariable', 'hazard', 'hr', 'adjusted', '0006', 'category', 'arguing', 'refining', 'truly', 'passengers', 'fall', 'algorithms', 'polyphen', 'adzhubei', 'classifier', 'predictsnp', 'bendl', 'adversely', 't1343i', 'gnad', 'molina', 'vila', 'accurate', 'explicit', 'motivated', 'considerations', 'tkds', 'bresler', 'accelerated', 'm1166', 'y1278', 'g1128a', 'r1192p', 'bourdeaut', 't1151', '11196', 'i1183', 'a1200', 'r1231', 't1343', 'd1349', 'signifying', 'd1270g', 'autophosphorylated', 'd1270', 'mg2', 'l11240v', 'peptide', 'kcat', 'auto', 'phe', 'stabilizes', 'autoinhibited', 'bossi', 'le e', 'mutating', 'autoinhibitory', 'participate', 'cyan', 'magenta', 'grey', 'l11196m', 'sch', 'nherr', '2011a', 'fell', 'narrow', 'saturated', 'efficiencies', 'trends', 'i1170n', 'quantitation', 'correspondence', 'yields', 'straight', 'outliers', 'm1166r', 'slope', 'zero', 'and44', 'argue', 'judging', 't1151m', 'and4', 'oncogenicity', 'paints', 'picture', 'choi', 'mere', 'myofibroblast', 'zhang', 'solely', 'assuming', 'f1245c', 'f1245v', 'y1278s', 'i1170s', 'i1171n', 'argued', 'atpa', 'katayama', 'competition', 'resembling', 'accounting', 'concurs', 'ceritinib', 'i1171t', 'friboulet', 'johnson', 'sakamoto', 'macrocyclic', 'pf', '06463922', 'r1192', 'g1128', 'performance', 'informatics', 'assessing', 'trajectories', 'bonding', 'intramolecular', 'bonds', 'configurations', 'contacts', 'readily', 'monitoring', 'solvent', 'sasa', 'classed', 'favor', 'perturbation', 'fep', 'simulations', 'destabilizes', 'principal', 'motions', 'trajectory', 'shih', 'eigenvalue', 'destabilization', 'predictions', 'agree', 'perturbing', 'hydrophilic', 'elevate', 'computationally', 'special', 'removes', 'folding', 'reasons', 'excepting', 'plagued', 'reiterate', 'druggable', 'powered', 'capability', 'cataloguing', 'inss', 'congenital', 'adolescents', 'adults', 'atrxi', 'poorest', 'modal', 'myeloablative', 'conditioning', 'marrow', 'transplant', 'immunotherapy', 'retinoic', 'opportunity', 'drove', 'chand', '2011b', 'success', 'exceeds', 'blindly', 'assesses', 'elevating', 'improve', 'employed', 'park', 'formulate', 'predictive', 'leverage', 'upfront', 'responder', 'incorporates', 'pharmacologic', 'backbone', 'contemporary', 'regimens', 'facts', 'carr', 'wilkinson', 'confounding', 'expect', 'yield', 'advance', 'management', 'predisposition', 'siblings', 'implementing', 'surveillance', 'practices', 'nrf2', 'interfaces', 'proteasome', 'distribute', 'behaved', 'r554q', 'w544c', 'n469fs', 'p318fs', 'g333c', 'suppre

ssion', 'hypomorphic', 'r320q', 'r470c', 'g423v', 'd422n', 'g186r', 's243c', 'v155f', 'augmented', 'intriguingly', 'super', 'binder', 'superbinder', 'ubiquitinate', 'strengthen', 'mechanics', 'phenotypically', 'hypomorphs', 'dead', 'catalogued', 'genotypes', 'instruct', 'principles', 'presumed', 'proteosomal', 'liberate', 'spread', 'heme', 'oxygenase', 'hmox1', 'nadph', 'quinone', 'nqo1', 'strengthens', 'defense', 'neutralize', 'reactive', 'oxygen', 'roses', 'xenobiotic', 'reprogram', 'modulating', 'anabolic', 'suit', 'demands', 'uncertain', 'compromised', 'interventions', 'concepts', 'mind', 'biochemically', 'connects', 'a549', 'h2228', 'authenticates', 'tandem', 'passaged', 'resuscitation', 'mefs', 'thomas', 'kensler', 'nobunao', 'wakabayshi', 'humidified', 'mef', 'imdm', 'hd', 'rnaimax', 'ggu', 'cuc', 'cug', 'aaau', 'acc', 'ucu', 'cau', 'uau', 'gaa', 'ugu', 'gga', 'cag', 'uuc', 'acg', 'uca', 'ucg', 'aca', 'agu', 'aua', 'cac', 'buffers', 'ha', 'polyclonal', 'a2066', 'tubulin', 't7816', 'proteintech', 'ab290', 'h300', 'slk', 'bethyl', 'a300', '499a', 'dpp3', 'ab97437', 'mcm3', '123a', 'ikk', '2678', 'p62', 'sqstm', 'sc25575', 'ab13248', '2759', '8727', '4499', '2622', 'a190', '131a', 'np', '50mm', '2mm', 'ripa', '25mm', 'deoxycholate', '150mm', 'immunopurification', 'resin', 'washing', 'eluting', 'celsius', 'ne', 'codes', 'hnqo1', 'kind', 'jeffrey', 'quantification', 'renilla', 'cytomegalovirus', 'firefly', 'ub1', 'venus', 'npm1', '1mm', 'r320', 'hi5', 'insect', 'ion', 'ubch5', 'rbx1', 'neh2', 'immunostaining', 'coverslips', 'cytoskeletal', 'minutes', 'antifade', 'lsm5', 'pascal', 'confocal', 'laser', 'equipped', 'planapo', 'lenses', 'streptavidin', 'fasp', 'raw', 'mzxml', 'resultant', 'sorcerertm', 'sequest', 'build', 'transproteomic', 'tpp', 'v4', 'uniprotkb', 'swiss', 'prot', 'contaminants', 'porcine', 'p00761', 'p00760', 'trypsin', 'concatenated', 'reversed', 'decoy', '494', '4500', 'amu', 'cleavages', 'tolerance', 'binning', 'peptideprophet', 'semi', 'tryptic', 'static', 'carbamidomethyl', 'oxidation', 'serines', 'threonines', 'tyrosines', 'proteinprophet', 'cutoffs', 'apms', 'prohits', 'abundance', 'prey', 'spectral', 'counts', 'bait', 'dividing', 'calculating', 'deviations', 'unfiltered', 'normalizations', 'uniform', 'reiterated', 'ovals', 'truncated', 'reactivity', 'electrophilic', 'l231v', 's224y', 'p3181', 'r711', 'keap1g333c', 'annotate', '049', 'attempt', 'segregate', 'impacted', 'g480w', 'subcellular', 'indistinguishable', 'dimerize', 'sensor', 'biosensors', 'homodimer', 'stoichiometry', 'butylhydroquinone', 'tbhq', 'dimer', 'dimerization', 'rigorously', 'hemagglutinin', 'affinities', 'immunopurified', 'complexes', 'gather', 'fail', 'interactors', 'amer1', 'ikbkb', 'refer', 'superbinders', 'mean', 'artificial', 'compartment', 'turnover', 'cycloheximide', 'pulse', 'chase', 'behaves', 'exogenous', 'unbound', 'eluate', 'degrade', 'denaturation', 'latitude', 'classify', 'consider', 'rarely', 'cbiportal', 'loosely', 'fitness', 'surprising', 'analogously', 'sh2', 'antagonizes', 'said', 'tertiary', 'casts', 'doubt', 'cullin', 'neddylation', 'transitioning', 'slow', 'unfolded', 'deubiquitination', 'proteolysis', 'ubiquitinates', 'hinder', 'immediate', 'perturb', 'accepted', 'saturation', 'compromising', 'multifunctional', 'viability', 'prolonged', 'nf', 'investigating', 'looking', 'marginal', 'attempts', 'inputs', 'speculative', 'orient', 'ivr', 'kelch', '181', 'discernible', 'biochemistry', 'dictated', 'biosensor', 'add', 'cysteines', 'pka', 'attack', 'spatial', 'localizations', 'folded', 'extension', 'ideas', 'gateway', 'outsourced', 'derive', 'twenty', 'personal', 'pursuing', 'vus', 'segregation', 'pathogenicity', 'participants', 'unclassified', '98', 'adaptable', 'life time', 'prophylactic', 'chemopreventive', 'myriad', 'lake', 'ut', 'biologic', 'ethnic', 'start', 'causality', 'makes', 'ineligible', 'carriers', 'segregating', 'sides', 'settings', 'loh', 'signed', 'ascertained', 'sources', 'ohio', 'genetics', 'mayo', 'eligibility', 'records', 'reclassified', 'eosin', 'microdissected', 'cored', 'lymphblastoid', 'proteinase', 'nucleic', 'ethidium', 'bromide', 'alphaimage', 'innotech', 'leandro', 'visually', 'ambiguous', 'uncut', 'undercalling', 'exosap', 'usb', 'cleveland', 'traces', 'dnastar', 'reviewers', 'consisted', 'carrier', 'evolutionary', 'grantham', 'gvgd', 'agvg

d', 'iarc', 'fr', 'php', 'uses', '1641', '1863', '2401', '3200', 'assigns', 'c65', 'probabilities', 'c35', 'c45', 'c55', 'c15', 'c25', 'c0', 'odds', 'esefinder', 'splicesitefinder', 'theoretical', 'aberrantly', 'spliced', 'size s', 'extra', 'lethality', 'fanconi', 'anemia', 'discriminating', 'nond deleterious', 'appendix', 'online', 'multifactorial', 'combines', 'causation', 'variables', 'likelihoods', 'interpretation', 'suspected', 'a2', 'a3', 'nonmutate d', 'surprised', 'and3', 'ivs2', 's127n', 'v1804d', 'a1170v', 'e2856a', 'h1966y', 'i2285v', 'm784v', 'n1878k', 'y42c', '433', 'avgd', 'and6', 'a4', 'a5', 'intronic', 'ivs20', 'ivs23', 'ivs8', '12delta', 'ivs5', 'weighted', 'incorrectly', 'woman', 'phenocopy', 'inherited', 'ivs13', 'l1764p', 'warrant', 'suggestive', 'adapted', 'weighting', 'kconfab', 'cytokeratin', 'routine', 'onset', 'delay', 'decisions', 'offered', 'tuberous', 'sclerosis', 'tsc', 'disorder', 'malformations', 'neuronal', 'tsc2', 'cns', 'neurological', 'presentations', 'seizures', 'hydrocephalus', 'cognitive', 'psychological', 'impairments', 'emphasis', 'perinatal', 'cortical', 'worsening', 'born', 'neuropsychiatric', 'hyperactivity', 'circuit', 'heterozygote', 'synaptic', 'plasticity', 'learning', 'memory', 'keywords', 'mental', 'retardation', 'neurogenesis', 'epilepsy', 'autism', 'sega', 'spine', 'developmental', 'heart', '1998', 'hamartin', 'tuberin', 'european', 'slegtenhorst', 'kwiatkowski', '2003a', 'manning', 'tsai', 'haploinsufficiency', 'symptoms', 'morbidity', 'vries', 'begin', 'infantile', 'spasms', 'unresponsive', 'autistic', 'traits', 'weber', 'presently', 'cures', 'thanks', 'etiology', 'horizon', 'explores', 'anatomical', 'disturbances', 'connectivity', 'recapitulating', 'recapitulate', 'subcortical', 'uterus', 'morphogenesis', 'deficits', 'tbc1d7', 'heteromeric', 'homolog', 'rheb', 'inoki', 'nakashima', 'tee', 'gap', 'sato', 'mtorc1', 'mtorc2', 'sabatini', '2012a', 'actions', '2012b', 'eif4e', 's6k1', 'simplified', 'canonical', 'regards', 'unaddressed', 'mosaic', 'seminal', 'mendelian', 'penetrance', 'genetically', 'parents', 'au', 'novo', 'mosaicism', 'rose', 'sporadically', 'harris', 'generalized', 'jansen', 'sancak', 'kinetics', 'goedbлоed', 'mayer', 'pointed', 'suffer', 'hits', 'tubers', 'subependymal', 'astrocytoma', 'roberts', 'haploinsufficiency', 'debate', 'becomes', 'severity', 'mortem', 'epileptic', 'brains', 'matter', 'heterotopia', 'lissencephaly', 'dysplasias', 'historically', 'disorders', 'glioneuronal', 'holmes', 'outgrow', 'seizure', 'cerebral', 'pharmacologically', 'intractable', 'magnetic', 'resonance', 'mri', 'electroencephalography', 'eeg', 'neurotransmitter', 'pet', 'scans', 'kato', 'candidacy', 'epileptiform', 'realization', 'surgeries', 'cognition', 'social', 'behaviors', 'contention', 'electrically', 'progressively', 'progressive', 'boer', 'gaba', 'mori', 'maturation', 'cortex', 'experience', 'prone', 'organization', 'neurons', 'excitatory', 'excitation', 'astrocytes', 'ectopicall y', 'dysmorphic', 'yamanouchi', '1997a', '1997b', 'white', 'hypo', 'myelination', 'occasionally', 'cystic', 'calcified', '4x', '10x', 'square', '20x', 'p28', 'cre', 'transgenic', 'intensively', 'glial', 'fibrillary', 'acidic', 'gfa p', 'severe', 'floxed', 'fl', 'crossed', 'recombinase', 'mimicked', 'dysplastic', 'crossing', 'radial', 'glia', 'e12', 'macrocephaly', 'hippocampal', 'astrocytosis', 'nestin', 'goto', 'former', 'postnatal', 'hallmarks', 'dlx5', 'gabaergic', 'precursors', 'carson', 'fu', 'neurodevelopmental', 'susceptibility', 'prenatal', 'reverses', 'limitation', 'introducing', 'lateral', 'ventricles', 'embryos', 'pulses', 'uterine', 'versatile', 'electroporated', 'nakajima', 'knocked', 'electroporations', 'e16', 'corpus', 'callosum', 'hypertrophic', 'dendritic', 'advancement', 'creation', 'behavioral', 'deficit', 'demyelination', 'ages', 'neonates', 'supplementing', 'discrepancies', 'explanations', 'revert', 'coincidentally', 'zeng', 'missing', 'piece', 'puzzle', 'elegantly', 'vacuolated', 'demonstration', 'takes', 'addresses', 'deletes', 'recapitulates', 'rescuing', 'reversible', 'sen', 'sens', 'contiguous', 'lvs', 'bender', '1980', 'multinucleated', 'enlarge', 'diameter', 'ess', 'astrocytic', 'spinal', 'fluid', 'csf', 'remove', 'porter', 'ablate', 'halted', 'everolimus', 'hydroxyethyl', 'derivative', 'carries', 'protracted', 'subventricular', 'sv

z', 'adulthood', 'rodents', 'infant', 'npcs', 'olfactory', 'bulb', 'nucleus', 'seki', 'arai', '1991', 'biallelic', 'chan', 'hypothesize', 'zhou', 'creert 2', 'ventricle', 'conclude', 'migratory', 'p7', 'p10', 'corresponds', 'perinatal', 'hamartomas', 'neuron', 'forebrain', 'ganglia', 'le', '1988', 'gallagher', 'inoue', 'delete', 'condition', 'upregulation', 'neun', 'patch', 'clam p', 'recording', 'biophysical', 'survived', 'received', 'prevented', 'zhu', 'defect', 'expansion', 'baker', 'concentrating', 'simplicity', 'placement', 'dendrites', 'hope', 'rescue', 'alternations', 'decline', 'dysfunctions', 'synapses', 'communication', 'axons', 'behave', 'antennae', 'receive', 'wiring', 'relays', 'connective', 'postnatally', 'p14', 'branching', 'apical', 'thickness', 'polarity', 'granule', 'spines', 'membranous', 'protrusions', 'terminate', 'plastic', 'shapes', 'adapting', 'constantly', 'bear', 'postmitotic', 'slice', 'injection', 'biopsy', 'postmortem', 'shortened', '1984', 'densities', 'purkinje', 'diffusion', 'tensor', 'axonal', 'cues', 'contextual', 'environment', 'nie', 'axon', 'specification', 'neurite', 'destined', 'serves', 'playing', 'polarization', 'sad', 'tau', 'collapsing', 'determines', 'morita', 'melanogaster', 'photoreceptor', 'guidance', 'persistent', 'projections', 'neurites', 'retinal', 'ganglion', 'rgc', 'connections', 'misregulated', 'ephra', 'ephrin', 'detrimental', 'acquiring', 'malformation', 'neurocognitive', 'iq', 'subtler', 'speech', 'impairment', 'manipulations', '2010b', 'potentiation', 'depression', 'strengthening', 'weakening', 'strength', 'rats', 'depen ded', 'anxiety', 'yuan', 'synapse', 'unchecked', 'architectural', 'distortion s', 'hamartoma', 'syndromes', 'neurofibromatosis', 'fragile', 'johnston', 'rett', 'fkbp1a', '2010a', 'documentation', 's235', '236', 'rsk', 'roux', 'moore', 'faithful', 's240', '244', 'slices', 'paradoxically', 'auerbach', 'themes', 'propelled', 'reorganization', 'notch', 'karbowniczek', 'fight', 'avenues', 'nutrients', 'integrates', 'sensing', 'lysosomal', 'subpopulations', 'reside', 'associates', 'lysosome', 'serve', 'integrate', 'multicellular', 'organisms', 'autonomous', 'endocrine', 'properly', 'niche', 'relaying', 'diabete s', 'raptor', 'mlst8', 'shifts', 'catabolic', 'stimulating', 'lipids', 'howell', 'costly', 'carbon', 'evolved', 'exquisite', 'cytokines', 'hormones', 'igf 1', 'relay', 'cascades', 'responds', 'senses', 'stems', 'lying', 'rag', 'gtapas', 'heterodimer', 'complexed', 'kim', 'gator1', 'recruits', 'endomembrane', 'farnesylation', 'clark', 'saito', 'takahashi', 'scaffolds', 'effector', 'potter', 'acutely', 'establishment', 'turns', 'disassembly', 'cai', 'pan', 'dan', 'plas', 'spatially', 'localizes', 'dissociates', 'farnesylated', 'temporal', 'hela', 'unstimulated', 'immunoprecipitations', 'recognize', 's939', 't1462', 'quaternary', 'exclusion', 'fractionated', 's1e', 'heterotrimeric', '365', 'kda', 'pools', 'loosen', 'intermolecular', 'yip', 'turning', 'colocalizing', 'punctate', 'cytosol', 'golgi', 'discernable', 'colocalization', 'mitochondria', 'peroxisomes', 'starvation', 's2e', 'immunoblots', 'disrupts', 's2f', 'stochastic', 'stronger', 'lamp1', 's2g', 'lysosomes', 'nocodazole', 'microtubule', 'cytoskeleton', 's2h', 'dispersion', 'colocalizes', 'heavy', 'tethered', 'p18', 'myristoylation', 'palmitoylation', 'moieties', 'nada', 'assemble', 'exogenously', 't389', '293e', 'proper', 'hara', 'anticipated', 'refeeding', 'reciprocally', 'reciprocal', 'rheb1', 'perinuclear', 'colocalized', 'anchoring', 'farnesyl', 'fti', 'mobility', 'dynamically', 'withdrawal', 'influenced', 'converging', 'predominant', 's6c', 's6d', '6i', 'property', 'unusual', 'amongst', 'gaps', 's6e', 'weakened', '6j', 'wortmannin', 'mk2206', 'coincided', 'acceptor', 'pras40', '7e', 'linking', '7f', 'worth', 'noting', 'pc3', '7g', 's7d', '7h', 's7e', '7i', 'brings', 'lipid', 'castro', 'preferenc e', 'preferentially', 'bos', 'daumke', 'conversion', 'partner', 'holds', 'gdi', 'assembly', 'courses', 'hoogeveen', 'westerveld', 'pentameric', 'arrangements', 'constitute', 'oligomeric', 'cooperatively', 'organizational', 'positioned', 'constituents', 'lymphangiomyomatosis', 'lam', 'megalencephaly', 'nellist', 'branch', 'cool', 'muscle', 'contractions', 'jacobs', 'dialyzed', 'replacement', '63x', '40x', 'objectives', 'apotome', 'engaged', 'lsm', 'axio

vision', 'pseudo', 'velocity', 'calculation', 'thresholded', 'coefficients', 'percentages', 'pixels', 'fields', 'regarding', 'immunoblotting', 'hypomelanotic', 'macules', 'facial', 'angiofibromas', 'shagreen', 'patches', 'cephalic', 'plaques', 'ungual', 'fibromas', 'intellectual', 'disability', 'illness', 'angiomyolipomas', 'cysts', 'arrhythmias', 'enlarging', 'neurosurgery', 'thre atening', 'neurologic', 'occasion', 'recommended', 'embolization', 'sparing', 'topical', 'symptomatic', 'cardiac', 'complications', 'afterward', 'asymptomatic', 'periodically', 'ventricular', 'enlargement', 'mrис', 'deemed', 'visit s', 'abdomen', 'angiomyolipoma', 'glomerular', 'filtration', 'echocardiogram', 'infants', 'dyspnea', 'shortness', 'breath', 'visit', 'tomography', 'sign s', 'baseline', 'annual', 'dermatologic', 'dental', 'ophthalmologic', 'vision', 'circumstances', 'nephrectomy', 'relatives', 'counseling', 'thirds', 'off spring', 'inheriting', 'pregnancies', 'fibrous', 'plaque', 'rhabdomyoma', 'confetti', 'scattered', 'arms', 'legs', 'pits', 'nonrenal', 'achromic', 'revise d', 'definite', 'proband', 'duplication', '2015', 'click', 'pdf', 'simplex', 'probands', 'jones', 'dabora', 'detects', 'duplications', 'multiplex', 'mlpa', 'rendtorff', 'concluded', 'lovd', 'inferring', 'kozlowksi', 'persons', 'card', 'intrafamilial', 'milder', 'virtually', 'disfigurement', 'neuroimaging', 'obstruction', 'ascertainment', 'epilepticus', 'shepherd', 'refers', 'interrelated', 'academic', 'neuropsychological', 'psychosocial', 'difficulties', 'concerns', 'ever', 'krueger', 'asd', 'chung', 'language', 'lewis', 'seriously', 'debilitating', 'task', 'goh', 'disabilities', 'humphrey', 'intelligence', 'quotient', 'dropped', 'underscore', 'aggression', 'eden', 'self', 'questionnaire', 'addressing', 'overestimated', 'kidneys', 'oncocytoma', 'adenomatus', 'patel', 'vessels', 'sheets', 'smooth', 'adipose', 'bleeding', 'parenchyma', 'eosinophilic', 'polycystic', 'pkd1', 'compress', 'berry', 'aneurysms', 'rcc', 'pea', 'fat', 'immunologic', 'hmb', 'regress', 'disappear', 'outflow', 'obstructive', 'hemoptysis', 'chest', 'radiographs', 'reticular', 'ct', 'infiltrates', 'pneumothorax', 'nhlbi', 'mccormack', 'avila', 'multifocal', 'proliferations', 'pneumocytes', 'physiologic', '2009b', 'predilection', 'confused', 'premalignant', 'eye', '3573', 'healthy', 'newborns', 'detachment', 'neovascular', 'glaucoma', 'shields', 'aml', 'retrospective', 'grossman', 'parathyroid', 'insulinoma', 'islet', 'pheochromocytoma', 'carcinoids', 'sal eem', '919', 'truncation', 'vs', 'connor', 'wentink', 'snvs', 'webb', 'osborne', 'inappropriate', 'eliminated', 'adenoma', 'sebaceous', 'ash', 'leaf', 'shape', 'births', '1989', 'infancy', 'epithelial', 'urothelial', 'heritable', '1983', 'vitiligo', 'nevus', 'vogt', 'harada', 'angiofibroma', 'acne', 'mistaken', 'distinguishes', 'omit', 'dermal', 'trauma', 'holding', 'fibromatosiss', 'uncommonly', 'becker', 'schneider', '1975', 'transmit', '2013a', 'evaluations', 'conference', 'exam', 'hypertension', 'electroencephalogram', 'video', 'subclinical', 'dilated', 'funduscopy', 'electrocardiogram', 'ecg', 'conduction', 'walk', 'creatinine', 'gfr', 'consultation', 'geneticist', 'counselor', 'permits', 'obviate', 'neurosurgical', 'encephalopathy', '2009a', 'anticonvulsants', 'baumgartner', 'aid', 'surgeon', 'radiolabeled', 'accuracy', 'staged', 'bilateral', '2013b', 'underway', 'davies', 'tolerability', 'preferable', 'safety', 'corticosteroids', 'exophytic', 'acceptable', 'hemorrhage', 'avoided', 'insufficiency', 'childbearing', 'oophorectomy', 'production', 'transplantation', 'willmore', 'educated', 'adolescence', 'determination', 'adolescent', 'preventive', 'measures', 'ultrasound', 'clinicaltrials', 'cultural', 'ethical', 'substitute', 'professional', 'ed', 'leukocyte', 'child', 'mildly', 'sibs', 'purpose', 'paternity', 'maternity', 'assisted', 'reproduction', 'undisclosed', 'planning', 'pregnancy', 'risks', 'reproductive', 'banking', 'storage', 'methodology', 'preimplantation', 'fetuses', 'trimester', 'fetus', 'staff', 'umbrella', 'organizations', 'registries', 'medline', 'trait', 'alliance', '801', 'roeder', 'road', '750', 'silver', 'spring', 'phone', '225', 'toll', '301', '562', 'fax', 'email', 'info', 'aes', 'place', 'east', '332', '320', 'cincinnati', '513', '777', '877', '287', 'omim', 'elsewhere', 'hg

md', '16p13', 'hgnc', 'uniprot', 'entries', '191092', '191100', '605284', 'han', 'calmodulin', 'cdk', 'autophagy', 'el', 'lawrence', 'yeung', 'mak', 'releasing', 'hamartias', 'disorganized', 'arrangement', 'modify', 'inactivates', 'prevents', 'longest', '000368', 'noncoding', '1350', 'recur', '512', 'homologies', 'coiled', 'coil', 'erm', 'kappa', 'globular', 'ezrin', 'moesin', 'lamb', 'tnf', 's6kinase', 'vegf', '000548', 'comprises', 'polymorphic', '3600', '1442', '2714g', '4508a', 'conventions', 'hgvs', 'quick', 'rap1a', 'rab5', 'xiao', 'gsk3', 'pecoma', 'dickson', 'sj', 'v599e', 'ikenoue', '8132', '8137', 'iec', 'intestinal', 'g468a', 'g468e', 'f467c', 'r461i', 'i462s', 'g463e', 'irrelevant', 'transducers', 'reportedly', 'melanomas', '1796', '599', 'gxgxxg', 'covers', 'anchors', 'shielding', 'solvents', 'glycines', 'g361s', 'g463v', 'pcdna3', 'anne', 'vojtek', 'ann', 'arbor', 'mi', 'k482m', '442', 'bacterially', 'pgex', '2t', 'bucksinghamshire', 'kingdom', 'v12', 'palo', 'alto', 'pfa', 'elk1', 'gal4', 'pfr', 'riken', 'tsukuba', 'japan', 'geneticin', 'poolled', 'radioimmunoprecipitation', 'placid', 'ny', '470', 'morpholinepropanesulfonic', '525', 'imager', 'ch', 'immunoblotted', 'p44', 'prl', 'tk', 'pnf', 'violet', 'dimethylthiazol', 'yl', 'diphenyl', 'tetrazolium', 'mtt', 'isopropanol', 'formazan', '570', 'suspended', 'dish', 'polyp', 'oropharynx', 'insr', 'phosphorylase', 'phk', 'apk', 'boxed', 'phosphospecific', 'cotransfection', 'ss32', '36aa', 'bs', 'k600e', 'refractile', 'flattened', 'constituting', 'distinguishable', 'subline', 'microscopic', 'visualize', '9c', 'contrasts', 'lid', 'replaces', 'avoidance', 'fgfr3', 'reinforce', 'n540k', 'k650e', 'r669g', 'consolidates', 'azd4547', 'bgj', '398', 'tki258', 'jn42756493', 'ap24534', 'replacements', 'translatable', 'precision', 'fgfs', 'fgf', 'endometrial', 'fgfr3tacc3', 'fgfr2', 'fgfr4', 'owing', 'nonselective', 'tkis', 'ponatinib', 'brivanib', 'nintedanib', 'lenvatinib', 'dovitinib', 'lucitanib', 'vegfrs', 'accumulative', 'liming', 'focusing', 's249c', 'kd', 'quantitatively', 'figure1', '397', '806', 'icgc', 'biomuta', 'covering', 'figure1a', 'numbering', 'k650', 'g697', 'n540', 'g697c', 'figure1a1a', 'figure1b1b', 'figure1b', 'i538', 'r669', 'figure1c', 'ef', 'hot', 'spot', 'brake', 'figure1d', 'broaden', 'ranking', 'impacts', 'assume', 'deleteriousness', 'condel', 'sophisticated', 'predictors', 'saap', 'energetic', 'foldx', 'ranked', 'table1', 'kds', 'figure2a', 'k650n', 'n540s', 'r669q', 'i538v', 'figure2b', 'figure2a2a', 'v555m', 'd641g', 'd641n', 'az12908010', 'v555', 'twelve', 'and2b', 'd617g', 'g637w', 'd617', 'hrd', 'g637', 'figure2b2b', 'protective', 'figures2a2a', 'figure2c2c', 'figure3', 'figure3a', 'figure3b', 'figure3c3c', 'figure4a4a', 'figure2', 'e466k', 'underpins', 'figure4b', 'r675g', 'figure4c', 'apo', '4uw', '3p', '3gqi', 'adopts', 'r675', 'h650', 'tight', 'y653', 'inset', 'disturbed', 'favoring', 'repositioning', 'nmr', 'cps', 'promotion', 'forcing', 'dissociate', 'triggered', '669', 'pockets', 'rationalize', 'soaking', 'preformed', 'crystals', 'crystallographic', 'asymmetric', 'monomers', 'rmsd', '09', 'monomer', 'figure5a5a', 'bjg', 'pd173074', 'quinoxaline', 'a564', 'dimethoxyphenyl', 'orientated', 'perpendicular', 'v561', 'methoxy', 'atoms', 'atom', 'd641', 'pyrazole', 'solubilizing', 'away', 'shallow', 'n628', '1630', 'a640', 'pit', 'indentation', 'isoxazole', '4f64', '4f65', '4nk9', '4nka', '4nks', 'baf3', 'efficacies', 'figure6', 'figure6a', 'figure6b6b', 'figure6c', 'biggest', 'figure6d', 'ic50s', 'v561m', 'rigidity', 'occupancy', 'figure5', 'situations', 'figures11', 'figures22', 'figure22', 'oral', 'oscc', 'hypochondriasis', 'figure4', '675', 'consolidate', 'n549h', 'k659n', 'v651m', 'limitations', 'figures6c6c', 'and6d', 'v550l', 'v550e', 'v550m', 'inform', 'jhuem', 'figures6a6a', 'broadly', 'n550k', 'conclusive', 'juvenile', 'aml', '17q11', 'massively', '485', 'cytogenetics', 'karyotype', 'disappeared', 'leukemic', 'hematol', '306', '311', 'wiley', 'periodicals', 'leukemogenesis', 'microdeletions', 'caf', 'lait', 'multicenter', 'enrolled', 'alfa', 'french', 'rq', 'feasible', 'presenting', 'utr', '000267', 'mononuclear', 'fab', 'm3', '394', 'modalities', '0701', 'banding', 'karyotypes', 'iscn', 'categorized', 'inv', '7q', '11q23', 'anomaly', 'wga', 'cytologic', 'cr', 'blasts', 'eliminate', 'c'

nv', 'customized', '105k', 'v3', 'hs', 'predesigned', 'registered', 'tcag', 'nondeleted', '7900', 'ht', 'genotyping', 'rnasep', 'cycling', '95c', 'sec', 'quantifications', 'copycaller', 'preparations', 'mpo', 'kreatech', 'amsterdam', 'netherlands', 'sureselect', '466', 'space', 'bioscope', '936', 'igv', 'viewer', 'bigdye', 'capillary', '3130xl', 'analyzer', 'seqscape', 'mann', 'whitney', 'rr', 'microdeletion', 'mdr', 'eligible', 'id', 'discordant', 'wileyonlinelibrary', 'anomalies', 'nd', 'm4', 'xy', 'p11', 'itd', 'm6', 'm0', 'q21', 'mar', 'm5', 'xx', '786', '297', '476', '012', 'q24', 'q11', 'cebp', '879', '722', '897', 'm1', 'p36', '421', '819', '451', 'xxy', 'm4eo', 'p13q22', 'p13', '124', '299', '970', '434', '853', '065', '330', '196', '578', '380', 'abbreviations', 'wbc', 'british', 'nc', 'sex', '203', '191', '125', '122', '008', '220', '216', '386', '374', '383', '372', '382', '371', '358', '348', '349', '339', '393', '329', '379', '3y', 'snapshot', 'purple', 'boxes', 'circle', 'metaphasic', 'cebpa', 'idh1', '083', '132', 'introns', 'intron', '553', 'bi', '702', '653', 'q32', '703', '388', '773', '701', 'q10', 'q21q26', 'idem', '585', '352', 'homogeneous', 'lengths', '27b', 'nfi', 'limit', 'preamplified', '179', 'underestimate', 'monosomies', 'haferlach', 'parkin', 'q26', 'attribute', 'cmml', 'recovered', 'gilliland', 'persistence', 'nonhematological', 'chou', 'idh', 'implication', 'checked', 'suz12', 'myelofibrosis', 'Oh', 'n17', 'phosphoinositides', 'pc12', 'cos', 'elevation', 'elucidation', 'identities', 'cellu', 'lar', 'teins', 'assumes', 'sort', 'incorpor', 'ated', 'p1', '20ga', 'stimulatory', '20gap', 'neurofibromin', 'nfj', 'gene3', 'sole', 'mediators', 'behaviour', 'coimmunoprecipitates', 'ras35', 'phosphorylating', 'phos', 'phorylated', 'inositol', 'phospholipids', 'reasonable', 'aga', 'sf', 'coinfected', 'p85a', 'hydrolyable', 'analogues', 'val', 'ala', 'peptides', 'sixfold', 'lb', 'fb2', 'phosphotyrosine', 'neutralizing', '238', '25918', 'activ', 'ity', 'mobilized', 'protei', 'hydrolysa', 'ble', 'gm', 'pnp', 'ncu', 'bated', 'bou', 'ncou', 'pled', 'subtracted', 'background', 'pli', 'cate', 'standa', 'rd', 'errors', 'hi', 'bition', 'io', 'tj', 'i5', 'resid', 'ues', 'sos1', 'nhi', 'asso', 'ciation', 'nti', 'bodies', 'tei', 'bacu', 'lovirus', 'descri', 'bed', 'a38', 'rala', 'pu', 'rifed', 'col', 'baculovi', 'rus', 'p85', 'encodi', 'p110', 'supernat', 'ant', 'superdex', 'pid', 'nase', 'im', 'munopreci', 'pitates', 'rther', 'experi', 'ments', 'ruses', 'usion', 'ufactu', 'ers', 'cou', 'covalently', 'affigel', 'biorad', 'followi', 'ufacturers', 'nter', 'incubated', 'wheel', 'petition', 'prei', 'nclusion', 'mgci', 'tlc', 'qua', 'ntified', 'usi', 'ambis', 'ner', 'p85j3', 'hydrolysable', 'gmp', 'ting', 'p58a', 'p85f3', 'tracks', 'cl', 'puri', 'fied', 'kl', 'inac', 'tive', 'vis', 'ualized', 'brilliant', 'track', 'protocols33', 'chemilumi', 'nescence', 'rae', 'ra', 'rap2a', '120gar', 'incubating', 'r916p', 'hiles', 'contaminating', 'sulphate', 'polyacrylamide', 'elec', 'trophoresis', 'accumulated', 'indicat', 'ing', 'subfraction', 'prepa', 'ration', '68k', 'prewashing', 'prob', 'ably', 'oligomerized', 'fore', 'asn', 'phosphory', 'lated', 'phaeochromocytoma', 'pcl', 'rous', 'sarcoma', 'rsv', 'nerve', 'phosphatidylino', 'sitol', 'metabolically', 'labelled', 'orthophosphate', 'phosphoino', 'sitides', 'ngf', 'egf', 'phosphatidyl', 'bisphosphate', 'ptdns', 'p2', 'phosphatidyl', 'nositol', 'trisphosphate', 'ptdns', 'p3', 'incomplete', 'con', 'trol', 'simplest', 'ruled', 'breakdown', 'immunoprecipitates', 'phosphotyro', 'sine', 'domi', 'nant', 'phosphates', 'pho', 'spholipase', 'unaltered', 'kin', 'ases', 'elevates', 'cooperated', 'presum', 'rael', 'par', 'tially', 'termin', 'ally', 'liraf', 'inositides', 'autocrine', 'previ', 'ously', 'feel', 'irs', 'cd19', 'cd2821', 'stimulations', 'brane', 'sh3', 'lck14', 'mitogenic', 'vice', 'versa', 'pdgf27', 'typesn', 'nck', 'sites29', 'receives', 'lapetina', 'yi', '259', 'nl', '1500', '3000', 'zt', 'pci2', 'py', 'ps', 'multiples', 'immunopreci', 'polyphosphate', 'vol', 'august', '994', 'ni7', 'calcium', 'rsvneoni', 'g4i8', 'picked', 'ana', 'ortho', 'analysed34', 'chcl3', 'radioactivity', 'redissolved', 'oxalated', 'silica', 'deacylated', 'methylamine', 'jected', 'hplc', 'cerenkov', 'emission', 'ioo', 'oo', 'imm', 'naf', 'na3v04', 'pnpp',

'pmsf', 'aprotinin', '4gi0', 'ubi', 'coexpressed', 'expres', '531', '25000', '4jp2', 'ptd', 'lns', '30000', '15000', '10000', 'c1l', '0c0', 'oco', 'label ling', 'p1dlns', 'untransfected', 'sion', 'compelling', 'produc', 'tion', 'as e', 'observa', 'tions', 'pmt2p85a', 'psg5p110', 'stand', 'check', 'mecha', 'n isms', 'redundancy', 'synergize', 'gliomas', 'proteasomal', 'pkc', 'expandin g', 'gliomagenesis', 'destruction', 'destabilized', 'gbms', 'elucidating', 'd ownward', 'rasgap', 'martin', 'xu', 'triggers', 'basu', 'declue', 'neurofibro mas', 'sarcomas', 'pheochromocytomas', 'riccardi', 'andersen', 'thiel', 'glio blastomas', 'parsons', 'refined', 'precocious', 'honda', 'pagano', 'cichowsk i', 'bortezomib', 'degraded', 'masking', 'ubiquitinated', 'attenuation', 'not hing', 'interrogate', 'bisindolylmaleimide', 'marcks', 'ro', '8220', 'phorbo l', 'myristate', 'pma', 'heckman', 'marais', 'drnf1', 'dictating', 'gbm', 'ma ckay', 'twelves', 'da', 'rocha', 'u87', 'yazaki', 'polyubiquitinated', 'smea r', 'unmodified', 'degradable', 'tumorigenic', 'neurosphere', '3g', 'looked', 'dbtrg', '05mg', 'galactosidase', 'gal', 'gli36', 'sf539', 'u251', 't98g', 'm 059j', 'u138', 'diploid', 'courtois', 'senescing', 'cdkn1a', 'doxorubicin', 'senesce', 'ikediobi', 'meir', 'mitigated', 'ois', 'restrain', 'narita', 'lo we', 'pilocytic', 'rodriguez', 'spontaneously', 'restricting', 'triggering', 'p15ink4b', 'senescent', 'collado', 'michaloglou', 'archived', 'p15', 'restr ained', 'parada', 'cooperate', 'precedes', 'ln229', 'ln319', 'tallen', 'permi ssive', 'messiaen', 'reasoned', '141', 'beroukhim', 'jensen', 'grd', '217', '6789', '6792delttac', 'pros', 'c4103delt', '211', 'c896', '897duptt', 'c349 6g', 'coincidental', 'lacked', 'squares', 'ts543', 'ts565', 'c2195g', 'c470 g', '657', 'dasgupta', 'johannessen', 'cloughesy', 'cawthon', 'wallace', 'gef s', 'specifying', 'marshall', 'consequential', 'gfrs', 'furnari', 'feldkamp', 'concomitantly', 'stommel', 'sensitized', 'shares', 'restrict', 'choosing', 'undefined', 'pd0325901', 'relieved', 'pmek', 'rebound', 'mitogens', 'facili tates', 'raf1', 'unprecedented', 'prolongation', 'grim', 'brafwt', 'raswt', 'nraswt', 'occult', 'relief', 'mel', 'taha', 'merghoub', 'alan', 'houghton', 'mskcc', 'mewo', 'malme3m', 'snf96', 'm308', 'antoni', 'ribas', 'ucla', 'wm39 18', 'katherine', 'nathanson', 'meenhard', 'herlyn', 'wistar', 'fingerprintin g', 'g12a', 'g12d', 'q61r', 'q61l', 'v600k', 'v600r', 'k601e', 'sequenom', 'h iseq', 'rin', '75bp', 'fastq', 'casava', 'adapter', 'cutadapt', 'trimmed', 'g rch37', 'tophat', 'cufflinks', 'visualizations', 'gplots', 'chemiluminesenc e', 'fuji', 'las', 'lifesciences', '718', '519', '520', 'actinin', '17829', '1862335', 'perk', '9101', '9102', '9121', '9122', 's338', '9427', 'facs', 'trypan', 'vi', 'xr', 'coulter', 'equation', 'rasgtp', 'rbd', 'pull', 'sinf 1', 'smartpool', 'gipz', 'shnon', 'tripz', 'shnf1', 'cloneid', 'v2ths', '2608 06', 'v3ths', '380114', '380110', 'doxycycline', 'genotyped', 'd594g', '264', 'n581s', '215', 'surrogate', 'sixth', 'g12c', '285', 'hemizygous', 'q1336', '266', '1161', 'q282', 'monoallellic', 'hodis', 'melanomagenesis', 'sheath', 'mpnts', 'mpnst', '286', 'kiapp', '1nm', '239', '50nm', '500nm', 'augment', '9nm', 'potencies', 'equipotent', 'mek162', 'spry4', 'durably', 'devoid', 'q1 070', 'nf1wt', 'wcl', 'sufficiently', 'analogous', '350', '2818', 'pigmente d', 'lisch', 'optic', 'leukemias', 'underestimation', 'attributable', 'challe nges', 'advent', 'dimerizing', 'homo', 'melanocyte', 'cooperativity', 'abroga tion', 'maximally', 'diminishing', 'unsuccessful', 'torc1', 'dictates', 'star k', 'attenuates', 'monophosphorylated', 'abrogating', 's212', 'rational', 're troviral', 'hijacked', 'kirsten', 'harvey', 'searches', 'transfectable', 'ret rovirus', 'sequences1', 'controlling', 'approx20', 'undergone', 'genes2', 've ssel', 'formation3', 'man', 'slime', 'mould', 'worms', 'flies', 'frogs4', 'pl ayers', 'proteins5', 'tumours6', 'come', 'essential7', 'ubiquitous', 'transla tionally', 'principally', 'mislocalized', 'recruit', 'translational', 'interv ention8', 'engage', 'gdp4', 'slowly', 'catalysed', 'recently12', 'overview', 'sos2', 'son', 'sevenless', 'caenorhabditis', 'elegans', 'grb2', 'sos', 'kin ases15', 'comthe', 'shc', 'simplification', 'differently14', 'opposed', 'orde rs', 'circumstances18', 'proteins19', 'depth3', 'relocated', 'activation20',

'maps', 'fos', 'jun', 'ap1', 'heterodimers22', 'cycle23', 'pi3ks', 'pleckstrin', 'paid', 'pdk1', 'pkb', 'pkcs', 'p70s6k', 'targets26', 'ras27', 'rac', 'kappab', 'pathways28', 'comonc'e', 'initiate', 'ralgds', 'ral', 'phospholipase', 'cepsilon', 'plcepsilon', 'catalyses', 'diacylglycerol', 'mobilization', 'stores', 'glycogen', 'synthase', 'pla2', 'pld', 'p90', 'stimulator', 'rgl', 'rsb2', 'rgl2', 'rlf', 'cdc42', 'ralbp1', 'family30', 'kip1', 'bim', 'fas', 'ins', 'matrix', 'metalloproteinases', 'gene32', 'crest', 'neu', 'carcinomas33', 'domain34', 'overactivated', 'hard', 'approx70', 'approx15', 'p110alpha', 'akt2', 'tumours35', 'messengers', 'approx30', 'tumours36', 'covalent', 'attachment', 'isoprenoid', 'caax', 'peptidomimetics', 'pyrophosphate', 'mimicking', 'bisubstrate', 'pharmaceutical', 'companies', 'ftis', 'libraries', 'therapies37', 'expectations', 'mmtv', 'peptidomimetic', 'mammary', 'fti39', 'realized', 'increasingly', 'successes', 'root', 'geranylgeranyltransferase', 'ggt', 'geranylgeranylation', 'attempting', 'ggtis', 'combination40', 'put', 'repeatedly', 'geranylgeranyl', 'protein41', 'arguments', 'made37', 'unresolved', 'recently42', 'rebranding', 'leukaemias44', 'oligonucleotides', 'protein545', 'directing', 'rnaseh', 'cells46', 'isis', 'pharmaceuticals', 'phosphothioate', 'derivatives', 'isis2503', 'isis5132', 'cells47', 'conceptual', 'isis6957', 'trials47', 'braf49', 'delivering', 'ensuring', 'mainstream', 'industry', 'adept', 'leap', 'leukaemia51', 'entered', 'ago', 'pd98059', 'conditions55', 'immunodeficient', 'carcinomas56', 'orally', 'redesignated', '1040', 'awaited', 'bay43', '9006', 'site57', 'melanomas58', 'ester', 'pathway59', 'egfr60', 'erbb', 'inhibitors61', 'humanized', 'domains63', 'transphosphorylation', 'osi', 'tarceva', 'genentech', 'diarrhoea', 'carcinoma64', 'randomized', 'gemcitabine', 'cisplatin', 'paclitaxel', 'carboplatin', 'disappointing', 'vitro65', 'toxicities', 'pki116', 'gw2016', 'ekb', '569', 'wyeth', 'ayerst', '1033', 'pd183805', 'pfizer', 'therapy58', 'internalization', 'cells67', 'inc', 'c225', 'imclone', 'bristol', 'myers', 'squibb', 'cetuximab', 'erbitux', 'tolerated68', 'trastuzamab', 'months69', 'licensed', 'expensive', 'patent', 'protection', 'expires', 'distantly', 'achievement', 'elimination', 'neighbours', 'reliance', 'harming', 'oncogene72', 'achille', 'unclear73', 'isoprenylation', 'rce1', 'endoprotease', 'isoprenylcysteine', 'icmt', 'methylates', 'ras74', 'farnesylthiosalicylic', 'ras75', 'aptamer', 'application76', 'culture77', 'impractical', 'hydrolyse', 'diaminobenzophenone', 'phosphoramidate', 'dabp', 'hydrolysed', 'gtp78', 'imagined', 'gamma', 'learnt', 'bifurcation', 'talk', 'deluding', 'thinking', 'surely', 'rationally', 'scrutinized', 'fundamental', 'reliant', 'successors', 'inevitably', 'unstable', 'cue', 'battle', 'infectious', 'arsenal', 'clinicians', 'revise', 'discarded', 'labour', 'cautiously', 'optimistic', 'nfl', 'inconsistent', 'underex', 'pressed', 'neurofibrosarcoma', 'myclogenous', 'jcml', 'drome', 'neur', 'leukaemias', 'fre', 'quentilly', 'recessive', 'conforma', 'catalyze', 'accel', 'erating', 'accelerates', 'physiologically', 'gtp16', 'connection', 'compli', 'cated', 'uncommon', 'microinjection', 'transforma', 'lines20', 'straightforward', 'demon', 'strate', 'cases23', 'allele9', 'marrows', 'malig', 'conspicuously', 'lines25', 'chil', 'dren', 'lorm', 'll', 'f1', 'ne6', 'rrti', 'nz', 'oi', 'ill', 'tim', 'devia', 'nalysis', 'ermeabilized', 'iduals', 'triangles', 'ctivities', 'ure', 'pl20', 'activat', 'cryo', 'gent', 'maltoside', 'ddm', 'intro', 'duced', 'preloaded', 'oflabelled', 'mononu', 'mal', 'devel', 'oped', 'myelodysplastic', 'completin', 'astrocy', 'toma', 'uniformly', 'thin', 'radioanalytic', 'dp', 'tentily', 'ia', 'cfu', 'granulocyte', 'macrophage', 'ship', 'hematopoiesis', 'gestational', 'dl', 'dl3', 'defects9', 'dl2', 'cellulose', 'recombi', 'cul', 'tures', 'csp', 'litter', 'mates', 'mil', 'fers', 'elevations', 'mgm', 'rofibromin31', 'exper', 'iments', 'reproduce', 'rare13', 'paradox', 'differentia', 'entiation32', 'unproved', 'diamonds', 'homozy', 'gous', 'nft', 'simpler', 'sys', 'tem', 'func', 'nega', 'tively', 'lines14', 'reduc', 'immuno', 'substrate35', 'cells28', 'hypersensitivity', 'line36', 'ligands34', 'impli', 'transplanted', 'fatal', 'jcml39', 'infiltration', 'cells40', 'dif', 'ferentiate', 'argues',

'blastic', 'jcml41', 'lig', 'extrapolate', 'unidentified', 'neurotropin', 'in depen', 'dent', 'ngf42', 'ations', 'disposed', 'somehow', 'accumulat', 'ogeni c', 'firm', 'therapies43', 'relentlessly', 'surplus', 'cali', 'fornia', 'clin ician', 'caring', 'exami', 'nation', 'described27', 'mgc12', 'dithiothreito l', 'pefabloc', 'peptin', 'ambient', '2s', '0c', 's00', 'c02', 'cx32p', 'war m', 'resus', 'pended', 'permeabilization', 'pipes', 'kcl', 'cac12', 'units', 'strep', 'tolysin', 'difeo', 'tube', 'appropri', 'ate', 'aliquots', 'yl3', '2s9', 'elution', 'pelleting', 'oso', 'dode', 'cyl', 'precleared', 'iof', 's epharose', 'natants', 'lof', 'soo', 'ooso', 'radioactive', 'elut', '3s', 'chr omatographic', 'separation', 'pei', 'lici', 'uti', 'lized', 'committees', 'in diana', 'califor', 'nia', 'mated', 'pregnant', 'sacrificed', 'cervical', 'dis location', 'gestation', 'abdominal', 'incision', 'embryon', 'ic', 'dissecte d', 'iscove', 'supple', 'mented', 'pulled', 'gauge', 'needle', 'syringe', 'pu shing', 'thoroughly', 'plating', 'microscopy', 'ferent', 'institutions', 'investigator', 'unaware', '314', 'microsatellites', 'endpoints', 'unambiguousl y', '162200', '3500', 'axillary', 'freckling', 'tibial', 'translated', 'repea ts', 'rep', 'zeste', '015355', 'pseudogene', 'recurring', 'encompass', 'imprecise', 'reliable', 'distinction', 'microsatellite', 'france', 'grant', 'coordinator', 'henri', 'mondor', 'teil', '561', '1083', 'donors', 'leukocytes', 'n anodrop', 'coleman', 'orlando', 'typing', 'extragenic', 'genescan', '2500', 'rox', 'urea', '310', '672', 'practical', 'exponential', 'alb', '4q11', 'calibrator', 'subtracting', 'computer', 'oligo', 'blastn', 'nr', 'triplicates', '689', 'mer', 'multistep', 'spacing', 'spanned', 'mismatches', 'paralogous', '44k', 'genomically', 'jmd', 'march', '15k', 'june', 'alui', 'rsai', 'priming', 'cy5', 'cy3', 'microcon', 'ym', 'sandwiched', 'rotations', 'g2565ba', 'qc', 'fine', '712', '905', '3130', 'automatic', 'regularly', 'noise', 'metric', 'snr', 'herein', 'junction', 'exactly', '836', '791', '17q', '943', '375', '780', '167', 'numbered', '1185', '175', '609', '630', '794', 'zoom', 'entirely', 'noteworthy', 'interspersed', '307', 'joining', 'nhej', 'b2', 'blmh', 'bleomycin', 'hydrolase', 'notoriously', 'laborious', 'stepwise', 'yac', 'salsa', '082', 'mrc', 'holland', 'inexpensive', 'fist', '637', '493', '998', 'al18', '821', 'prospective', 'erks', 'p74raf', 'pp60src', 'myristoylate d', 'p21rasgtp', 'bring', '340', '341', 'p21ras', 'tyr340', 'recruiting', 'words', 'nematode', 'worm', 'fly', 'vertebrate', 'elk', 'gille', 'janknecht', 'phas', 'haystead', 'pause', 'payne', 'alessi', 'guan', 'cr1', 'cr2', 'cr3', 'heidecker', 'finger', 'fabian', 'wane', 'posttranslational', 'williams', 'kikuchi', 'leever', 'fantl', 'freed', 'irie', 'activations', 'iras', 'r891', 'destroys', 'p21ras', '3136', 'press', 'traverse', 'wartmann', 'davis', 'stokoe', 'pp60src', 'membrane localized', 'activators', '527', 'y527fsrc', 'cartwright', '1987', '9e10', 'evan', '1985', 'morrison', 'retarded', 'electrophoretic', 'samuels', 'gsterk2', 'arg891leu', 'r891raf', 'ib', 'cterminal', 'ras4b', '6kcaax', 'r891raf6kcaax', 'translocate', 'gly2', 'gly2ala', 'buss', 'sefton', 'y527fsrcg2a', 'operation', 'r1751', '18a', 'baltimore', 'supertifurg a', 'mam', '3137', 'lc', 'cc', 'rafexpressed', 'gibco', 'brl', 'mdck', 'micro injected', 'y13', '3138', '0t', 'hl', 'rni', 'lo', 'cd0', '0cl', 'ar0c', '80000', '60000', '40000', 'xec', 'td', 'y527fsrcw1', 'y527fsrci751', '325', '648', 'araf6kcaax', 'arafdd', 'aspartic', 'arafdd6kcaax', 'arafff', 'arafff6kcaax', 'kras4b', 'malian', 'howe', 'rasindependent', 'route', 'gibbs', '3139', 'rafempty', 'r891rafcaax', 'phosphothreonine', 'phosphoserine', 'phosphorlma ger', 'phosphoamino', 'nps', 'y5', '27fs', 'fi', 'traf', 'k6c', 'aa', 'fq', 'immunoprecipitated', 'electophoresed', 'pvdf', '32plabelled', 'ninhhydrin', 'raseexpressing', '3140', 'pps', '1y', 'y527fs', 'ip', 'rafy34ody341d', 'enc', 'uz', 'uo', 'tt', 'rc', 'ln', 'lf', 'pp60y527fsrc', 'reinforces', 'tyr', 'y340fy340f', 'rafff', 'raffdd', 'rafy34ofy341f', '1y340fy341f', 'y340fy341', 'tyr341', 'y340dy341', 'vras', 'y340fy341f', '1y340dy341d', '3141', 'y340dy34', 'separable', 'y340dy341d', 'arafddcaax', 'preferred', 'degenerate', 'songyan g', 'translocating', 'paterson', 'situation', 'pl2orasgap', 'boguski', 'mccor

mick', 'y527fsrceexpressing', 'wholesale', 'sithanandam', 'mark', '3142', 'untransformed', 'jielb', 'iran', 'ras5', 'cellia', 'cowley', 'il2', 'turner', 'stimulus', 'cdnas', 'pefplink', 'pefbos', 'mizushima', 'nagata', 'elongation', 'ggeqkliseedl', '3143', '5x', 'gg', 'psa', 'kc1', 'mercaptoethanol', 'na3vo4', 'um', 'microcystin', 'ug', 'benzamidine', 'aminoethyl', 'benzene', 'sulfonyl', '160', 'sheared', 'microfuge', '1251', 'brij35', 'sg', 'mercaptoethanol1', 'glutathionesepharose', 'pharmacia', 'pa', 'glutathionine', 'mgci2', 'spotting', 'p81', 'phosphoric', 'subtraction', 'blank', 'dialysed', 'mci', 'jg', 'immobilon', 'excised', 'kamps', 'microinjections', 'workstation', 'carl', 'oberkochen', 'delineated', 'scalpel', 'blade', 'petri', 'formaldehyde', 'quenching', 'ascitic', 'furth', '1982', 'fluorescein', 'isothiocyanate', 'conjugates', '389', 'phenotyped', 'official', 'journal', 'dysmorphism', 'overgrowth', 'liss', 'mim', 'carey', 'pseudogenes', 'micrornas', 'supp', 'pez', 'correa', 'jenne', 'raedt', 'riva', '2004a', 'plexiform', 'friedman', 'expressivity', 'upadhyaya', 'castle', '2004b', 'cardiovascular', 'spiegel', 'utilising', 'adjusting', 'entitled', 'constitution', 'methodologies', 'spred1', 'sprouty', 'evh1', 'cardiff', 'unavailable', 'fulfilled', 'fourteen', 'scoliosis', 'sli', 'lamp', 'cranial', 'height', 'circumference', 'coarse', 'occiput', 'asymmetry', 'forehead', 'frontal', 'bossing', 'ptosis', 'downslanting', 'eyes', 'eversion', 'eyelid', 'folds', 'nasal', 'bulbous', 'ears', 'malar', 'hypoplasia', 'philtrum', 'micrognathia', 'chin', 'posterior', 'hairline', 'scholar', 'repeating', 'motor', 'skills', 'compression', 'neurofibroma', 'cerebrovascular', 'artery', 'stenosis', 'tibia', 'vertebrae', 'wing', 'xanthogranuloma', 'meaningful', 'edgar', 'query', 'cgi', 'usa', 'covariates', 'bonferroni', 'nominal', 'statview', 'cary', 'customised', 'estimation', 'know', 'publicaton', 'dub', '830', '870', '169', 'pb', 'huson', 'birch', '004', 'overrepresented', 'explainable', 'prepupal', 'covariate', 'chi', 'ad', 'hoc', 'luca', 'lcr', 'belong', 'paralogs', '17q24', '19p13', 'douglas', 'medial', 'impossible', 'claimed', 'ruling', 'abolish', 'signifi', 'cance', 'identifi', 'vuss', 'easy', 'defi', 'cient', 'unclassifi', 'confi', 'ned', 'brct', 'geneticists', 'classifi', 'cation', 'discov', '1142', 'affi', 'liations', 'centre', 'taconicartemis', 'cologne', 'cancerdiscovery', 'aacrjournals', 'klijn', 'south', 'jos', 'jonkers', 'plesmanlaan', '1066', 'cx', '5122000', '5122050', 'mail', 'niki', 'doi', '1158', '2159', '8290', 'cd', '0094', 'hboc', 'enigma', '273', '920', 'nontruncating', 'cosegregation', 'cooccurrence', 'calculate', 'preexisting', 'diffi', 'cult', 'elegant', 'technically', 'demanding', 'fulfi', 'noncancelous', 'sdm', 'rmce', '2017', 'onlinefirst', 'july', '1144', 'october', 'bouwman', 'selectable', 'sco', 'brca1sco', 'loxp', 'split', 'rosa26', 'hydroxy', 'oht', 'effi', 'specifi', 'frt', 'flp', 'ciency', 'flpmiated', 'variegation', 'modifi', 'dutch', 'belgian', 'r26', 'profi', 'singlecopy', 'phosphoglycerate', 'pgk', 'hbrca1', 'r26creert2', 'brca1sco5', 'r26rmce', 'r26hbrca1', 'zsgreen', 'hyg', 'neor', 'mbrca1', 'neos', 'puros', 'puror', 'ef1', 'neo', '2013cancer', '1145', 'm1400v', 'l1407p', 'm1411t', 'palb2', 'bic', 'nhgri', '185delag', '5382insc', 'y105c', 'r866c', 'e1250k', 'designation', 'agency', 'verifi', 'rmed', 'artifi', 'cial', 'ef1a', 'rst', 'switched', 'shut', 'workfl', 'es', 'outline', 'timecourse', '1146', 'switching', 'sulphorhodamine', '224', '1366', 'interstrand', 'crosslink', 'icl', 'stressed', 'reversion', 'platinum', 'functionality', 'resazurin', 'tted', 'constrained', 'squared', 'rse', 'bayesian', 's4f', 'm18t', '53t', '69delag', 'k45q', '133a', 'c61g', '181t', 'c64g', '190t', 'd67y', '199g', '314a', 'n132k', '396c', 'p142h', '425c', 'l147f', '441g', 'l165p', '494t', 'r170w', '508c', 's186y', '557c', 'v191i', '571g', 't231m', '692c', 'd245v', '734a', 'l246v', '736t', 'v2711', '811g', 's308a', '922a', '923g', 'l358r', 'c360r', 'e362h', '1073t', '1078t', '1084g', '1086g', '1668f', '2002c', 'd695n', '2083g', 'p7981', '2393c', 'n810y', '2428a', 't826k', '2477c', '1147', 'r841q', '2522g', 'y856h', '2566t', '2596c', 's988a', '2962t', 'm1008i', '3024g', 'e1060a', '3179a', 's1101n', '3302g', 'k1110del', '3328', '3330delaag', 's1140g', '3418a', 'e1214k', '3640g',

'n1236k', '3708t', '3748g', '11267s', '3800t', 'e1282v', '3845a', 's1297de1', '3891', '3893delttc', 's1301r', '3903t', 'e1346k', '4036g', 'v1378i', '4132g', '4198a', '4220t', '4232t', 'r1443g', '4327c', 's1448g', '4342a', 's1486c', '4456a', 's1497a', '4489t', 'v1534m', '4600g', 'r1589p', '4766g', 'm1628t', '4883t', 's1651p', '4951t', 's1651f', '4952c', 'm1652i', '4956g', 's1655f', '4964c', 'h1686r', '5057a', 'h1686q', '5058t', 'v1688del', '5062', '5064delgtt', 't1691i', '5072c', 'r1699w', '5095c', 'r1699q', '5096g', 'g1706e', '5117g', 'g1706a', 'a1708e', '5123c', 'w1718c', '5154g', 't1720a', '5158a', 'e1735k', '5203g', '1148', 'stressing', 'irradiation', 'diminish', 'g1770v', 'classi', 'usefulness', 'questionable', 'dsb', 'unrepaired', 'v1736a', '5207t', 'd1739g', '5216a', 'd1739v', 'h1746q', '5238c', 'r1753t', '5258g', '5266dupc', '5291t', 'c1767s', '5300g', '5309g', 'w1782c', '5346g', 'a1789t', '5365g', 'e1794d', '5382g', '5411t', 'p1812r', '5435c', 'w1837r', '5509t', 'h1862l', '5585a', '1149', 'waterfall', 'chart', 'pinheads', 'encircled', '1646', '1706', '1766', '1826', 'mdlsalrveevqnvinamqkilecpiclel', 'vnkrmsmvsgltpeefmlvykfarkhhitlnliteetthvvmktdaefvcert1kyfl', 'ikepvstckdhifckfcmlk1lnqkkgpsq', 'giaggkwvvsyfwvtqsikerkmlnehdfevrgdvnngrnhqgpkraresqdrkifrgle', 'cplcknditkrsqliqestrfsqlveellkii', 'cafqltdtgleyanansynfak', 'iccygpftnmpdglewmvqlcgasvvkelssftlgvgvhpivvqpawtedngfhaig', 'qmceapvvtrewvldsvalyqceldtylipqiphshy', '1150', 'shuttled', 'pim1', 'scei', 'meganuclease', 'mcherry', 'uorescent', 'dsbs', 'rming', 'cantly', 'parp1', 'isogenic', 'olaparib', '1151', 'cytotoxicity', 'cant', 'damaged', 'alleviates', '53bp1', 'mask', '1152', 'recombineering', 'consuming', 'easier', 'transgenics', 'integrations', 'concatemers', 'swiftly', 'obviates', 'belgium', 'causally', 'ambivalent', 'refl', 'ect', 'aws', 'embryoid', 'misclassifi', 'applies', 'l1407', '1153', 'facilitatin', 'fancd2', 'catalyzing', 'minigene', 'instrumental', 'transacting', 'implementable', 'neoadjuvant', 'ct2', 't4an0m0', 'downstaging', 'pt0', 'ptis', 'cystectomy', 'urologic', 'community', 'ercc1', 'cdk7', 'ddb1', 'xpa', 'xpc', 'xeroderma', 'pigmentosum', 'xp', 'modulate', 'ercc2', 'responders', 'nonresponders', 'turbt', 'contamination', 'ranges', '130', '0003', 'fidelity', '277', 'afforded', '007', 'hochberg', 'binomial', 'mutsigcv', 'dashed', 'denotes', 'reaches', 'unselected', 'guo', 'chinese', '429', 'queried', 'stick', 'scc', 'archaeabacterial', '3crv', 'cockayne', 'trichothiodystrophy', 'ttd', 'anova', 'gm08207', 'coriell', 'plx304', 'lane', 'expresses', 'rescues', 'repairs', 'aducts', 'uv', 'clonogenic', 'd609g', 'ercc2s', 'spreads', 'v242f', 'genotoxic', 'unwinds', 'haploinsufficient', 'delays', 'inadequate', 'rad3', 'operative', 'justify', 'nonresponding', 'played', 'comorbidities', 'exceptional', 'rarity', 'extraordinary', 'chemotherapies', 'decision', 'designs', 'extravesical', 'prechemotherapy', 'irb', 'dana', 'farber', '021', '334', 'memorial', 'slloan', 'kettering', '076', '025', 'genitourinary', 'barletta', 'signoretti', 'dfci', 'ahmadie', 'ffpe', 'nonurothelial', 'source', 'certified', 'qiaamp', 'shearing', 'barcoded', 'adaptors', 'v2', 'multiplexed', 'bam', 'dbgap', 'phs000771', 'file', 'firehose', 'infrastructure', 'managed', 'executed', 'content', 'depthofcoverage', 'oncotator', 'orthogonal', 'fluidigm', 'invalid', 'misalignment', 'artifacts', 'viewable', 'tgfb1', 'depdc4', 'aggregated', 'polyphen2', 'dinucleotide', 'tabulated', 'rateresponders', 'ratenonresponders', 'read', 'ifc', 'chips', 'tagging', '3770', 'rev', 'c1', 'bravo', 'barcodes', 'sequencers', 'miseq', 'chemistry', 'nonreference', 'attb', 'attp', 'frames', 'pdonr', 'recombined', 'pentr', 'attl', 'attr', 'top10', 'anion', 'lentivirus', 'envelope', 'supernatants', 'sv40', 'pseudodiploid', 'ercc2r683w', 'blastocidin', 'containment', 'coriell', 'repositories', 'camden', 'nj', 'authenticated', 'serially', 'celltiterglo', 'luminescence', 'biotek', 'sigmoidal', 'nunc', 'irradiator', 'vigorously', 'dry', 'irradiated', 'colcemid', '075', 'wright', 'stain', 'metaphases', 'chromatid', 'rings', 'tri', 'quadraradial', 'di', 'tricentric', 'thawed', 'navo4', 'suspensions', 'colorimetry', 'rabbit', 'chemiluminescent', 'perkinelmer', 'film', 'pik3cb', 'pik3cd', 'tailor', 'commences', 'pip2', 'pip3', 'spurred', 'nonredundant', 'unnecessary', 'i

soformderived', 'pik3caare', 'dld1', 'bt549', 'u87mg', 'ptendeficient', 'wrote', 'conflict', 'statement', 'employees', 'biomedical', 'pnas', 'dmmitri', 'wiederschain', 'sanofi', 'aventis', 'quai', 'jules', 'guesde', '94403', 'vity', 'sur', 'seine', '0802655105', 'dcsupplemental', 'academy', '1073', 'september', '13057', '13062', 'doxycyclineinducible', 'regulatable', 'si', 'measurable', 'h1047r', 'e545k', 'wellcharacterized', 'pik3cawould', 'phosphofoxo', 'pik3caknockdown', 'shRNAexpressing', 'cleavage', 'dox', 'p473akt', 'pfox03a', 'pp70s6', 'pfox01', 'pp85s6', 'p308akt', 'ppras40', 'ctrl', 'x105', 'scramble', '246', 'foxo3a', 's6k', 'doxycycline', 'semisolid', 'hoechst', '33342', 'photographed', 'nikon', 'imagepro', '13058', 'wee', 'correspondingly', 'unperturbed', 'pik3cain', 'implanted', 'continuously', 'drinking', 'gavage', 'multiplied', 'phosphoakt', 'pik3cadid', 'dispensable', 'tetr', 'mts', 'attach', '13059', 'p110d937a', 'converting', 'afg', 'counterpart', 'tgx221', 'ic87114', 'ps6', '456', 'implant', '1200', '1400', 'libitum', 'qd', 'freely', 'caliper', 's236', 'btranscript', 'p473', '13060', 'tgx256', 'indication', 'alters', 'p308', 'd937a', '123456', 'pgsk3', 'tumorsin', 'calipers', '13061', 'unabated', 'incremental', 'coexist', 'a2780', 'penetrant', 'dictate', 'broadsht', 'retard', 'thoughtfully', 'mcrpc', 'biopsies', 'spondin', 'zbtb16', 'plzf', 'atm', 'graphical', 'unlabelled', 'yearly', 'adt', 'abiraterone', 'bono', 'ryan', 'enzalutamide', 'beer', 'scher', 'cooper', 'pflueger', 'e26', 'hieronymus', 'lalonde', 'taplin', 'beltran', 'gundem', 'hong', 'mehra', 'allen', '2014a', 'piloted', 'enrollment', 'longitudinal', 'centralized', 'cerami', 'robinson', 'thorvaldsd', 'ttir', 'adequately', 'morrissey', 'rosenberg', 'su2c', 'pcf', 'idt', 'schema', '189', 'zygosity', 'epstein', 'mismatch', 'mlh1', 'ms h2', 'rows', 'histogram', 'cnvs', 'colors', '13q', '16q', 'chr11q23', 'mcprc', 'pie', 'diameters', 'chimeras', '247', 'fli1', 'rspo2', 'representation', 'kiaa1549', 'ex10', 'ex9', '020910', '004333', 'ubn2', 'ex3', 'ex11', '173569', 'lsm12', '152344', 'golga4', 'ex17', 'ex8', '001172713', '002880', 'sm12', 'anticodon', 'tkc', 'hun', 'hpc2', 'ubinuclein', 'ubn', 'ab', 'smc', 'atpases', '440', '550', 'kmt2d', 'gnas', 'snv', 'putatively', 'roychowdhury', '2014c', 'frameworks', 'modifier', 'ncor1', 'geng', 'agonism', 'flutamide', 't878a', 'w742c', 'glucocorticoids', '1702h', 'exposures', 'q903h', 'carreiraa', 'v7', 'antonarakis', 'ntad', 'dbd', 'lbd', 'prad', 'referring', 'ncor2', '5schematics', 'schematics', 'c2', 'tad', 'fh', 'math', 'meprin', 'doamin', 'btb', 'poz', 'narrowed', 'cao', 'kikugawa', 'hsieh', 'chr11', 'marks', 'wa42', 'wa4', 'e17k', 'pik3', 'pk3ca', 'tbl1xr1', 'cds', 'acpp', 'voeller', 'rnf43', 'znrf3', 'adrenocortical', 'assi', 'giannakis', 'seshagiri', 'organoid', 'porcupine', 'armadillo', 'eb1', 'grhl2', 'scope', 'cdk4', 'finn', 'comstcock', 'keys', 'actionability', 'mateo', 'helc', 'ob', 'frap', 'trrap', 'pik3c', 'brc', 'hypermutated', 'fanca', 'rad51b', 'rad51c', 'hypermutation', 'corroborating', 'msh6', 'pritchard', 'implement', 'feasibility', 'indifferent', 'warranting', 'schwartz', 'specificin', '106th', 'palanisamy', 'prospectively', 'kaufman', 'weston', 'warrants', 'iterative', 'sampling', 'matching', 'multiregional', 'asangani', 'uveal', 'choroid', 'plexus', 'ciliary', 'iris', 'alterations6', 'metastasize', 'ota', 'intradermal', 'hyperpigmentation', 'periorbital', 'dermis', 'microscopical', 'nevi', 'gnaq', 'gna11', 'melanocytic', '209', 'q209', 'couple', 'guanosine', 'hydrolyzed', 'diphosphate', 'locked', '227', 'q227', 'arginines', 'r201', 'i2', 'broadened', 'revisited', 'retrieved', 'nonobese', 'diabetic', 'immunodeficiency', 'nod', 'scid', 'melan', 'cells22', 'nontumorigenic', 'nejm', 'itable', '713', '2table', 'r183', 'ctg', 'ccg', 'cta', 'ctt', 'q2091', 'cgc', 'tgc', 'gtc', 'agg', 'att', 'ultraviolet', 'taa', 'tta', 'segmental', 'sparsely', '118', '048', '1figure', 'epithelioid', 'aberrations7', 'requisite', '2figure', 'immunocompromised', 'serve d', 'latency', 'effected', 'countered', 'originating', 'endothelin', 'dismiss', 'enucleated', 'obscured', 'oncoproteins', '254890', 'toxin', '922', 'histopathological', 'exhaustive', 'peculiar', 'schwann', 'solitary', 'epidemiologic', 'occupational', 'fair', 'complexion', 'irides', 'g12vkras', 'hyper', 'pi

gmentation', 'cardinal', 'invade', 'destroy', 'muscles', 'gnefs', 'returning', 'mammals', 'whilst', 'g12', 'g13', 'trapping', 'specialised', 'pigment', 'hair', 'tone', 'societies', 'doubling', 'cgp', 'predominate', 'g12vnras', 'inefficient', 'g12vhrras', 'tyrosinase', 'ulcerated', 'colonized', 'epidermis', 'subcutis', 'inefficiently', 'license', 'authority', 'lsl', 'intravital', 'allograft', 'krasg12v', 'flanks', 'np40', 'heidelberg', 'ncl', 'novocastra', 'newcastle', 'llc', 'separately', 's100', 'ki67', 'labvision', 'citrate', 'ph6', 'envision', 'aec', 'chromogen', 'elite', 'abc', 'ak', '5001', 'ap', 'sera', 'immunoresearch', 'mp', '606', 'italy', '157', '1207', 'goat', 'backs', 'tails', 'dome', 'superficial', 'immersed', 'collagenous', 'vague', 'nests', 'junctional', 'pleomorphism', 'atypia', 'negligible', 'macroscopically', 'arrowheads', 'occasional', 'extensions', 'overt', 'infiltrating', 'destructive', 'melanin', 'laden', 'scant', 'hpfs', 'borderline', 'perianal', 'necessitated', 'sacrifice', 'asymmetrical', 'ostensibly', 'inclusions', 'characterisation', 'summarised', 'grossly', 'infiltrative', 'pleomorphic', 'promptly', 'limits', 'took', 'stabilised', 'hmb45', 'gp100', 'mitf', 'antigens', 'ihc', 'dc t', 'eyelids', 'deposits', 'skeletal', 'fibres', 'tail', 'veins', 'recipient', 'absorbed', 'ingested', 'periphery', 'juxtaposition', 'circumstantial', 'vein', 'injections', 'ulceration', 'proliferating', 'governs', 'microenvironments', 'extrinsic', 'succumb', 'doxycyclin', 'transgene', 'adapt', 'powell', 'g12dkras', 'darkening', 'rafts', 'noonan', 'cfcs', 'costello', 'elusive', 'catalyzed', 'switches', 'vetter', 'wittinghofer', 'accelerate', 'scheffzek', 'ahmadian', 'herrmann', 'fiegen', 'mobile', 'sprang', 'cylinders', 'barb acid', 'kranenburg', 'malumbres', 'seeburg', '163950', '218040', 'faciocutaneous', '115150', 'craniofacial', 'schubbert', '2007a', '117', 'denayer', 'estep', 'gremer', 'gripp', 'sol', 'church', 'cirstea', 'k5', 'v14', 'p34', 'i36', 't58', 'g60', 'v152', 'd153', 'f156', 'cfc', 'carta', 'kratz', 'nava', 'niihori', '2007b', 'zenker', 'pathophysiological', 'gelb', 'tartaglia', 'tidyman', 'rauen', 'k5n', 'v14i', 'q22e', 'q22r', 'p34l', 'p34r', 't58i', 'g60r', 'd153v', 'f156l', 'atg', 'mutnomen', 'peyfp', 'xho1', 'bamh1', 'ptachras', 'tucker', '333', '1198', '1531', '4t', 'ecori', 'cdc25', 'lenzen', 'ptac', 'gppnhp', 'mantgdp', 'mantgtp', 'mantgppnhp', 'hemsath', '30mm', '5mm', '3mm', 'dtede', 'photophysics', 'apparatus', '10mm', 'kpi', 'dithioerythritol', 'fluoromax', 'horiba', 'jobin', 'yvon', 'fluorimeter', '366nm', '450nm', 'constants', 'kobs', 'fitting', '32gtp', 'fitted', 'equations', '30mmtris', '10mmkpi', '5mmmgcl2', 'exchanged', 'load', 'unmonitored', 'hightech', '100mm', 'titrate d', 'quadratic', 'monkey', 'herbrand', 'sparks', 'ras10', 'danvers', '221', 'phosphor', 'thr202', 'tyr204', 'ser473', 'kraswt', 'milburn', '4q21', 'pai', '1ctq', 'resolutions', 'p120rasgap', '1wq1', 'margarit', '1nvv', 'nassar', '1c1y', 'pacold', '1he8', 'byr2', '1k8r', '1lfd', 'bunney', '2c51', 'e153', 'e153v', 'rasg12v', 'rasf281', 'reinstein', 'permissible', 'deduced', 'nearby', 'edge', 'buried', 'rasg60r', 'dxxg60qe', 'gasper', 'faces', 'tremendously', 'rasv14i', 'rasq22e', 'rasf156l', 'fluorescently', 'f281', 'completeness', 'concerning', 'g12v', 'nonhydrolyzable', 'analog', 'slower', 'pechlivans', 'dissociations', 'acceleration', 'abolishes', 'insensitivity', 'rasp34', 'rasg60', 'rasp34l', 'rasp34r', 'rast58i', 'capabilities', 'hydrolyze', 'stone', 'krasv14i', 'krasq22e', 'krasf156l', 'perceivably', 'krasp34r', 'krasg60r', 'pmek1', 'perk1', 'pakt', 'disabled', 'interactive', 'interfered', 'and4d', 'compensated', 'krasp34l', 'tuned', 'disturbs', 'multifaceted', 'tuning', 'asseessment', 'krask5n', 'krast58i', 'krasd153v', 'nrast50i', 't50', 'inspecting', 'abankwa', 'superposed', 'krasq22r', 'krasq22k', 'tsukuda', 'pathogenic', 'guanidinium', 'coordination', 'quilliam', 'ineffectiveness', 'hereby', 'e37', 'invariant', 'eberth', 'corrupt', 'ford', 'sung', 'g60a', 'hwang', 'hrasg60a', 'damps', '1884', '1886', 'karnoub', 'weinberg', 'quinlan', 'settlemen', 'causal', 'expressions', 'handicaps', 'aoki', 'capacities', 'g13d', 'crc', 'sw48', 'cecum', 'macroscopic', 'lymphatic', 'chemokine', 'cxcr4', 'emboli', 'integrin', 'vegfa', 'angiopoietin', 'intravasation', 'nicolantonio', 'lo

pez', 'bardelli', 'gly', 'asp', '4268', 'anoikis', 'immunosuppressed', 'secti onnext', 'recombinants', 'adeno', 'aav', 'paisley', 'penicillin', 'weighing', 'charles', 'river', 'housed', 'sterile', 'sterilized', 'creu', 'sant', 'pau', '700', 'disaggregated', 'killed', 'pain', 'necropsy', 'observers', 'mm2', 'ne crotic', 'keratin', 'carpinteria', 'dilutions', 'integrins', 'chemicon', 'atl anta', 'ga', 'abgent', 'diego', 'cambridge', 'cd34', 'ready', 'franklin', 'la kes', 'contrasted', 'multiplication', 'ebioscience', 'ibm', 'armonk', '190', '188', '265', '1201', '1395', '964', 'necrosis', 'dissemination', 'macro', 'bigger', 'krasg12d', '019', '037', 'submucosal', 'buds', 'magnified', 'snai l', '158', '1937', 'colonization', 'micrometastasis', 'adhesion', 'unveil', 'modestly', '039', '4g', 'bulk', '033', '022', '045', 'conveyed', 'engineering', 'val12', 'displays', 'sw480', 'collagen', 'vitronectin', 'lm', 'urokinase', 'plasminogen', 'upa', 'upar', 'crosstalk', 'suggestion', 'lymphovascula r', 'extravasation', 'micrometastases', 'paracrine', ' sdf1', 'migrate', 'hypoxic', 'permeability', 'disregulated', 'loosening', 'microvessel', 'grows', 'entities', 'adopted', 'k117', 'homozygosity', 'shortly', 'panitumumab', 'waiver', 'authorization', 'dneasy', 'iplex', 'maldi', 'tof', 'spectrometer', 'request', 'ms', 'mislabeling', 'misidentification', '244k', 'segmentation', 'a0', 'coordinates', 'hg18', '2xmyckras4b', 'quickchange', 'xlII', 'temecula', 'bab', 'pressurized', 'ventilated', 'cages', 'compliance', 'iacuc', 'injecting', 'basement', 'formulated', 'hydroxypropyl', 'methylcellulose', 'spectrometri c', '322', 'a146v', 'k117n', 'q22k', 'e31k', 'curative', 'followup', 'pairwise', '07', 'hek', '293ft', '351', 'mislabelling', 'specially', 'cccl', '281', 'ls1034', 'listing', '014', 'hypertriploid', 'isochromosome', 'variably', '100nm', 'cccl18', 'purely', 'cytostatic', 'overwhelming', 'definitive', 'meaningfully', 'agencies', 'underrepresentation', '698', '449', 'mutagenic', 'insults', 'nalm6', 'p120', 'arg789', 'gapp120', 'g12s', 'reluctance', 'advocate', 'accelerating', 'facilitating', 'lethals', 'fate', 'follicles', 'crypts', 'bronchiolo', 'convey', 'permissiveness', 'overrides', 'immortal', 'prototypic', 'nonreceptor', 'phosphoinositol', 'pdk', 'opening', 'basically', 'obscure', 'unveiling', 'manipulate', 'opened', 'meuwissen', 'knockin', 'steroids', 'bicistronic', 'immortalization', 'serrano', 'ires', 'friedrich', 'soriano', 'hydro', 'flanked', 'hygromycin', 'imagefigure', 'polyadenylation', 'xbai', 'sal i', 'kbp', 'triangle', 'endonuclease', 'c57bl', 'blastocysts', 'backcrossed', 'transmission', 'midgestation', 'initiator', 'tga', 'polyadenosine', 'e13', 'crosses', '266kb', 'transcriptionally', 'ert2', 'ert', 'passages', 'kinetic', 'steady', 'cook', 'elicit', 'permanent', 'retroviruses', 'ras4a', 'cooperating', 'e1a', 'p19arf', 'p21cip', 'p25', 'adenovirus', 'ruley', 'karyotyping', 'sky', 'acentric', 'dicentric', 'tetraploid', 'breathing', 'hyperplasias', 'surfactant', 'spc', 'cc10', 'bronchioles', 'histiocytic', 'anal', 'papillomas', 'sp', 'cdk4r24c', 'hemangiosarcomas', 'ducts', 'tall', 'mucin', 'temporally', 'prerequisite', 'harderian', '4oht', '8h', 'testicular', 'seminiferous', 'tubes', 'bowman', 'capsules', '8c', '8d', 'cerebellum', '8e', 'hepatocytes', '8f', 'exocrine', '8g', 'depicting', 'convergent', 'unscheduled', 'thymic', 'crypt', 'morphologies', 'ink4', 'sv', 'psk', 'saci', 'stui', 'leon', 'inserting', '560', '552', 'gta', 'bglii', 'sanchez', 'garc', 'nagy', 'lif', 'linearized', '240', 'genepulser', 'gancyclovir', 'annealing', '403', '621', 'analyristo', '106k', 'geneamp', 'polyacrilamide', 'mechanically', 'dakocytomation', 'c12', 'fdg', 'sorted', 'moflo', 'todaro', '1963', 'chloride', 'pbabe', '9272', 'ab3', 'painting', 'incorporating', 'dyes', 'isotonic', 'carnoy', 'fixative', 'specified', 'hogan', 'convincingly', 'mandatory', 'leu19phe', 'lys17asn', 'ala146thr', 'arg164gln', 'outwith', 'personalised', 'cancerresearchuk', 'cancerstats', 'polyps', 'colonic', 'chapelle', 'polyposis', 'hybridisation', 'localised', 'leslie', 'andreyev', 'conlin', 'etienne', 'grimaldi', 'winder', 'vectibix', 'lievre', 'benvenuti', 'khambata', 'freeman', 'karapetis', 'ramos', 'loupakis', 'cutsem', 'mcneill', 'krieken', 'tol', 'transduce', 'influencing', 'focussed', 'locking', 'ellis', 'orita', 'miyakura', 'akagi', 'we

llbrock', 'wan', 'ninewells', 'dundee', 'royal', 'infirmary', 'invited', 'caucasian', 'icd', '153', '154', 'bowel', 'ulcerative', 'colitis', 'diverticular', 'tayside', 'brought', 'theatre', 'wizard', 'southampton', 'venous', 'containers', 'crawley', 'sussex', 'rflp', 'services', 'pef', 'plink', 'pefha', 'gifts', 'professor', 'london', 'queensferry', 'edinburgh', 'q61h', '119f', 'r164q', 'adhere', 'trypsinised', 'cm3', 'watford', 'hertfordshire', 'mlb', 'gentle', 'rocking', 'lds', 'orthovanadate', 'eppendorf', 'hemel', 'hempstead', 'cast', 'whatman', 'maidstone', 'kent', 'merck', 'nottingham', 'dh20', 'nano', 'labchip', 'cxr', 'cyanine', 'ctp', 'cys3', 'quickamp', 'hybridised', 'v9', 'bioconductor', 'visualised', 'java', 'treeview', 'tttaatacttttatgtattcagggtgttg', 'ttaccatcttgcacatctttcttt', 'fam', 'tgatgccttctatacattagtcgaaattcgaaaa', 'tamra', 'warrington', 'deparaffinised', 'vp2000', 'coplin', 'warmed', 'maidenhead', 'meanwhile', '707g18', 'bluegnome', '12q', '89h19', 'lsi', 'wcp', 'pipetted', 'coverslip', 'denature', 'rubber', 'cement', 'seal', 'morning', 'bubbles', 'sealed', 'nail', 'varnish', 'bx60', 'g57t', 'wave', 'lys', 'arg', 'gln', '164', '173', 'localisation', 'equivalence', 'spectrophotometrically', 'subcluster', 'composite', 'ptpre', 'arhgap6', 'ereg', 'hbegf', 'btc', 'ngef', 'ceacam', 'serpinb2', 'pak3', 'eii2', 'creb1', 'tcf4', 'e2f2', 'mmp3', 'glut', '1264', '1233', 'linearity', 'specificities', '1271', '12p12', 'dots', 'hsr', 'dhplc', 'kuklin', 'underestimated', 'mis', 'recommend', 'summarises', 'higashi', 'kong', 'america', 'edkins', 'gelsi', 'boyer', 'tyner', 'tumourigenicity', 'wojciech', 'logically', 'dusps', 'keyse', 'croonquist', 'warmka', 'bloethner', 'transporter', 'yun', 'practise', 'recommendation', 'adjuvant', 'shp2', 'transducer', 'ptp', 'ptyr', 'frees', 'leopard', 'ls', 'conundrum', 'competing', 'affording', 'ptks', 'ptps', 'delicate', 'obesity', 'autoimmune', 'perceived', 'ptpn11', 'underlies', 'nterminus', 'entrance', 'pubs', 'acs', 'dx', '1021', '4151', 'liberating', 'acronym', 'lentiginous', 'electrocardiographic', 'ocular', 'hypertelorism', 'pulmonic', 'genitalia', 'deafness', 'stature', 'nsassociated', 'gof', 'heightened', 'residing', 'lsassociated', 'incite', 'potentiated', 'y279c', 'weakens', 'gab1', 'compensates', 'perturbs', 'engender', 't468m', 'q506p', 'q510e', '21a', 'xhoi', 'neb', 'fulllength', 'purities', 'desalting', 'packed', 'nitrophenyl', 'start', 'hydroxide', 'spectramax', '384', 'spectrophotometer', 'michaelis', 'menten', 'sigmaplot', 'quartz', 'cuvette', 'dephosphorylated', '305', 'ls50b', 'crystallization', 'hanging', 'drops', 'reservoir', 'peg3350', 'kf', 'licl', 'cryoprotectant', 'flash', 'beamline', 'bm', 'photon', 'aps', 'refinements', 'iteratively', 'phenix', 'refine', 'h2o', 'quench', 'endoproteinase', 'autosampler', 'c18', 'waters', 'submerged', 'finnigan', 'gradients', 'steep', 'acetoneitrile', 'electrosprayed', 'ltq', 'peptic', 'xcorr', 'hx', 'hyclone', 'e76k', 'c459s', 'hgf', 'paxillin', 'pico', 'tablet', 'harm', 'rigorous', 'autoinhibition', 'aryl', '0009', '0023', '071', '090', '0034', '0004', '055', 'sprouty1', 'py118', 'py992', 'py589', '1167', '06', 'statistics', 'p21212', 'dimensions', '206', '219', 'deg', 'rmerge', 'reflections', 'rwork', 'rfree', '3996', '3956', '7824', '3984', '186', '068', 'ramachandran', 'disallowed', 'parantheses', 'shell', '930', 'supposed', 'dephosphorylation', 'interdomain', 'three-dimensional', 'proceeded', 'diffraction', 'crystallized', 'superimposed', 'colored', 'wpd', 'superimposition', 'angstroms', 'superimposing', 'penetrates', 'walled', 'occluding', 'rmsds', '418', 'abrogates', 'd61', 'y62', 'eliminates', 'polar', 'sequentially', 'displaces', 'imidazole', 'introduces', 'repulsion', 'y279', 'pushes', 'bridged', 's365', 'oxygens', 'welldefined', 'moves', 'losing', 'abolishing', 'q506', 'n58', 'a72', 'noticeable', 'carboxylate', 'inherently', 'electrostatic', 'existed', 'surfaces', 'attacking', 'nucleophilic', 'ptp1b', 'labels', 'scheme', 'y992', 'phosphorus', 'd425', 'functioning', 'buildup', 'inorganic', 'q510', 'nucleophile', 'ep', 'pentacoordinate', 'r465', 'amides', 'stacking', 'cognate', 'yersinia', 'equatorial', 'a461', 'g464', 'defines', 'electropositive', 'imposes', '464', 't468', 'closure', 'illuminate', 'furnish', 'nsh2', 'competitively', 'hydrogens', 'extents',

'engagement', 'persisted', 'stay', 'residence', 'dephosphorylate', 'twodimensional', '4145', 'dephosphorylating', 'sure', 'normalization', '4146', 'rasopathies', 'alleviating', 'cardiofaciocutaneous', 'transducing', 'evoked', 'zebrafish', '4147', 'pluripotent', 'multidisciplinary', 'mimetic', 'adapters', 'engenders', 'tyrosinephosphorylated', 'tyr542', 'erythroid', 'e2', 'cnc', 'ech', 'protect', 'toxicants', 'carcinogens', 'cytoprotective', 'chemoresistance', 'geared', 'chemoprevention', 'cap', 'collar', 'zipper', 'bzip', 'antioxidant', 'combat', 'harmful', 'xenobiotics', 'traditionally', 'regarded', 'neurodegenerative', 'aging', 'photo', 'fibrosis', 'motohashi', 'yamamoto', 'jeong', 'lau', 'conducive', '2008b', 'neh1', 'neh7', 'conjugation', 'mcmahon', 'tong', '2006b', 'cullinan', 'kobayashi', 'furukawa', 'xiong', 'neh4', 'neh5', 'neh3', 'neh6', 'moi', 'ubcm2', 'conjugating', 'plafker', 'dsgis', 'dsapgs', 'transducin', 'trcp', 'cul1', 'roc1', 'rada', 'chowdhry', 'coactivators', 'chromo', 'chd6', 'coactivator', 'ntoi', 'ch3', 'cbp', 'creb', 'katoh', 'fahl', 'seventh', 'neh', '316', 'repress', 'tramtrack', 'bric', 'brac', 'dgr', 'mulahey', 'itoh', 'ogura', '2006a', 'latch', 'docks', 'unoccupied', 'accepting', 'polyubiquitination', '26s', 'usp15', 'peptidase', 'deubiquitininating', 'deubiquitinates', 'villeneuve', 'cys151', 'evade', 'oxidized', 'sulfenic', 'disulfides', 'adducted', 'electrophiles', 'dinkova', 'kostova', 'hannink', 'thiols', 'attached', 'translocates', 'dimerizes', 'masculoaponeurotic', 'fibrosarcoma', 'maf', 'ares', 'cytoprotection', 'nguyen', 'hirotsu', 'rtgabnnngcr', 'capitals', 'rushmore', 'wasserman', 'hayes', 'gcs', 'specializing', 'gpx', 'metabolizing', 'udp', 'glucuronosyl', '1a1', 'multidrug', 'mrp1', 'gordon', 'battery', 'detoxify', 'carcinogen', 'benzo', 'pyerene', 'gomez', 'burdens', 'azoxymethane', 'osburn', 'khor', 'nitrosobutyl', 'hydroxybutyl', 'amine', 'faheny', 'dimethylbenz', 'anthracene', 'tetradecanoylphorbol', 'hirayama', 'morit o', 'hyperoxia', 'cho', 'rs6721961', 'marzec', 'suzuki', 'singh', 'shibata', '2008a', 'jiang', 'solis', 'partly', 'radioresistance', 'mitsubishi', 'pentose', 'g6pd', 'phosphogluconate', 'pgd', 'transketolase', 'tkt', 'transaldolase', 'taldo1', 'nicotinamide', 'malic', 'me1', 'phosphoribosyl', 'amidotransferase', 'ppat', 'methylenetetrahydrofolate', 'mthfd2', 'isocitrate', 'flux', 'purines', 'building', 'redox', 'antioxidants', 'reddy', 'ishimoto', 'etoposide', 'fluorouracil', 'ohta', 'homma', 'lister', 'shim', 'platin', 'chemotherapeutics', 'ionizing', 'h460', 'carbonyl', 'satoh', 'urethane', 'intermediates', 'padmanabhan', 'ooi', 'gallbladder', 'konstantinopoulos', 'accessing', 'hypermethylated', 'muscarella', 'hanada', 'coppole', 'prognoses', 'mirnas', '200a', 'bartel', 'sangokoya', '271', '377', 'eades', 'bryan', 'yates', 'wakabayashi', 'sequestosome', 'sequesters', 'autophagosomes', 'faraonio', 'gst1', 'stalling', 'eliciting', 'komatsu', 'aggregates', 'organelles', 'stge', 'lc3', 'autophagosome', 'blockage', 'autophagosomal', 'sequestered', 'noncanonical', 'taguchi', 'twofold', 'lowered', 'adds', 'wilms', 'camp', 'hast', 'diptidyl', 'denicola', 'brafv619e', 'mycert12', 'employing', 'concrete', 'fumarate', 'krebs', 'metabolite', 'krebs', 'metabolized', 'malate', 'hydratase', 'eng', 'modifies', 'adam', 'plant', 'phytochemicals', 'sulforaphane', 'curcumin', 'epigallocatechin', 'gallate', 'resveratrol', 'cafestol', 'kahweol', 'cinnamonyl', 'zerumbone', 'garlic', 'oganosulfur', 'lycopene', 'carnosol', 'detoxification', 'carcinogenic', 'cruciferous', 'vegetables', 'broccoli', 'protected', 'conaway', 'kuroiwa', 'consumption', 'spitz', 'seow', 'ambrosone', 'joseph', 'sprouts', 'glucosinolates', 'isothiocyanates', 'glucosinolate', 'glucoraphanin', 'china', 'sprout', 'aflatoxin', 'balogun', 'dietary', 'adduct', 'liviers', 'pyrene', 'garg', 'thimmulappa', 'oltipraz', 'pyrazinyl', 'dithiole', 'thione', 'roebuck', 'rao', 'moon', 'clapper', 'nishikawa', 'iida', 'sharma', 'forestomach', 'bbn', 'urinary', 'inconclusive', 'hydroxylated', '1998a', 'polyaromatic', 'hydrocarbon', 'kelley', 'flatulence', 'irritation', 'paresthesia', 'fingertips', 'dmf', 'alkylate', 'phillips', 'bg', 'ischemia', 'reperfusion', 'ashrafian', 'tecfidera', 'administraion', 'kirsner', 'brusatol', 'brucea', 'javanica', 'seeds', 'ren', 'nanomolar', 'gclm', 'gclc', 'ascorbic', 'lut

eolin', 'ochratoxin', 'trigonelline', 'atra', 'magesh', 'rar', 'wonder', 'max uitenko', 'safe', 'react', 'cddo', 'dithiolethione', 'bardoxolone', 'withdrawn', 'unspecified', 'tayek', 'kalantar', 'zadeh', 'michael', 'nucleophiles', 'cocystal', 'mimetics', 'pollutants', 'repaired', 'detoxifying', 'thiol', 'foods', 'advocated', 'chemically', 'nonstressed', 'holoenzyme', 'modulation', 'nad', 'quinine', 'oxidoreductase', 'peroxiredoxins', 'prdx', 'stresses', 'predispose', 'lk2', 'ebc1', 'ho1', 'u1', 'hn', 'lk', '0247', 'ppt', 'sq', 'lcnec', 'chromatograph', 'dfs', 'segregated', 'hereafter', 'dc', 'isothermal', 'calorimetry', 'itc', '33neh2', '77dxetge82', 'd77v', 'e79q', 't80k', 'observable', 'biphasic', 'w24c', '130f', '231xxqxdxdlg31', 'monophasic', 'unequivocally', 'titration', 'thermograms', 'isotherms', 'ka', 'e82d', 'aqua', 'pymol', '2flu', 'modeller', '9v1', 'detailing', '2h', 'r483', 's508', '2i', 's363', 'r380', 'disturb', 'e79k', 'doublet', 'h1650', 'ebc', 'peroxiredoxin1', 'prd x1', 'mrp3', 'mg132', 'lives', 'l79k', 'accomplish', 'immunofluorescent', 'en counter', 'affords', 'peroxide', 'exporting', 'nonsilencing', 'h2o2', 'nontreated', 'governed', 'cullin3', 'ikb', 'hif', 'contemporaneous', 'd27', 'd29', 'e79', 'e82', 't80', 'q75', 'l30', 'nonconserved', 'w24', 'i28', 'v32', 'r34', 'vulnerable', 'gate', 'indubitably', 'degrons', 'timely', 'hippel', 'lindau', 'pvhl', 'brac1', 'bard1', 'centrosome', 'jeopardizing', 'hyperkeratoses', 'digestive', 'preferential', 'deviated', 'complicated', 'temporary', 'dev astating', 'adaptation', 'evolvement', 'microdissection', 'lm200', 'arcuturus', '44b', 'vp', 'microcal', 'mbl', 'calbiochem', 'nacalai', 'wako', 'immunization', 'diazepam', 'dz', 'hypothalamus', 'let', 'striatal', 'publicly', 'dep tor', 'hyperactivating', 's2215y', 'r2505p', 'l2431p', 'entirety', 'encyclopedia', 'ccle', 'thousands', 'cataloged', 'dozens', 'hypersensitive', 'rapalog s', 'enrich', 'mining', 'c1483', 'e1799', 't1977', 's2215', 'l2427', 'r2505', 'pombe', 'tor2', 'l1460p', 'c1483f', 'e1799k', 'f18881', 't1977r', 'v2006i', 'i2500f', 'd2512h', '4ebp1', 'implying', 's2215f', 'c1483r', 'f1888i', 't1977k', 'i2500m', 'a41p', 'f1888v', 't1977s', 'v2006l', 's2215t', 'r2505q', 'd2512g', 'l1433s', 'a1459p', 'e1519t', 'n2206s', 'l2209v', 'a2210p', 'l2216p', 'r2217w', 'l2220f', 'q2223k', 'a2226s', 'rictor', 'y35n', 'e139k', 'r788c', 's1101l', 'mtors', 'c1483y', 'frb', 'fkbp12', 'torin1', 'trex', 'packaged', 'mln0128', 'gdc0980', 'downregulate', 'retreatment', 'molt16', 'hec59', 'snu349', 'jhuem7', 'tuhr10tkb', 'r2152c', 'sw780', 'nprl2', 'solved', 'prototypical', 'upregulate', 'rapalog', 'commonplace', 'arriving', 'regimen', 'cbio', 'freeze', 'seeing', 'sliding', '463', '2549', 'k463', '463k', '502549', 'equals', 'c29f4', '49d7', '108d2', 's473', 'd9e', 'gt239', 'genetex', 'titer', 'tremegene', 'taconic', 'selleckchem', 'hagst', 'yaffe', 'lab', 'mops', 'ncih446', 'blenis', 'lippard', 'etron', 'synthesizing', 'gblock', 'prk5', 'pcdna5', 'ph8', 'rotated', 'rcf', '2x', 'pulldown', 'volts', '1x', 'caps', 'naoh', 'tbst', 'palpable', 'approached', 'delivered', 'supervision', 'hematologic', 'tailored', 'antiestrogen', 'mgmt', 'activin', 'orphan', '2p23', 'neurotrophic', 'definitively', 'y1282', 'homodimerization', 'sibling', 'nonfamilial', 'tae684', 'nonhematologic', 'npm', 'alcl', 'dlbcl', 'imt', 'echinoderm', '2p21', 'kinesin', 'kif5b', '10p11', 'trk', 'tfg', '3q12', 'conserve', 'oligomerization', 'signet', 'interconnected', 'jak3', 'stat3', 'player', 'scf', 'departments', 'happen', 'miss', 'consist', 'yellowish', '2p', '02341066', 'answers', 'arbitrary', 'rodig', 'tyramide', 'd5f3', 'nucleolus', 'tpm3', 'tpm4', 'pore', 'msn', 'popupdownload', '2341066', 'formulation', 'gtl', 'caki', 'h441', 'tubulogenesis', 'fatigue', 'terminology', 'nausea', 'vomiting', 'diarrhea', 'disturbance', 'alt', 'clearance', 'cyp3a4', 'fasting', 'widening', 'pfs', '1005', 'inversion', '1007', 'pemetrexed', 'docetaxel', 'gsk1838705a', 'igf', '1r', 'sr', 'alcls', 'tae', '684', 'tks', 'stat5', 'cryptic', 'sec31a', 'bioavailability', 'cep', '14083', 'ap26113', 'ariad', 'g1269s', 'cytochrome', 'nms', 'company', '263', '338', 'barrier', 'lu', 'resistances', 'guarante e', 'caution', 'eagerly', 'preselection', 'optimization', 'reductions', 'costs', 'stratify', 'nb1s', 'nbl', 'underscoring', 'cacgtg', 'paolo', 'modulato

r', '4s', 'hydroxylase', 'migrating', 'adrenal', 'medulla', 'superior', 'mesenteric', 'smg', 'sy5y', 'u2os', 'manners', 'previously29', 'tc', 'pgl4', 'box1', 'reports23', 'previously10', 'halts', 'described33', '894', 'consistency', 'ch5424802', 'alkf11741', 'potentiate', '709', 'patients25', 'pivotal', 'ptk', '620', 'phylogenetic', 'sharing', 'ltk', 'imts', 'vcl', 'orchestrator', 'fertility', 'object', 'fruit', 'jelly', 'belly', 'ortholog', 'nucleophosmin', 'consequent', 'cltc', 'atcs', 'rms', 'boosted', 'abl1', 'undergoes', 'acquires', 'manifest', 'hundreds', 'eml', 'asian', 'caucasians', 'cribriform', 'infrequently', 'a20', 'e6', '13th', 'clinics', 'sputum', 'pleural', 'effusion', 'bronchial', 'lavage', 'ones', 'determinable', 'historical', 'intercalated', 'polymer', 'klc1', 'boy', 'k1062m', 't1087i', 'alks', 'l1198f', 'g1201e', 'c1156y', 'deeply', 'thr790', '11152r', '1151tins', 'g1202r', 's1206y', 'g1269a', 'secondarily', 'standpoint', 'magic', 'pill', 'rituximab', 'runx1', 'pml', 'rara', 'entity', 'world', 'promyelocytic', 'apl', 'addicted', 'indispensable', 'tell', 'glu', '659', '1248', '753', 'abolition', 'proteintyrosine', '185', 'intracytoplasmic', 'immunohistological', 'ethylnitrosourea', 'substitutes', 'salivary', 'gland', '4446', 'simian', 'smai', 'drai', 'linkers', 'att', '833', '834', 'psv2neo', 'transfectants', 'seaplaque', 'antiserum', 'immunizing', 'antiphosphotyrosine', 'uci', 'solubilized', 'hydrochloride', 'nonident', 'immunocomplexes', 'adsorbed', 'hydroxyethylpiperazine', 'ethanesulfonic', 'flm', 'mncl2', 'boiling', 'ile', 'peninsula', 'phosphopeptide', 'phosphocellulose', 'thinlayer', 'utilizes', 'jig', 'clonally', 'biol', 'sw', 'thi', 'detergent', 'clarification', 'phosphorylations', 'locate', '1239', '1255', 'comigrated', 'phosphorylated', '1s', 'av', 'aq', 'tsr', 'wss', 'yk', 'e8', 'iti', 'apt', 'iw', 'els', 'phe12', 'weights', '837', 'ir', '185k', 'cii', '838', 'y1248', '1025', 'ffu', 'microgram', 'fiore', 'moloney', 'yarden', 'schlessinger', 'dimers', 'dimeric', 'cerione', '839', '248', 'hazan', '1173', 'haley', 'gill', 'characterizations', 'segatto', '840', '200k', '116k', '92k', '155k', '45k', '165k', 'transforms', 'hetero', 'oligomers', 'delta', 'presumptive', 'hes', 'r481g', 'l507p', 'i562m', 'h570r', 'h650q', 'n659s', '1705p', 'r748g', 'y849s', '32d', 'expressors', 'antagonized', '849', 'mpns', 'neutrophilic', 'polycythemia', 'vera', 'thrombocythemia', 'mastocytosis', 'sm', 'nonreactive', 'eosinophilia', 'cel', 'pdgfrb', 'lymphoid', 'jak2', 'eosinophilopoietic', 'gists', 'race', 'cytogenetically', 'fusions2', 'declaration', 'helsinki', 'hemi', 'nested', 'ttcagccaggatggaaaggctgtca', 'f2', 'cagacggtgagggtcacag', 'tgtcaaagatgctcaggagc', 'bidirectionally', 'antibiotics', 'iu', 'wehi', 'walter', 'eliza', 'hall', 'c3h', 'hej', 'janvier', 'authorities', 'rhine', 'westphalia', 'hendrik', 'heldin', 'ludwig', 'upsala', 'sweden', 'gatccactcccgagacggaaagtaccgtggagggc', 'gccctccacgtacttccgtctcgaggatggatc', 'gctaagaatctccctggagctgagaacc', 'ggttctcagctccaggagattcttagc', 'cgctggagggtcatggaaatcaatcagccc', 'gggctgattgattccatgaccctccagc', 'caatcagcccagatggacgtgaatatatttatgtgg', 'ccacataaaatattcacgtccatctggctgattg', 'ctgaagataatgactcaactggggccaatttgaac', 'gttcaaatgtggccccagttgagtcattatcttcag', 'catttgaacattgtaagcttgctggagcctg', 'caggctcccagcaagttacaatgttcaaatg', 'gagaagccaaagaaagagccggatattttggattgaacc', 'ggttcaatccaaagatatccggctctttttggcttc', 'cagtatgtccccatgctaaggccaaagaggttctaaatattc', 'gaatatttagaaacctttgccttctagcatggggacatactg', 'gcatgattcgaactctgtcgaaaggcag', 'ctgccttcgacacagagttcaatcatgc', 'pmy', 'ig', 'toshio', 'kitamura', 'tokyo', 'egfp', 'cytometric', 'plat', 'sortin', 'logarithmic', 'carboxymethoxyphenyl', 'sulfophenyl', 'microliters', 'phenazine', 'methosulfate', 'pms', 'od490', 'methocult', 'm3231', 'stemcell', '000g', 'pstat5', '495', 'sc835', '4058s', '3164', 'ppdgfra', '2992', 'spleen', 'akc', 'machine', 'retrovirally', 'syngeneic', 'moribund', 'retro', 'orbital', 'kilogram', 'nonparametric', 'p567p', 'jak2wt', 'jak2v617f', 'depicts', '180', 'glycosylated', 'aqueous', '72h', 'aad', 'gates', '0nm', '159nm', '6nm', '8nm', '689nm', 'pdfgra', 'intravenous', 'diseased', 'splenomegaly', 'lymphadenopathy', 'recipients', 'granulocytic', 'blastlike', 'photographs', 'deplete', 'spleens', 'hematopoietic', 'spl', 'nace', 'naphthyl', 'chloro', 'ester'

ase', 'prolongs', 'postinjection', 'supposedly', 'surviving', 'overproduce', 'immunochemotherapy', 'fluctuated', '4560', 'expressor', 'cord', 'ablhigh', 'abllow', 'aml14', 'cd3', '627', 'insert', 'separates', 'nonpolar', '691', '795', 'd842v', 'd846y', 'y849c', 'd842', 'd846', 'v561d', 's566', 'n659k', 'fibroid', 'ifps', '571', '824', '848', 'f8081', 'n870s', 'fifth', 'a509d', 'c235y', 'w349c', 'v536e', 'deletions34', 'suffering', 'tki', 'learn', 'ifps 31', 'infiltrate', 'adenopathy', 'mpn', 'k179m', 'g161v', 'liability', 'broad en', '7554', 'elife', '03751', 'divide', 'manage', 'restrictions', 'encourag e', 'wrong', 'deactivate', 'vivanco', 'external', 'kill', 'pearce', 'encourag ed', 'hirai', 'gsk690693', 'heerding', 'rhodes', 'gdc0068', 'dependently', 's upplement', 'okuzumi', 'gherardi', 'h1993', 'h1648', 'mk2006', 'addiction', 'calleja', 'wu', 'w80a', 'erbb3', 'accompany', 'chandarlapaty', 'antagoniz e', 'franke', 'coexistence', 'phospholipid', 'k181m', 'mcf10a', 'r86', 'r86 a', 'ericson', 'r25c', 'parikh', 'wondered', 'precedence', 'i289m', 'h355y', 'r368c', 'dko', 'tpa', 'arita', 'landgraf', 'equimolar', 'g311d', 'barretin a', '4h', 'document', 'suspect', 'richest', 'selleck', 'houston', 'tx', 'bioc hem', 't246', 'pakt1', 'pakt2', 's474', 'rxrxxs', 'pbad', 'pmet', 'y1349', 't 202', 'y204', 'pfoxo3a', 's253', 'vinculin', '16b12', 'covance', 'princeton', 'alexa', 'phalloidin', '546', 'conjugate', 'pip', 'echelon', 'japanese', 'bio resources', 'jcrb', 'silvia', 'giordano', 'instituto', 'candiolo', 'bert', 'j ohns', 'hopkins', 'h1975', 'hcc827', 'hcc4006', 'sciencell', '2200', 'mccoy', 'lonza', 'allendale', '2201', 'lafayette', '358718', '237948', '325553', '325 558', 'plncx1', 'makt1', 'birnbaum', '15990', 'plncx', 'lightning', 'myr', '1 5989', '27295', 'clai', 'subclone', 'plpcx', 'william', 'sellers', '9004', 'h tert', 'p53dd', 'christopher', '11128', 'gctgcacaaacgaggaaatcatcaagacctg', 'caggtcttgatgtactccctcggtgcagc', 'ccgctactatgccatgatgatcctcaagaaggagg', 'cctccttctgaggatcatcatggcatagtagcgg', 'ccgctactacgcatgatgatcctcaagaaggaa g', 'cttccttctgaggatcatcatggctagtagcgg', 'aagacctggggccatgtacttcctcc', 'g gaggaagttagcatggcccgaggctt', 'cagtggaccactgtcatcgaaggccacccatgt', 'cacatg gaagggtggcttcgatgacagtggccactg', 'tccgctgcctgcaggcgaccactgtcatcg', 'cgatgacag tggcgcctgcaggcagcgg', 'gctactacgcatgatgatcctgcggaa', 'tccttccgcaggatcat catggctagtagc', 'caaactccttggcaaggtaaccttggcaaagtca', 'tgactttgccaaggattacc ttgccaaggagttt', 'aaagtcatcctggatggagaaggccactg', 'cagtgccctctcccacaccagg atgactt', 'tcatggaagagatctgcttccgcgcacg', 'cgtgcgcggaaagcagatcttccatga', 'cttctacaaccaggactatgagcccttcgagc', 'gctcgaagaggcgtcatagtcctggttgtaa g', 'tggacaaagatggccacatgaagatcaactgacttt', 'caaagtcaacttcatgtggccatctttg tcca', 'sonicated', 'amphotropic', 'ampho', 'imgenex', 'particles', 'immortal ize', 'ihem', 'tet', 'pmd2g', 'ppax2', 'hairpins', 'g148', 'unattached', 'vic ell', 'unequal', '9840', 'normalize', 'slurry', 'pmoles', 'upright', 'axio', 'camera', 'orca', 'hamamatsu', 'photonics', 'donkey', 'leica', 'tcs', 'sp5', 'wll', 'multiphoton', 'overactivation', 'phosphodiesterase', 'reactivating', 'deadliest', 'year1', 'migration2', 'pathway3', 'kinases5', 'pathway8', 'cel ls9', 'cells10', 'acti', 'pathway12', 'msh', 'melanosome', 'pde', 'pital', 's aint', 'paris', 'denis', '3department', 'nicolas', 'dumaz', 'inserm', 'januar y', 'april', '1038', '2022', 'andr', 'martine', 'rights', 'reserved', 'deregu lating', 'nontransformed', 'ex', 'stim', 'ulate', 'melana', 'rasraf', 'coimmu noprecipitated', 'scr', '586', 'precipitate', 's17n', 'wm852', 'nhems', 'prei ncubated', 'immunogen', 'situated', 'phosphomimetic', 'interacted', 'pbraf', '587', 'uncoupled', 'ibmx', 'forskolin', 'adenylyl', 'cyclase', 'mc1r', 'nhe m', 'pde6', 'contributor', 'pde4d', 'comigrates', 'comigrate', 'sbcl2', 'wm17 91c', 'zap', '588', 'ascertain', 'reactivate', 'tritiated', 'wanted', 'dyin g', 'annexin', '170', '589', 'neous', 'inhibitors30', 'cells15', 'ksr1', 'de f', 'heterodimerization32', 'regu', 'lation', 'recycled', 'pp2a', 'prolyl', 'persisting', 'pin1', 'bypass', 'r10', '590', 'described38', 'monocytes', 'r eactivates', '10a', '10b', 'asthma', 'rheumatoid', 'ribonucleoproteins', 'snr nps', 'spliceosomes', 'spliceosome', 'u2', 'u4', 'u5', 'catalyzes', 'metazoan s', 'sf3b1', 'srsf2', 'zrsr2', 'sf1', 'sf3a1', 'u2af2', 'prpf40b', 'mds', 'sf

3b155', 'snrnp', 'huntingtin', 'r625', 'k666', 'k700', 'sideroblasts', 'rars', 'analyzes', 'uqcc', 'adam12', 'gas8', 'crnde', 'ankhd1', 'gusbp1', 'f8', 'abcc5', 'gusbp11', 'aggregation', 'k700e', 'k666t', 'g742d', 'h622q', 'd781g', 'r625l', 'r625h', 'k666m', 'cytopenia', 'multilineage', 'rcmd', 'rs', 'misclassified', 'mds44', 'um48', 'splicesome', 'p95h', 'k413e', 'v304m', '161', 'euclidian', '124k', 'undistinguishable', 'latest', 'deregulations', 'intervene', 'orchestrated', 'matera', 'wahl', 'dehm', 'manley', 'yoshida', 'ogawa', 'uvm', 'skcm', 'biankin', 'landau', 'malcovati', 'ss', 'ribonucleoprotein', 'particle', 'bps', 'polypyrimidine', 'ag', 'auxiliary', 'u2af', 'gozani', 'snrna', 'u6', 'lariat', 'horowitz', '622', '781', 'electron', 'golas', 'deboever', 'furney', 'gentien', 'sf3b1mut', 'neomorphic', 'sf3b1wt', 'nmd', 'kervestin', 'jacobson', 'bioinformatic', 'predetermined', 'sf3b1k700e', 'tpm', 'whiskers', 'tukey', 'd8', 'a1jh', 'a19s', '5022', '0078', 'miscellaneous', 'n626d', 'd781e', 'k666e', 'convergence', 'panc05', 'q699h', 'k666n', 'h662q', 'naln', 'sf3b1k666n', 'sf3b1h662q', 'sf3b1k700k', 'nanostring', '699', 'annotations', 'sf3b1k700r', 'panc10', 'quartiles', 'midline', 'mxsf3b1mut', 'mxsf3b1wt', 'mxsf3b1k700r', 'mxsf3b1', 's2d', 'k700r', 'lose', 'proposing', 'ess1', 'ags', 'adenines', 'mxsf3b1k700e', 's4d', 'uxs1', 's4e', 'trimer', 'minigenes', 'bold', 'zdhhc16', 'tracts', 'e9', 'i9', 'e10', 'i9noag', 'i8', 'pyt', '10n', '6n', '2n', 'selex', 'lim', 'coolidge', 'ilagan', 'transesterification', 'bpss', 'k562', 'mercer', '151', 'coincide', 'aaa', 'ggg', 'tmem14c', 'sept6', 'dlst', 'dip2a', '8912', 'erroneous', 'termination', 'rivas', 'isotope', 'silac', 'oda', 'abcb7', 'akap81', 'mzb1', 'phgdh', 'secretion', 'anken', 'y623c', 'iron', 'ringed', 'boultwood', 'papaemmanuil', 'sideroblastic', 'ataxi', 'overload', 'allikmets', 'ptcs', 'chx', 'ptc', 'sf3b1y623c', 'sf3b1e622d', 'importin', 'lukl', 'electric', '30a', '34a', 'crosslinking', 'spliceostatin', 'corrionero', 'compensating', 'pairing', 'slu7', 'chua', '1999a', '1999b', 'zn', 'knuckle', 'spf45', 'hspf45', 'rbm17', 'corsini', 'lallena', 'unanticipated', 'engaging', 'porse', 'kalyna', 'sickness', 'rps14', 'cskn1a1', 'ebert', 'downregulating', 'ribosome', 'shirai', 'almeida', 'kfir', 'khan', 'bapt', 'eif1ax', 'hematological', 'dismal', 't3', 'enucleation', 'humanomni2', 'tetraploidy', '1p', '6p', 'preceded', 'trisomy', 'commensurate', 'inversions', 'acral', 'mucosal', '1629', '2604', '2112', '8000', 'dinucleotides', 'commonest', '2q33', 'reactome', 'genosplice', 'dexseq', 'mats', 'homogenous', 'negates', 'spliceosomal', 'vitamin', 'ubiquinol', 'reductase', 'chaperone', 'institut', 'curie', 'december', 'biobank', 'ended', 'mulv', 'manufacturers', 'cery', 'pontoise', 'gdna', 'sequencher', 'hiseq2000', '100bp', 'mirneasy', 'ambion', '100ng', 'robustness', 'console', 'unadjusted', 'confident', 'archive', 'sra', 'unspliced', 'quantitect', 'courtaboeuf', 'device', 'css', 'pheno', 'ega', '447', 'hcc78', 'slc34a2', '428', 'shrinkage', 'sdc4', 'cd74', 'ur2', 'avian', 'aside', 'u118mg', 'aka', 'gopc', 'cholangiocarcinoma', 'ezr', 'lrig3', 'rat1', 'il3', 'shp', 'jak', 'stat', 'surgically', 'istituto', 'clinico', 'humanitas', 'milan', '623n3', '654', '117o13', '971', '59k17', '626', 'ctd', '2314k7', '438', 'beak', 'blinded', 'colorado', 'ttf1', 'marque', '8g7g3', 'p63', 'biocare', 'bc4a4', 'ck5', 'cytomination', 'd5', 'b4', 'micron', 'conditioner', 'blocker', 'iview', 'recoverall', 'austin', 'manufacture', 'hogenized', 'nucleoproteins', 'organic', 'nuclease', 'cu', 'xalkori', 'nct00585195', 'superscript', 'rev1', 'transcribe', 'dntps', 'rnasin', 'thermocycler', 'clean', 'circularization', 'grand', 'pshai', 'ipswich', 'invpcr', 'thermocycle', '3730', 'e34r', 'e4f', 'e5f', 'e34f', 'e2f', 'touchdown', 'e1f', 'e43f', 'pcdh', 'mcs1', 'mountain', 'r8', 'transit', 'mirus', 'john', 'minna', 'theodorescu', '04217903', 'york', 'nvp', 'py2274', '3078', '6347', 'py542', '3751', '610621', 'ps437', '4058', '2920', 'py705', '9145', '9139', 'pt202', '9107', '8035', 'mab274', '20ul', '490nm', 'gallios', 'cytometer', 'brea', 'modfit', 'verity', 'house', 'topsham', 'ary001', 'minneapolis', 'mn', 'slc34a', 'anschutz', 'campus', 'smoker', '250mg', 'po', 'bid', 'recist', 'suv', '48hr', 'heparan', 'proteoglycan', 'takeuchi', 'bergethon', 'histologies', 'trying',

'202', '1476', 'christensen', 'scl34a2', 'priority', '073', 'exploring', '6q2', 'massachusetts', 'mgh', '574', 'vanderbilt', '443', 'irvine', 'fudan', 's hanghai', 'censored', 'joint', 'tnm', '835i21', '1036c2', 'deparaffinized', 'codenatured', 'hybrite', 'slipped', 'bx61', 'genetix', 'jose', 'fusing', 'a gencourt', 'formapure', 'tcggatttctacttttcgtg', 'ctcctgttcaa', 'gaaatgcgttatttgg', 'tgaaacttgttctgttatccaa', 'mastercycler', 'hamburg', 'pc9', 'mgh006', 'h3122', 'hcc', 'chemietek', 'fitchburg', 'centro', '960', 'luminometer', 'berthold', 'oak', 'ridge', 'tn', 'mediatech', 'superfect', 'ethyleneglycoltetracetic', '3flag', 'geneart', 'regensburg', 'dichotomous', '0b', 'identifier', 'institution', 'enrolling', 'sponsored', 'withros1', 'rearranged', 'overrepresentation', 'asians', '663', '607', '472', 'bronchioloalveolar', 'improvement', 'restaging', 'continues', 'coronal', 'reconstructions', 'nonsmoking', 'predisposing', 'nonsmokers', 'subclassification', 'expertise', 'nsclcs21', 'commercially', 'learned', 'began', 'registration', 'ddr1', 'circuitry', 'humdisease', 'today', 'pi3kca', 'greenman', 'sea', 'corrupted', 'kinzler', 'blume', 'hunter', 'weinstein', 'rush', 'wolf', 'gu', 'immunoaffinity', 'sampled', 'semiquantitative', 'tyr1234', 'parenthesis', 'phosphoproteome', 'olsen', 'phosphopeptides', '4551', '2700', 'phosphosite', 'website', 'papers', 'html', 'approximate', 'speaking', 'simplify', 'epha2', 'eph', 'fak', 'ullrich', 'pies', 'amann', 'h3255', 'paez', 'tracy', 'calu', 'lutterbach', 'minami', 'predominating', 'nonkinase', 'predominately', 'glycolytic', 'complicating', 'grouping', 'unusually', 'summed', 'phosporylation', '847', '233', '662', '1295', 'ephb4', '912', 'axl', '761', '772', 'ephb1', '1439', 'epha4', 'ephb3', 'epha1', 'epha5', 'tyro3', 'minus', 'omitted', 'readable', 'lynch', 'pao', 'dele746', 'a750', '1858r', '1747', '1833v', '1838v', 'e746', 't751', 'h1703', 'corner', 'hernandez', 'solute', 'ending', 'mif', 'leng', 'curiously', 's3e', 's3f', 'mellinghoff', 'strict', 'ship2', 'prominently', 'charest', 's3h', 's3g', 'giving', '42mapk', 'tyr1068', 'scansite', 'druker', 'mahon', 'h1373', 'correlating', 'ong', '926', '962', 'recycling', 'endocytosis', 'metalloproteinase', 'ceramide', 'occasions', 's4g', 's4h', 'expansive', 'circuits', 'classifying', 'fyn', 'hck', 'frk', 'brk', 'unexplored', 'deserves', 'egfris', 'mahadevan', 'individualized', 'mitelman', 'oncokinase', 'adenocarcinomas2', 'nsclc4', 'histocompatibility', 'respectively4', 'method2', 'pcr2', 'k17', '304', 'consecutively', 'respectively1', 'consumed', 'q12', 'q15', 'q22q25', 'q14', 'homodimerization', 't8', 'n35', 'r32', '596', 's13del2046', '544', 'c6', '116', 'ladder', 'captures', 'k15', 'r12', 'k16', 'k22', 'k23', 'k24', 'r11', '22q11', '432', 'e20', '913', 'e14', 'ins11del49a20', '454', 'ins69a20', '501', 'del14a20', '814', '999', 'ccdc6', 'r121', 'fomed', 'cancer12', 'service', '482', 'vandetanib', 'preserves', 'thirteen', 'thirty', '00088', 'diagnoses', 'ntrk3', 'pattern7', 'fusions10', 'summarize', 'previously1', 'tropomyosin', 'syndecan', 'ret51', 'ret9', 'ret14', 'r12s', 'proliferated', 'gene15', 'chernobyl', 'accident16', 'atomic', 'bomb', 'survivors17', 'patients18', 'desirable', 'fukushima', 'accident', 'mutations19', 'treatment20', 'roach', 'benvenuto', 'chong', 'kopera', '900', '770', 'dunlop', 'hedges', 'nonterminating', 'methodstsc2', 'variantstsc2', 'x75621', 'gi', '450351', 'r611q', 'antiseraexpression', 'midi', 'venlo', 'leiden', 'af013168', '2331280', '2b4', '1a5', 'lincoln', '680', 'zymed', '3g9d9', 'variantsthe', 'verviers', 'dioxide', 'polyethyleneimine', 'polysciences', 'coeveots', 'x100', 'biochemicals', 'woerden', 'criteriontm', 'powder', 'campina', 'melkunie', 'eindhoven', '1895', 'pbst', 'poole', 'odysseytm', 'default', 'offset', 'pixel', 'border', 'alamut', 'biosoftware', 'rouen', 'intolerant', 'tolerant', 'server', 'evs', 'washington', 'resultsfunctional', 'assessmentin', 'pm1', 'v241del', 'e1552del', 'l1750afs25', 'e75g', 'p6701', 'v11441', 'h1773p', 'viewer download', 'slidefunctional', '1525', '1770', 'slideimmunoblot', 'coexpressing', 'n1643s', 'l1493v', 's13791', 'l1584r', '889', '1552', '1653', '241', 'a889p', 't1203k', 'g1567d', 'v1673d', 'i427m', 'r462c', 'r505q', 'y598c', 'q1554h', 'slideoverview', 'brakets', 'disagreement', '901', '1807', 'shading', '4

662g', '1514g', 'discussionwe', 'think', 'r1200w', 'y1571n', 't1623i', 'p1675l', 'ahead', 'epitopes', '844', 'r462h', 'r462p', 'y598h', '1281c', '1477c', 'langkau', 'fulfill', '758', '4136c', '496', '2753c', 's918c', 'disagreemen t', 'variantsgroup', 'variantsvariants', 'v241del4', '1410r', '1493p', 'h597y', 'v705e', '1792r', '1826p', '1830r', 'e1552del4', 'h1620r', 's1653p', 's1653f2', '1146r', 'v299g', '1448p', 'r462c1', 'y598c1', 'v705m', '1844r3', '11750afs254', 'i427m2', '1493v2', 'r505q2', 'q1503p', 'q1554h1', '11584r1', 'w1610g', 'v1673f', 'p170915', 'g1596v', 'a328p6', 'p911', 'r245h', 's433c', 'v1144m', 'p1145l', 's1207i', 's1454g', 'v1500m', 'r1706c', 'v1711m7', 'e114k', 's235n', 'e254k', 'a447v', 'e498k', 'h522y', 'm649t', 'g654c', 'g661r', 'g661v', 'a772e', 'n958s', 'y1033h', 'v1120e', 'p1315s', 'p1358l', 's137913', 'p1381l', 'n1522s', 'e1679k', 'd1734e', 'p1771l', 's1774tour', 'activty', 'huntington', 'tor1', 'fatc', 'companion', 'tor', '2430', '2450', 'caffeine', 'fission', 'nitro', 'nitrosoguanidine', 'mutagen', 'i2017t', 'addictive', 'formati ons', 'a8s', 'p2476l', 'firmly', 'sapiens', 'canis', 'lupus', 'familiaris', 'taurus', 'mus', 'musculus', 'rattus', 'norvegicus', 'gallus', 'danio', 'rerio', 'anopheles', 'gambiae', '938', '1456', '1479', '2273', '2284', '2294', '2288', 'choices', 'pcmv6', 'rockville', '2821', 'k2100', 'tbs', 'ultrasonic ation', 'microcentrifuge', 'rotator', 'lisa', 'emd', 'immunoprecipitate', 'pt389', 'tmb', 'lyzed', 'radioimmune', 'precipitation', 'skim', 'sliced', '2972', '2855', '7985', '4858', '2215', '2217', 'films', 'costar', 'instructed', 'multilayered', 'eclipse', 'ti', 'tr', '1004', 'maxtract', 'cultivation', '214010', 'axiovert', '200m', 'inserts', 'polyethylene', 'pores', 'falcon', 'mu ltiwell', 'noninvasive', 'hsd', 'foxn1nu', 'harlan', 'frederick', 'aliquot', 'needles', 'definitely', 'ensembl', 'reconfirmed', 'rps6', 'reflection', 'adherent', 'electropherogram', 'adherence', 'thickening', 'twisting', 'shapin g', 'stacked', 'interested', 'avid', 'ubiquitously', 'preclinically', 'assuri ng', 'pikk', 'pi3kc', 'yeasts', 'rt112', 't24', 'valid', 'intratumor', 'prognostication', 'aided', 'available1', 'compilation', '947', '479', 'ahr', 'slfn11', 'topoisomerase', 'schemata', 'speed', 'regimens2', 'mainstay', 'facile', 'lineage3', 'types6', 'affirming', '392', 'genotyping13', 'tumorscape14', 'expo', 'mile', '10215', 'previously3', '1022', 'proxies', 'logistical', 'amax', 'ec50', 'hill', 'hdac', 'lbh589', 'panobinostat', 'harvard', '02142', '02115', '02139', '92121', 'emeryville', '94608', 'howard', 'hughes', 'chevy', 'maryland', '20815', 'bionetworks', '1100', 'fairview', 'ave', 'seattle', '98109', '10065', '603', 'macmillan', 'publishers', 'techniques6', 'modelling', 'b ayes', 'elastic', 'analysis16', 'bootstrapped', 'andnraswere', '0325901', 'ptpn5', 'overexpressionfor', 'erlotinib8', 'lapatinib17', 'nutlin', 'ext2', 'glycosyltransferase', 'heparin', 'biosynthesis', 'sensitivity21', 'hsp90', 'aagg', 'aagh2', 'inhibitors23', 'suboptimally', 'sox10', 'vorinostat', 'romidepsin', 'tumorscape', 'squam', 'nsc', 'endometrium', 'ovary', 'oesophagus', 'liv erkidney', 'hodgkin', 'mesothelioma', 'burkitt', 'ewing', 'osteosarcoma', 'bi le', 'chondrosarcoma', 'medulloblastoma', 'freq', 'corr', 'excl', 'diagonal', 'converse', 'v56', 'oesophageal', 'letter', '604', 'aew541', 'inhibitors24', 'origfir', 'antagonist25', 'coexistent', '3i', 'cyp1a1', 'ipc', '298', '3k', 'chp', 'govern', 'rma', '3456789', 'nnms', 'cosmicms', 's100a4', 'gapdhs', 'cyt27b1', 'ncrna00173', 'grin2a', 'sulf2', 'c5orf39', 'acss3', 'bootstrapping', 'receiver', 'crossvalidated', '387', 'quartile', 'whisker', '605', 'irinotecan', 'camptothecin', 'top1', 'topotecan', '14b', '14d', 'hut', 'tt2609', 'ko52', 'hec', '1080', 'h1299', 'kmm', 'ichikawa', 'ons', 'mino', 'fgh', 'sh luc', 'shahr', '0123', 'ahrr', 'phrf1', 'c20orf173', '1234', 'oncomap', 'denote', 'curvesfornras', 'ofnras', 'mhh', 'hep', 'cervix', 'leiomyosarcoma', '174', '576', '542', 'peritoneal', 'fallopian', '245', 'uterus', '448', 'vulva', '427', 'testis', 'sept9', 'ulk4', 'tmem90b', 'spock3', 'mss', 'vertical', 'ordered', 'interquartile', 'option', 'irinotecanbased', 'oftopotecan', 'regimen s26', 'correlative', 'assembling', 'annotationofpreclinicalhumancancermodel s', 'ormechanistic', 'efforts27', 'lapatinib', 'nh2', 'energetically', 't798

i', 't790m', 'ckit', 'l755s', 't733i', 'exel', '7647', 'circumvent', 'reliev e', 'dasatinib', 'nilotinib', 'sunitinib', '690514', 'irreversible', 'hki', '272', 'yvma', 'g776', 'tykerb', 'presents', 'medicinal', 'exelixis', 'gp2', 'nonessential', 'azam', 'v659e', 'cooh', 'prv2', 'endofree', 'mega', 'pvsv', 'tablettop', 'centrifuge', 'allegra', '6r', 'precoated', 'takara', 'tccggggac gaattctgcaca', 'tcacattccccatcaggatcca', 'tctttggatcctcatcaag', 'agggtctg gacagaagaagc', 'ex24', 'attgactctgaatgtcgcc', 'caactggcacgtccagaccc', 'ggcgcctactcgctgaccctgc', 'sr1', 'gccaaaagcgccagatccaagca', 'sf2', 'ctcagtgacctgttt gg', 'sr2', 'aggtcaggttcacaccg', 'cagcagaagatccggaagta', 'atactcctcagcatccac ca', 'gcctcttagaccatgtcc', 'atccctcaggttagctcatcc', 'cgggagttggtgtctgaatt', 'ctctgggttctgcctgtag', 'aagcccccacacatgacc', 'ttcacatattcaggctgg', 'phred', 'phrap', 'consed', 'wa', 'operon', 'luminosity', 'multilabel', 'photograph y', 'afterwards', 'bicinchoninic', 'novex', 'perbb2', 'typhoon', '1m17', '1xk k', '2gs6', '2gs7', 'proprietary', 'glide', 'schrodinger', 'graphics', 'delano', 'passaging', 'unmutagenized', 'seventeen', 'l785f', 'singlet', 's783p', 'l726f', 'p7801', 'l726i', '555', '554', 'agreed', '610', '918', '861', '267', 'parallels', 'l777f', 'bulkier', 'aliphatic', 'rigid', 't798', 'l248r', 'q252h', 'l726', 'v2991', 's783', 'l755', 'packing', 'c630y', 'e717k', 'e719k', 'e719g', 'y803n', 'e812k', 'd821n', 's1002r', 'resynthesize', 'uninfected', 'piled', 'participates', 'vulnerability', '755', 'thermodynamic', 'incompatible', 'boundary', 'capecitabine', 'potentiates', 'nonmutation', 'cmet', 'manifestation', 'contrasting', '2of', '21q22', '3396', 'copa', '5v', '7p21', 'rnawas', '7300', 'calibrated', 'coralville', 'exon2', 'exon3', 'exon4', 'exon11', 'rlm', 'generacer', 'chinnaiyan', 'fellows', 'scientist', 'requests', 'reprints', '1301', 'street', '4237', '48109', '0602', '734', '647', 'umich', '5472', 'cancerres', 'exon7', '80jc', '37jc', 'posthybridization', 'digoxigenin', 'fluor', '594', '4v', 'ccd', 'bacs', 'underexpression', '3397', 'maxi', 'mixes', 'interrogating', 'overexpresses', 'pca3', 'lcm', 'diagrams', 'horizontal', 'coloring', 'tracing', 'tissues8', 'elf3', 'walking', 'analytic', 'subversion', 'ascribed', 'ews', 'workshop', 'rer', 'unify', 'endorsed', 'insta', 'bility', 'quency', 'assist', 'hnpcc', 'defrayed', 'payment', 'advertisement', '1734', '9500', 'ucsd', 'epn', 'blvd', '20892', '1the', 'nonpolyposis', 'mmr', 'replicative', 'seer', 'epidemiology', 'repetitive', 'nonhereditary', 'tu', 'mors', 'intersection', 'cer', 'descriptors', 'thibodeau', 'abbreviated', 'peltomaki', 'aaltonen', 'quences', 'obser', 'tors', 'europe', 'asia', 'intramural', 'extramural', 'academia', 'session', 'recommending', 'uniformity', 'indi', 'extracolonic', 'chaired', 'discussing', 'methodological', 'applicability', 'acronyms', 'disadvantages', 'mmp', 'mutator', 'evenly', 'controversy', 'certainty', 'hmsh6', 'superfluous', 'mono', 'tetranucleotide', 'overestimating', 'mononucleotide', 'bat25', 'bat26', 'bat40', 'racial', 'fishel', 'd5s346', 'd2s123', 'hmsh2', 'bat34c4', 'rii', 'actc', '635', '636', 'hamilton', 'gilbert', 'hamelin', 'kirsch', 'markowitz', 'bapat', 'proves', 'mation', 'identi', 'fication', 'recommends', 'merging', 'icg', 'hans', 'vasen', 'tected', 'enriching', 'pathol', 'encourages', 'ment', 'failures', 'reanalyze', 'reamplification', 'discern', 'felt', 'noninformative', 'concur', 'spe', 'cific', 'kindred', 'diagnose', 'metachronous', 'gas', 'tric', 'hepatobiliary', 'pelvis', 'ureter', 'irregular', 'spaces', 'med', 'investigative', 'jass', 'queensland', 'barbara', '303', 'reacting', 'lym', 'expansile', 'deserve', 'quent', 'serrated', 'managing', 'registry', 'gynecological', 'muta', 'crude', 'quote', 'jarvinen', 'fap', 'colectomy', 'bax', 'donald', '229', '1973', 'nonadenocarcinoma', 'ascending', 'exceed', 'commented', 'karolinska', 'sented', 'stephen', 'proficient', 'proapoptotic', 'russell', 'jacoby', 'wisconsin', 'nonsteroidal', 'b6', 'ranks', 'num', 'ber', 'relatedness', 'speculation', 'vated', 'declining', 'thioguanine', 'af', 'kindreds', 'sperm', 'relaxation', 'tis', 'frequen', 'cies', 'pres', 'stands', 'aaag', 'pheno', 'deficiencies', 'radicals', 'diet', 'proficiency', 'permitting', 'sidransky', 'jen', 'hmsh3', 'rh', '667', 'malt', 'truncate', 'forth', 'tide', 'explorations', 'anonymous', 'dau

ghter', 'urine', 'fluids', 'cataloguing', 'ation', 'ical', 'collect', 'disseminate', 'interpreting', 'decipher', 'pursue', 'msh3', 'ye', 'iir', 'initiative s', 'opinion', 'specify', 'pendent', 'colorectum', 'prog', 'radiosensitivity', 'anecdotal', 'turcot', 'levamisole', 'predispositions', 'ashkenazi', 'jews', 'education', 'mod', 'dominantly', 'peer', 'reflex', 'arch', '1269', '1277', '0035', 'manifests', '1913', 'henry', 'muir', 'torre', 'keratoacanthoma s', 'pertains', 'utah', 'drs', 'arup', 'financial', 'florida', 'characteristically', 'debated', 'quarter', 'mutl', 'prime', 'epcam', 'magnifications', '3100', '1270', 'pms2', 'mlh3', 'consanguinity', 'exonucleases', 'exonuclease', 'mutsa', 'nascent', 'mutla', 'web', 'accessed', 'precancerous', 'ensues', '1799t', 'sessile', 'tgfb2', 'nodal', 'algorithmic', 'workup', '1272', 'versions', 'expense', '1273', 'bat', 'contracted', 'sized', 'slippage', 'perfectly', 'imperfectly', 'methylated', 'nonmethylated', 'bisulfite', 'adequacy', '1274', 'antigenicity', 'd17s250', 'corporation', 'indeterminate', 'trinucleotide', 'eliminating', 'islands', 'try', 'dissection', 'dissecting', 'lasercapture', 'italian', 'haplotype', '1180', '1386', 'haplotypes', 'cancerogenesis', 'turin', 'father', 'policies', 'heating', 'g168', 'avant', 'transgenomic', 'omaha', 'amplifying', 'paxgene', 'repeatmasker', 'qiaquick', 'stretches', '000251', 'p5', 'hapmap', 'feb', 'gt', '504', '597', '425', '575', 'ancestry', 'pcrs', '3516', 'nonrepetitive', 'inverted', 'interposed', '1276', 'lambada', 'ancestor', 'reconstructed', 'segregates', 'alu', 'desai', 'al16', 'al17', 'eighth', 'southwestern', '7000', 'inhabitants', 'geographical', 'dating', '673', 'drift', 'traced', 'stella', 'beside', 'cumulative', 'kovacs', 'kp', 'generational', 'geographic', 'unmethylated', 'basolateral', 'intercellular', 'lef', 'abundantly', 'renewal', 'unraveled', 'yes', 'reprinted', 'lante t', 'copyright', 'permission', 'elsevier', 'duodenal', 'merits', 'enteropathy', 'absorption', 'parenteral', 'nutrition', 'epilogue', 'revelation', 'individualize', 'branches', 'northeastern', 'anticipation', 'offsprings', 'interindividual', 'norm', '120436', '609309', 'hnppccs', 'sscp', 'ptt', 'satisfied', 'aci', 'ta', 'haplotyping', 'p55', 'p48', 'p26', 'p61', '2p15', 'histories', 'interview', 'iqr', 'ltr', 'fb', 'fg', 'arrowhead', 'deriving', 'families', 'satisfying', 'crcs', 'nodule', 'oblige', 'sign', 'symbol', 'en', 'ass', 'pedigree', 'p22', '1906g', 'foulkes', 'genealogical', 'headed', 'italics', '270', '710', '455', 'predominated', 'transmitting', 'maternal', 'cb', '1941', '1961', 'sx', 'subfamily', 'fraumeni', 'heralded', '1936', 'youngest', 'emphasized', 'appreciably', 'precluding', 'intrauterine', 'gametogenesis', 'enzymatically', 'gem', 'ablates', 'elicits', 'brca1fh', 'i26a', 'cys', 'bred', 'pups', 'testes', 'tubules', 'elongated', 'spermatids', 'spermatozoa', 'spermatogenesis', 'sterility', 'littermates', 'brca1tr', 'pdx1', 'p53r172h', 'kra sls1', 'brca1flex2', 'p53flex7', 'flex7', 'succumbed', 'flex2', '2595', 'kp c', 'wapcre', 'p53lsl', 'r270h', '7502', 'sawtooth', '5197', 'abraxas', 'ccdc 98', 'bach1', 'fancj', 'ctip', 's1655', 'donates', 's1598f', 'brca1s1598f', 'intercrossed', '309', '4543', '2632', 'bach1wt', 'ubiquitinating', 'integral', 'inert', 'predicated', 'ablated', 'ubch5a', 'examines', 'zn2', 'prohibitive', 'permissions', 'journals', 'oxfordjournals', 'kings', 'hospitals', '8th', 'floor', 'tower', 'se1', '9rt', 'tel', '2071883699', '82579', '2071882585', 'jo', 'wish', '1093', 'hmg', 'ddi476', 'postulate', 'leaves', 'ambiguity', 'plethora', 'nucleation', 'decatenation', 'tumourigenesis', 'fewest', 'highaffinity', 'ordination', 'm18', 'wished', 'logged', 'ligating', 'm18k', 'competitor', 'triazole', '3at', 'c24r', 'c44f', 'h41r', 'c39r', '182d', 'orthologs', 'pathogenically', 'nonpolymorphic', 'shaded', '601', 'polyubiquitin', 'nineteen', 'i89t', 'q12y', '182p', 'c39s', 'c44r', 'c47r', 'c61l', 'parity', 'substrateindependent', 'ub', 'ease', '602', 'ubch7', '1337', '869', 'flatter', 'tighter', 'ordinate', 'brace', 'join', 'v11a', 'i15l', 'i21v', 'i42v', '187v', 'd67e', 'r71g', 'k38n', 's4p', 'r7c', 'i90t', 'd96n', 'g98r', 'i124v', 'ubch5b', '1jm7', '1fbv', '1w4u', 'beneath', 'brzovic', 'resonances', 'ubch5c', 'hands', 'atf', 'p300', 'tumourigenic', 'counselling', 'bundle',

'h41l', 'c47g', 'i31t', 't37r', 'k45e', 'glu75', '152f', 's59r', 'i68k', 's72r', 't69', 't77m', 's80', '163f', 'e29v', 'k4', 'k50l', 'd59', 'k56n', 's91', 'q60l', 'k65m', 'q92', 'i26n', 'i26', '128p', 'g57r', 'backbones', 'f79s', 'f79', 'i15t', 'distortion', 'fragility', 'extrapolated', 'e2s', 'v617f', 'jh2', 'f595', 'k539l', 'erythropoietin', 'epor', 'spare', 'jh', 'ferm', 'radixin', 'jh3', 'jh1', 'pseudokinase', 'jak1', '617', 'pv', 'pmf', 'v617', 'places', 'resided', 'facing', 'pointing', 'preservation', 'hydropathy', 'f617', 'transcriptional', 'elicited', 'contrarily', 'epo', 'indistinguishably', '592', 'bosc', 'pmx', 'pgrr5', 'prltk', 'reporters', 'immunobloted', 'pjak2', 'nonstimulated', 'megakaryoblastic', 'amkl', 'regained', 'rotation', 'hinderin g', 'v658f', 'u4c', 'appended', 'prefer', 'alleviate', 'inactivity', 'acknowl edge', 'decisively', 'productive', 'v617l', 'signaled', 'scissors', 'competitors', 'unwanted', 'thrombocytopenia', 'transductions', '9r', 'glomax', 'passi ve', 'spinning', 'iblot', 'tyr694', 'tyr202', 'romano', 'kroemer', '2b7a', 'ucsf', 'chimera', 'biocomputing', 'p41', 'thrombocytosis', 'imf', '3p21', 'gata1', 'tmd', 'yolk', 'sac', 'megakaryocyte', 'monozygotic', 'twins', 'multigen ic', 'precludes', 'hel', 'thrombopoietin', 'tpo', 'tips', 'odense', 'tfa', 'woburn', 'auburn', 'ultimate', 'thermofinnigan', 'mscv', 'mig', 'al29', 'wb', 'tritonx', 'phenylarsine', 'oxide', 'arlington', 'heights', 'tyr1007', '1008', 'y694', 'c17', 'c20', '4g10', 'cd41', 'rnasea', 'cellquest', 'willebran d', 'numeric', 'aperture', 'kalamazoo', 'adobe', 'photoshop', 'megakaryocytes', 'puregene', 'gentra', '3434', 'vancouver', 'bc', 'canada', 'enumerated', 'femurs', 'tibiae', 'spinfected', '1800g', 'spinfection', 'lethally', 'drew', 'dallas', 'immunoassay', 'granted', 'levine', 'al32', '875', 'crystallographically', 'megakaryocytic', 'cmk', '346', 'y1022', 'y1007', 'y1008', 'quadrant', 'tyk2', 'conformationally', 'exposes', '842', '3r', 'mpl', '3j', 'coil s', 'butyl', 'fluoro', 'dihydro', 'filling', 'gmcsf', 'maturing', 'pulp', 'singly', 'clumping', 'neutrophils', 'immunochemistry', 'extramedullary', 'peyer', 'reticulin', 'hematocrit', 'infiltrated', '4m', 'c57', 'bl6', 'leukocytosis', 'pg', '6100', 'immunophenotypic', 'mac1', 'gr1', 'ter119', 'megakaryopesis', 'gated', 'mk', 'burst', 'bfu', 'cfus', 'gemms', 'gms', 'replating', 'disparity', 'platelets', 'proplatelet', 'polyploidization', 'gemm', 'meps', 'organomegaly', 'nonautonomous', 'erythrocytosis', 'hypolobated', 'disparities', 'titers', 'mpd', 'socs', 'mpds', 'inducibly', 'baxter', 'zhao', '2005a', 'scott', 'stats', 'endo', 'naka', 'rui', 'socs1', 'socs3', 'rises', 'homodimeric', 'stat1', 'stripping', 'reprobining', '173k', 'adjust', 'set2', 'walz', 'stops', 'reprobed', '143k', 'migrates', 'yfp', 'colocalize', 'relocalized', 'kir', 'reconstitution', 'smart', '96k', 'inhibitible', '74k', 'corresponded', '2005b', 'bock', 'relocalize', 'csfr', 'crucially', 'hemoglobin', 'fibrotic', 'tefferi', 'bacher', 'reilly', 'yoshikawa', 'prolactin', 'interferon', 'ifn', 'cntf', 'cardiotrophin', 'oncostatin', 'osm', 'jh4', 'bmt', 'ihle', 'jude', 'y705', 'gab2', 'pim', 'stocks', 'duplicates', 'm3234', 'benzidine', 'immunecomplex', 'akaadgy694vkpqikqv', 'potentials', 'tyrosyl', 'phosphotyrosyl', 'hrs', 'rbc', 'neutrophilia', 'distorts', 'helped', 'rationalized', 'nonconservative', 'groove', 'rules', 'counselors', 'homologs', 'benefited', 'heterologous', 'msa', 'orthologous', 'u14680', 'troglodytes', 'af207822', 'u68174', 'af036760', 'af355273', 'u50709', 'xenopus', 'laevis', 'af416868', 'clustalw', 'gapped', 'psi', '1649', '1859', 'salign', 'command', '1jnx', '1l0b', '1kzy', 'limo', 'xrcc1', '1cdz', 'satisfaction', 'restraints', '1694', '1817', '1819', 'modbase', 'buriedness', 'dssp', '1654', '1657', '1659', '1662', '1663', '1666', '1676', '1678', '1698', '1702', '1773', '1774', '1811', '1813', '1834', '1836', 'isotropic', 'neighborhood', 'neighborhoods', 'hydrophobicity', 'lifwcmvy', 'patgs', 'hqrkned', 'blosum62', 'entropy', 'breaker', 'dino', '1396', 'lexa', 'm1775r', 'y1853x', 'm1689r', 'egy48', 'cotransformed', 'prb1840', 'operator', 'pg5luc', 'phrg', 'tentatively', '1652', '1855', 'generic', 'contoured', 'p1806a', 'hidden', 'dissimilar', 'g1743r', 'l1657p', 'k1702e', 'delimits', 'e1660g', 's1715n', 'f1761i', 'm1775v', 'a1843t', '30a3',

'90a3', 'hierarchically', 'annotating', 'rationalizes', 'neg', 'tolerable', 'unverified', 'presymptomatic', '1560', 'nonglobular', 'canine', 'pentapeptide', 'i1855', 'l1854', 'y1853', 'positional', '1863aa', 'zinc', 'rad51', 'sub nuclear', 'numerical', 'copurified', 'cip1', 'dearth', 'hurdle', 'intention', 'hf7c', 'mat', 'ura', 'lys2', 'ade2', 'gal80', '538', '17mers', 'cyc1', 'sfy5 26', 'cann', 'uas', 'operators', 'plex9', '24endt', 'gcggatcctcagtagtggctgtgg gggat', 'qualitatively', 'pcbrca1', '385', 'erdos', 'h1860x', 's9503101', 'cg gaattcgagggaacccctt', 'acctg', 's970074', 'gcggatcctcagggatctgggg', 'p1856 x', 's970073', 'gcggatcctcatatcaggtagggtgtcc', 'i1855x', '1855stop', 'gcggatcc tcacaggttaggtgtcc', 'l1854x', '1854stop', 'gcggatcctcagtaggtgtcccagc', 'pgbt9', 'pas2', 'pfu', 'dpni', 't1561i', 't1561if', 'cttggattcggaggaaatcccttaccccgagtc tgg', 't1561ir', 'ccagactcgaggtaagggattccctcgaattccag', 'l1564p', 'l1564pf', 'gagggtaccccttaccggaatctggatcag', 'l1564pr', 'ctgattccagatt', 'ccgggttaagg ggtaaccctc', 'd1733g', 'd1733gf', 'aaaaaatgctcaatgagcatg', 'gttttgaagtccgcgg a', 'd1733gr', 'ctccgcggacttcaaaaccatgctc', 'attcagcattttc', 'g1738e', 'g173 8ef', 'gagcatgatttgaagtcagagaaga', 'tgtggtaacccaag', 'g1738er', 'cttccgttaa ccacatcttctgacttcaa', 'aatcatgctc', 'p1806af', 'ggtaccgggtgtccacgcaattgtg', 'gttgtcagc', 'p1806ar', 'gctgcacaaccacaattgcgtggacaccggtacc', 'psh18', 'erica', 'golemis', 'stringency', 'streaked', 'uracil', 'thawing', 'liter', 'na2h po4', '7h2o', 'nah2po4', 'mgso4', 'bromo', 'indolyl', 'galactopyranoside', 'd imethylformamide', 'contructs', 'dextrose', 'inoculate', 'od600', 'shaker', 'recovery', 'isoamyl', 'alcohol', 'naoac', 'strathern', 'higgins', 'raffinose', 'bu', 'salts', '70g', '30g', 'p385', 'cggaattcggaggaaaccccttacctg', 's9503 098', 'gcggatccgtgtggctgtggggat', 'ncoi', 'miniprep', 'terminators', 'pht38 5', 'conjugative', 'transposon', 'tn4430', 'transconjugants', 'cointegrates', 'cointegrate', 'regenerating', 'kpni', 'transposition', 'digesting', 'religating', 'matings', 'selections', 'transconjugant', 'replica', 'tn4330', 'recloned', 'religation', 'od600was', 'cracking', '8m', 'bromophenol', 'electroblotted', 'wet', 'plexa', 'radish', 'p1749r', 'r1751q', 'y1769x', 'm1689k', 'k1727 r', 'l1786p', 's1722p', 'n1774y', 'q1811l', '1756', '1736', 'terminally', 'rough', '1723rgtpi', 'tolerates', '1824ggtpi', '1822gvplh', '1793gvplk', 'mtd', '1760', 'i1760', 'dog', '1524', '1661', 'yli', 'ylv', 'moment', 'contradict', 'm1652k', 'asp4', 'm1652', '1703', '1705', '1822', 'f6', 'w74', 'f1761', 'l2 5', 'l1780', 'destabilising', 'v1833e', 'v1833', 'destabilise', '1676rgtpl', '1717wgtpf', '1780gvpql', 'constraint', 'borders', 'misinterpretation', 'lex', 'recover', 'wealth', 'await', 'traps', 'nonfunctional', 'enforced', 'phenomena', 'marcotte', 'deng', 'understandable', 'physico', 'happens', 'futreal', 'miki', 'koonin', 'glover', 'chai', 'hartman', 'hainaut', 'hollstein', 'iwabuchi', '133', 'conservations', 'analytical', 'ultracentrifugation', 'fil l', 'charmm', 'conflicts', 'brooks', 'solvated', 'tip3p', 'neutralized', 'ions', 'steepest', 'decent', 'equilibration', 'nvt', 'ensemble', 'npt', 'fs', 'shake', 'bonded', 'saved', 'overhauser', 'noe', 'gsolv', 'mechanical', 'bradley', 'gne', 'solvation', 'grid', 'averages', 'brca1wt', 'coworkers', 'l2', 'l3', 'h1', 'h2', 'spite', 'r248', 'r181', 'arg248', 'snapshots', 'moved', 'attraction', 'sharply', 'drifts', '1737', '1744', 'probing', 'rmsfs', 'rmsf', 'fluctuation', 'unsurprisingly', 'attracts', 'r249', 'dysfunctional', 'retaining', 'avoids', 'attracting', 'ctbp', 'clapperton', 'lending', 'unregulated', 'trapped', 'stabilities', 'cellbased', 'psr', 'alvaro', 'monteiro', 'alberta', 'edmonton', 't6g', '2h7', '492', '2136', '0886', 'ualberta', 'moffitt', 'subscriptions', 'subscribe', 'disclosure', 'pmc', 'edited', '4880', '4890', '4563', 'problematic', 'gadd45', 'heterodimeric', 'ser1655', 'gly1656', 'lys1 702', 'h2ax', 'mdc1', 'doublestrand', 'plm1', 'tnt', 'biotinylated', 'srstpsptfnk', 'srstsptfnk', 'sigmaaldrich', '646', '858', 'backgrounds', '859', 'phgr', 'articles', '396', 'qualitative', 'frog', 'puffer', 'urchin', 'compile', 'kathleen', 'cunningham', 't7', 'promotor', '35s', 'peptidebinding', 'nnsplie', 'elastase', 'chymotrypsin', 'sdspage', 'c1697r', 'phosphopeptidebinding', 'nonphosphorylated', 'tended', 'tending', 'easton', 'm1775k', 'arg1699', 'met

1775', 'i1766s', 'g1788v', 'unfolded', 'tumorsuppressive', 'preponderance', 'w1837g', 'a1843p', 'v16961', 'h1746n', 'impede', 'g1656d', 'r16991', 'pms a', 'gv', 'gd', 'graded', '706', 'a1752p', '752', 'lends', 'dicroism', 'spectroscopy', 'm1783t', 'v1808a', 'v1809f', 'v1833m', 'thermodynamics', 'fills', 'proteinprotein', 'unambiguous', 'incorporate', 'coherent', 'roadblock', 'move', 'impulse', 'cterminus', 'scandinavian', 'a1669s', 's1715r', 'w1837x', 'v1665m', 'd1692n', 'strang', 'rockefeller', '1230', '10021', '7340567', 'ext', '4729471', 'monteia', 'rockvax', 'hosted', 'hhs', 'hum', 'genet', '2016', '353', '113705', 'rad50', 'mre11', 'p95', 'disperses', 'testable', 'yeastbased', 'posing', 'vallon', 'christersson', 'swedish', 'lund', 'frank', 'scalia', 'mother', 'grandmother', 'r1699', 'truncations', 'disulfide', 'linkages', 'multicentric', 'sister', 'mothers', 'bergthorsson', 's1715', '184', 'satisfactorily', 'uncharacteristic', 'suffered', 'a1669', 'underscored', 'contradiction', 'offunction', 'mirror', 'scrutiny', 'erroneously', 'message', 'v1668del', 'underlines', 'paradoxical', 'funded', 'hoffmann', '1799', 'nonmelanoma', 'activation15', 'inducer', 'v600', '720', 'examinations', 'amplichip', 'b9', 'plex xikon', 'a431', 'dimethylsulfoxide', 'oversight', 'senior', 'vouch', 'summarises', 'keratoacanthoma', 'chronically', 'torso', 'extremities', 'egr1', 'supervised', 'dmbo', 'keratinocytes', 'tetradecanoyl', 'hydrolyzable', 'photolysis', 'caged', 'freezing', '277k', 'gln61', 'proton', 'protonation', 'vesicular', 'rab', 'topology', 'associative', 'sn2', 'abstracts', 'favouring', 'disassociative', 'deprotonation', 'hydrolysing', 'gppch2p', 'interpretations', 'dnak', 'appnhp', 'photosensitive', 'crystallisation', 'genuine', 'laue', 'synchrotron', 'monochromatic', 'uncharacteristically', 'participation', 'protonated', 'get', 'wat173', 'bobschrift', 'ground', 'temperatures', 'pure', 'ethyl', 'diastereoisomer', 'isomers', 'isomer', 'diffracted', 'protecting', 'optimised', 'achieving', 'mercury', '313', 'unphotolyzed', 'cage', 'nitrosoacetophenone', 'mainchain', 'sidechain', '5p21', 'sphere', 'equivalently', 'asp30', 'glu31', 'o3', 'g12p', 'met67', 'glu62', 'sidechains', 'glu63', 'wat175', 'shelxl', 'omitting', 'tyr64', 'ne2', 'oe1', 'occupancies', 'modelled', 'questioned', 'really', 'icm', 'drop', 'rtest', 'stereo', 'wat172', 'ser17', 'thr35', 'gly60', 'plane', 'bridging', 'heteroatom', 'superposition', 'fo', 'fc', '2fo', 't2', 'wat189', 'wat190', 'l4', 'ser65', 'ala66', 'met76', 'arg68', 'anchored', 'refines', 'reorientation', 'shortening', 'symmetry', 'twinning', 'divalent', 'og1', 'octahedral', 'asp57', 'od1', 'protectant', 'scissile', 'carbamoyl', 'spherical', 'pear', 'oriented', 'lined', 'coming', 'pictures', 'liganded', '3o', 'lys16', 'neighbouring', 'tyr32', 'harmony', 'entropic', 'reorientate', 'unwind', '166', 'movements', 'reorientations', 'unimportant', 'adjustment', 'lattice', 'biomacromolecules', 'cryocrystallography', 'rationalise', 'dielectric', 'occupation', 'substances', '028', 'artefact', 'reacts', 'things', 'isomerization', 'isomerisation', 'mant', 'characterised', 'geometric', 'esd', 'internet', 'angle', 'concerted', 'donated', 'approaching', 'superpositioning', '0040', 'alf3', 'pro12', 'gly13', 'alf4', 'cdc42gap', 'aluminium', 'provisionally', 'pentavalent', 'fluorides', 'unliganded', 'fluorines', 'rhogap', 'axial', 'bifurcated', '2o', 'lo', 'proceeding', 'nh1', 'stabilising', 'virtue', 'arg903', 'breaking', 'reconciled', 'unprotonated', 'withdraw', 'electrons', 'electrophilicity', 'fourier', 'ftir', 'secondly', 'stabilisation', 'clearer', 'pentacovalent', 'thirdly', 'conceive', 'metaphosphate', 'favoured', 'stabilise', 'fourthly', 'constellation', 'populated', 'fulfil', 'fifthly', 'envisioned', 'rigidly', 'withdrawing', 'rendering', 'nucleophilicity', 'tune', 'justification', 'permanently', 'peg', 'batch', 'seed', 'sealing', 'tape', 'hampton', 'setup', 'arc', '6281', 'lot', 'oriel', 'ug5', 'monochromator', 'h10uv', 'spex', '2nm', 'fibre', 'lumatec', 'liquor', 'nylon', 'drying', 'illumination', 'identically', 'epoxy', 'sitting', 'kept', 'spray', 'ionisation', 'nanospray', 'cryostream', 'siemens', 'rotating', 'anode', 'highstar', 'hasylab', 'bw6', 'desy', 'mar345', 'xds', 'xscale', 'p3221', 'n1', 'n2', 'rcryst', '010', '404', '408', '475', '402', '168', '232', '256', 'brackets',

'quantities', 'hkl', 'amplitudes', 'magnesium', 'plor', 'ccp4', 'surrounding s', 'fobs', 'fcalc', 'converged', '2fobs', 'asn26', 'thr50', 'thr74', 'cys11 8', 'normalised', 'mapman', 'keeping', 'cgls', 'jobs', 'restraint', 'default s', 'defs', 'swat', 'anisotropic', 'scaling', 'sulfur', 'anisotropically', 'r tab', 'triphasphatase', 'aluminum', 'ecules', '789', 'mu', 'tations', 'transduc', 'resting', 'nases', 'completes', 'reg', 'ulated', 'kinds', 'wiesmuller', 'twein', 'schmitz', 'planck', 'molekulare', 'physiologie', 'abteilung', 'structurelle', 'biologie', 'rheinlanddamm', '201', '44139', 'dort', 'mund', 'kabsch', 'medizinische', 'forschung', 'jahnstrasse', '69120', 'hei', 'delberg', 'heinrich', 'pette', 'experimen', 'telle', 'virologie', 'immunologie', 'universitat', 'burg', 'martinistrasse', '20251', 'tpresent', 'chester', 'beatty', 'fulham', 'sw3', '6jb', 'tto', 'alfred', 'mpi', 'dortmund', 'mpg', 'lowing', 'pi20gap', 'totype', 'transmitted', 'acceler', 'attainment', 'supplying', 'hydro', 'tran', 'sition', 'rasgaps', 'rasgdp', 'temary', 'stoichiometric', 'mole', 'cule', 'rasrasgap', 'alumi', '2ou', 'sciencemag', 'utc', 'jstor', 'lesop', 'iswii', 't1g', 'vqa', 'jsl', 'molscript', 'raster3d', 'gram', 'aif3', 'ball', 'lys88', 'thr791', 'calculat', 'sequently', 'nucle', 'otide', 'rebuilding', 'idues', '952', 'finement', 'rected', 'temper', 'ature', 'simulat', 'correc', 'ry', 'peared', 'trigonal', 'tetragonal', 'aif4', 'gotil', 'gott', '981', '990', 'stereochemically', '1037', '811', 'rsym', '1042', 'rotationt', 'translationt', '059', '548', 'lhi', 'ih', 'lllhi', 'ith', 'teulerian', 'q1', 'tfractional', 'ehif', 'echi', 'ilvalue', 'cules', '16c', 'su', 'perimpositon', 'cax', 'gapc', 'triangular', 'penetrating', '3145', 'phate', 'x3', 'x6c', 'x7c', 'call', 'l1c', 'raplagppnhp', 'ter', 'gapex', 'cata', 'lytic', 'cx6c', 'leu902', 'l2c', '764', 'klesllctlndr', 'eismedeattlfrattlastlmeqymk', 'atatqfvhhalkdsilkimeskqs', 'celspsklek', 'hspl', 'hlyqllwmmfsk', 'eveladsm qt1frgnslaskimtfcfk', 'vygatylqklldpllrivitssdwqhvsfevdptrlepse', 'hsnf1', 'nk1vpfitavadl', 'dlkdtqdanaifrgnslatqclemmk', 'ivgghylkvtkpvldicesssks', 'ceidpvklke', 'rngap1', 'nlakeflcdlimk', 'evekldndhlmfrgntlatkamesfmk', 'lvaddyl dstlsdfiktvqlcds', 'cevdpkqlgnvs', 'cegap', 'eriapiikaladh', 'eishltdpttifrg nt1vskmmdreamr', 'lsglhylhqtlrpvlsqivaekkp', 'ceidpskikd', 'dmgap1', '1571', 'nashilvtellkq', 'eikraarsddilrrnscatralslytr', 'srgnkylkt1rpvlqgivdnkes', 'feidkmkpgs', 'sciral', '1717', 'nathivvaqlikn', 'eiekssrptdilrrnscatrslsmlar', 'skgneylirtlpqllkkiiqnrdf', 'feieklkped', 'scra2', 'nedvntnlthllnils', 'elvekifmaseilpptlryiygclqksvqhkwpnttm', 'rtr', 'vvsgfvflrllicpailnprmf', '1331', 'sle', 'enqrnllqmte', 'kffhaiisssefpqlrvchclyqvvsqrfpqnsig', 'avgsamfl rfinpaivspyeag', 'gdnvesnkenllyyvd', 'kvfsaivgssvscptvmcdi', 'fyslrqmaakrfpnn phv', 'qys', 'avssfvflrfavailsphafh', 'nssleknrallmryve', 'vawtkilnnvhqlpknl rdvfsalrcrleaq', 'nreal', 'lissiflrfcpailspslfn', 'rsavdtnlhnlqdyve', 'rvfeaiktsadrcpkvlcqi', 'fhndlrecagehfpsnrev', 'rys', 'vvsgfiflrfapailgpklfd', 'ens', 'ekmldlfekymt', 'rlidaitssiddfpielvdicktiynaasvnfpeayi', 'avgsfvflrfi gpalvspdsen', 'sda', 'erqielfvkymn', 'ellesisnsvsyfppplfyicqniykvacekfphaii', 'aagsfvflrfcpalvspdsen', '15', '917', 'iis', 'dspspiaa', 'rtlilvaksvqnlan', 'lvefga', 'epymegvnp', 'fiksnkhrrmim', 'f1delgnvpelp', '1405', 'ild', 'kkpppri e', 'rglklnmskilqsi', 'hvlft', 'eehmrpfnd', 'fvksnfdaarrffldiasdcptsd', '1469', 'lrb', 'hypdtqtv', 'rtltlisktqiign', 'wgcqsrkrsrf', 'kksvmcef1km', 'fqee ryftdvkkf1deisstetke', 'eypsptna', 'rnltliaktlqnlan', 'fskfgg', 'ephmefmne', 'fvdrewhrmkd', 'flriissesesksg', 'ltt', 'erldaqt', 'rtltlisktiqslgn', 'lvssrs sqqtc', 'keeftvelykk', 'fcteqhv davkhflevistpshas', 'vthahd', 'kpfitlakviqslan', 'grenifkkdi', 'eeflktcsdkifn', 'selckiptnn', 'iid', 'ishlse', 'rtfislakvi qnian', 'gsen', 'fsrwpalcsqkdf1kecs', 'drifr', 'flaelcr', 'tdrt', 'rn', 'ce', 'dm', 'iral', 'ira2', 'leu910', 'cx6', 'pro34', 'jle36', 'tar', 'phosphatidi c', 'arachidonic', 'perturba', 'asp33', 'glu37', 'asp38', '949', 'conspic', 'uously', 'lys949', 'a7c', 'asn942', 'carbonamide', 'gln938', '935', 'glu950', 'ser39', '1423', 'lys935', 'rofibromas', 'tants', 'tramolecular', 'lys1423', 'neurofibro', 'fixes', '335', 'interac', 'a6c', 'penetrat', 'ha7c', 'ke', 'x2', 'ox2', 'vation', 'arg749', 'medi', 'got', 'orienta', 'gly12', 'oncoge

n', 'ly', 'planation', 'tacts', 'clos', 'tact', 'hood', 'nism', '336', 'ga1', 'aifj', 'plexes', 'om', 'i7', 'posed', 'thr785', 'nh3', 'gln6', 'postulat', 'catal', 'ysis', 'addi', 'tionally', 'communicates', 'glu3', 'uncomplexed', 'phe788', 'ala790', '903', 'flr', 'print', '901flr903', 'pl20gap', 'sug', 'gests', 'myosin', 'ics', 'nucleoside', 'monophosphate', 'bipy', 'ramidal', 'associa', 'phorous', 'oping', 'chem', 'reac', 'gonal', 'geometry', 'ond', 'cleophilic', 'wat', 'guments', 'anticatalytically', 'lographic', 'ecule', 'invari', 'cept', 'rap', 'nucleo', 'philic', 'cogenic', 'pared', 'trimeric', 'topological', 'criti', '337', 'cholera', 'gai', 'gr', 'gdalf4j', 'transdu', 'gci1', 'aif4j', 'ty', 'bridg', 'sta', 'bilization', 'guanindinium', 'connect', 'sup', 'plied', 'proxim', 'gol', 'rgs', 'thiotriphos', 'dicates', 'rgs4', 'g011lgdpa1f4', 'gxi1', 'gions', 'accelerat', 'aif', 'lyzes', 'explana', 'gin61', 'sig', 'nal', 'ture', '1970s', 'bestowed', 'spearheaded', 'homeostatic', 'barriers', 'hanahan', 'sufficiency', 'limitless', 'metastasis1', 'plasticity2', 'endow', 'evolving', 'discusses', 'fuel', 'kras4a', 'serving', 'phosphat e3', 'hydrolysis5', 'incessant', 'pathophysiological', 'point6', 'aetiolog y', 'characteristics8', 'alleles11', 'hrasg12v', 'hrasg12d', 'assays15', 'leukaemias16', 'therapy18', 'permutations', 'charts', 'outputs', 'instigated', 'types19', 'approaches21', 'impinge', 'alleles22', 'hrasq61l', 'krasq61l', 'nrasg12d', 'haigis', 'hyperproliferation', 'haemato', 'poietic', 'allele0', 'revisit', 'continual', 'erosion', 'milieu', 'arbitrators', 'fuels', 'g0', 'arrested', 'factors31', 'amphiregulin', 'areg', 'receptors33', 'quiescence36', 'comoncogenic', 'establishes', 'exit', 'smad3', 'smad2', 'upregulates', 'cdki', 'dissolution', 'transactivate', 'ccne', 'ccna', 'inaccurate', 'panoply', 'culminates', 'srf', 'atf2', 'rac1', 'gds', 'gefs50', 'ubiquitylation', 'oncogene59', 'wnt1', 'provoke', 'overdrive61', 'ckis', 'cdks', 'levels42', 'collapsed', 'forks', 'ddr', 'overworking', 'exerting', 'favour', 'improper', 'nuclei22', 'checkpoints76', 'undoubtedly', 'defence', 'corruption', 'extrinsically', 'mitochondrion', 'converge', 'bcl', 'flip', 'cclar', 'iaps', 'pathways77', 'documented77', 'apoptosis78', 'cancers79', 'engages', 'nore1', 'ste20', 'mst1', 'jnk', 'killer', 'bak1', 'stronghold', 'downregulates', 'par4', 'cd95', 'augments', 'b80', 'pawr', 'nol3', 'x188', 'tnfrsf6', 'ras90', 'programmes92', 'mapkk', 'exacerbates', 'apoptosis94', 'role95', 'manner97', 'programmes', 'rassf5', 'stk4', 'apoptosis99', 'gene98', 'enacted', 'sheer', 'cell106', '1920s', 'otto', 'warburg', 'typified', 'mitochondrial', 'aerobic', 'glycolysis107', 'ramifications', 'deciphered', 'catabolism', 'glycolysis', 'macromolecular', 'biomass', 'impinges', 'reprogramming', 'hif1', 'arnt', 'biosynthetic', 'glut1', 'slc2a1', 'glucose112', 'hexokinase', 'phosphofructokinase', 'lactate', 'dehydrogenase117', 'growth106', 'comerk', 'lkb1', 's664', 's1798', 't1362', 'checkpoint200', 'eif4', 'hif1a', 'hk', 'synthesis205', 'pfk', 'pyruvate', 'lactose', 'ldh', 'persist', 'reactions109', 'tricarboxylic', 'tca', 'fatty', 'acids205', 'organelle', 'homeostasis122', 'qualities', 'autophagic', 'upkeep', 'vivo123', 'atg5', 'atg7', 'cargo', 'nutrients126', '127', 'initiates', 'sustains', 'remodelling128', 'coma', 'fgf2', 'sp2', 'ap2', 'translation206', '208', '210', 'prostaglandin', 'endoperoxide', 'ptgs2', 'cox2', 'ccaat', 'ebp', 'pea3', '214', 'ecm', 'releases', 'mmp2', 'mmp9', 'polysome', 'mrna144', 'gtpase220', 'sprouting', 'microvessels', 'macrophages', 'krasg13d', 'growth127', 'culminate', 'boosting', 'promoter132', 'cyclooxygenase', 'prostaglandins', 'vegfa129', 'bfgf', 'migration136', 'angiogenesis130', 'gro1', 'cxcl1', 'expression130', '138', 'factors138', '139', 'endotheliom', 'tumour129', 'confines', 'potentiating', 'sprouting129', 'thrombospondin', 'tsp1', 'thbs1', 'tsp2', 'thbs2', 'cells149', 'glycoproteins', 'chemokines', 'cells151', '152', 'vasculature', 'programme', 'foreign', 'evades', 'subvert', 'immunity', 'surfaced', 'mhc', 'immunogenicity', 'cells153', 'antigenic', 'loading154', 'tap1', 'tap2', 'lmp2', 'psmb9', 'lmp7', 'psmb8', 'expression160', 'responses162', 'anergic', 'tumour158', 'surveillance167', 'immunosuppressive', 'tregs', 'mdscs', 'site168', 'tumorigenesis170', 'elucidated171',

'catenin173', 'repressors', 'snai1', 'methylation177', 'relocalization', 'catenin174', 'attachments', 'complexes183', '187', 'polymerization', 'contraction', 'microtubules', 'products188', 'leave', 'execution', 'circulatory', 'anokis77', 'pathways182', 'background192', '194', 'pathway177', 'entering', 'instructive', 'backdrop', 'emerges', 'environments', 'operates', 'daunting', 'outputpouring', 'transformation198', 'advancements', 'spatiotemporal', 'propagates', 'culminating', 'unobserved', 'y32', 'ca2', 'mediation', 'w189', 'w175', 'raps', 'unhindered', 'cystein', 'crd', 'resolve', 'rasy32f', 'soak', 'p212121', 'pdb', '2rge', 'loop7', 'o1g', 'r97', 'v109', 'r68', 'nestled', 'semitransparent', 'hexacoordinated', 'w82', 'w392', 'w366', 'd107', 'y137', 'precatalytic', 'rasq611', 'h94', 'e98', 'k101', 'y96', 'q99', 'r102', 'm72', 'y71', 'd69', 's65', 'e62', 'w391', 'w384', 'w176', 'w372', 'undiscovered', 'w396', 'n86', 'w162', 'w101', 't35', 'donating', 'accepts', 'regaining', 'backsoak', 'f32', 'occludes', 'clue', 'occluded', 'masked', 'feat', 'succeeded', 'loose', 'character', 'augmenting', 'accept', 'donate', 'reconsidering', 'displace', 'outcompete', 'discourse', 'reconsidered', 'allosterically', 'mystery', 'completion', 'dimyristoylglycerol', 'phosphocholine', 'bilayer', 'r128', 'r135', 'q165', 'venue', 'lieu', 'futile', 'extracranial', 'a1099t', 'r1464stop', 'uncharacterised', 'd1091n', 'f1174i', 'a1234t', '1174', 'alkf1174', 'alki1250t', 'originates', 'car', 'martinsson', 'butrynski', 'characterise', 'nucleolar', 'shioda', 'sensory', 'vernérsson', 'enigmatic', 'griffin', 'nikova', 'arber', 'debelenko', 'alkf1174i', 'alkf1174s', 'alkm1166r', 'universally', 'encouragingly', 'alka1099t', 'alkr1464stop', 'initiative', 'alkt1087i', 'alkd1091n', 'alkt1151m', 'alka1234t', 'differentiating', 'greene', 'tischler', '1976', 'alkwt', 'mab31', 'pegfpn1', 'exploiting', 'warmuth', '1278', 'tyr1278', 'yy', 'tartari', 'tyr1604', 'pgmr', 'directs', 'photoreceptors', 'discs', 'hexagonal', 'ommatidia', 'disorganised', 'interommatidial', 'bristles', 'larval', 'draws', 'exercise', '1604', 'alkr1275q', 'met1166', 'irk', '1166', 'lemon', 'obstructs', 'inaccessible', 'cys1097', 'phe1174', 'initiation', 'ala1099', 'sticks', 'thr1151', 'twisted', '1464stop', 'transformation', 'mcduff', 'alkr1275', 'neurblastoma', 'shimada', 'gct', 'unfavourable', '908', 'morozova', 'r1646stop', 'constructions', 'eurofins', 'mwg', 'ebersberg', 'palk', 'pstat3', 'moog', 'lutz', 'waltham', 'electroporator', 'ingenio', 'lcc', 'neuritecarrying', 'cfl', 'judged', 'foetal', 'peprotech', 'rocky', 'stockholm', 'brien', 'naked', 'brand', 'perrimon', 'alk1234t', 'alk1464stop', 'puast', 'bestgene', 'chino', 'bloomington', '9146', 'glutaraldehyde', 'instar', 'larva', 'mab153', 'fluoromount', 'birmingham', 'cohesin', 'telophase', 'rad21', 'smc1a', 'combs', 'polycomb', 'homeotic', 'erythropoiesis', 'hscs', 'hpcs', 'h3', 'trimethylation', 'h3k27me3', 'h3k4me3', 'smc3', 'stag1', 'stag2', 'chromatids', 'cohesion', 'merkenschlager', 'odom', 'ctcf', 'ccctc', 'congruent', 'mart', 'nez', 'hsc', 'multiprotein', 'truncates', '1010', '401', 'pulling', 'embrace', 'h2b', 'reintroducing', 'dysregulating', 'mitosis', '1600', 'tf', 'visualizing', 'klf4', 'btg1', 'chop', 'slc', 'skewing', 'cbfb', 'fusion', 'repressive', 'prc2', 'retinoid', 'looping', 'tfs', '1060', 'sectional', 'dominance', 'bcor', '023', 'heterodimerize', 'expression12', 'acetylation', 'pluripotency', 'hdac8', 'lifespan', 'deacetylation', 'saml', 'pds5b', '149', '892', 'hgc', 'jp', 'heuristic', 'tcgahome2', 'jsp', 'prioritized', 'deseq2', 'u133plus2', 'gcrma', 'limma', 'ebayes', 'residuals', 'dichotomized', 'medians', 'sas', 'cooccurring', 'ipss', 'stratifying', 'lenalidomide', 'hypomethylating', 'decitabine', 'azacitidine', 'nonresponse', 'vaf', 'tet2', 'raeb', '0032', 'cytarabine', 'idarubicin', 'exemplary', 'vafs', 'nonmutational', 'runx', 'founding', 'cohesins', 'cooperative', 'programming', 'leukemogenic', 'omission', 'pathophysiologic', 'brd4', 'downmodulation', 'surveying', 'meiotic', 'pms1', 'ternary', 'exo1', 'ntg2', 'mip', 'acknowledgments', 'corrects', 'integrity1', 'maturation3', 'tissues4', 'eukaryotes', 'structures6', 'reaction9', 'constituted', 'latent', 'dqha', 'x4', 'subunits10', 'e707k', 'mmr11', 'pms2e702k', 'predisposition13', 'mutls', 'homodimeric14', 'ntd', 'dynamica

l', 'ntds', 'mmr18', 'ctds', 'heterodimerization21', 'ntd18', 'connector', 'p ale', 'subtilis', 'enumeration', '522', 'prokaryotes14', 'organized', 'subdoma in24', 'eubacteria', 'resident', 'excision28', 'prokaryotic', 'mutations29', 'debated30', 'mlh2', 'respectively15', 'sgs1', 'recq', 'glycosylase', 'apurinic', 'apyrimidinic', 'lyase', 'e682a', 'recombination36', 'cys769', 'mn2', 'activity26', 'aquifex', 'aeolicus', 'activity38', 'homoduplex', 'kdapp', 'probe18', 'anomalous', 'selenomethionine', 's24', '649', '822', '873', 'protrusion', '509', '524', '805', 'superimpose', 'connectors', 'rcsb', 'neisseria', 'gonorrhoeae', '3ncv', '3kdg', 'buries', 'homodimers24', 'pseudosymmetric', 'scer', 'spom', 'schizosaccharomyces', 'hsap', 'atha', 'arabidopsis', 'thali ana', 'dmel', 'cele', '506', 'structured', 'asp870p', 'arg547m', 'lys665m', 'superscripts', '759', 'glu767', 'arg768', 'farther', 'cys769m', 'ferc', 'y2 h', 'd870a', 'r547a', 'vivo36', 'c10', 'acr', 'hgrp', 'alignments12', 'glu707 p', 'his703p', 'cys817p', 'his850p', 'cys848p', 'endonucleolytic', 'vivo11', 'asp701p', 'capping', 'interaction39', 'd699n', 'asp701', '3rbn', 'tetrahedral', 'graphic', 'pbeq', 'solver', 'kcal', 'mlh1p', 'variant21', 'f33', 'conformation40', 'pinpointed', 'interaction36', 'glu682', 'studies36', 'glu682m', '1511a', 'm623a', 's625a', 'e680a', 'leu511', 'ser24n', 'tyr630m', 'tyr26n', 'phe27n', 'phe447e', 'n6', 'n8', 'n13', 'n16', 'n20', 'mp8', 'y26f', 'f27a', 'f27y', 'residues36', 'ser24', 'ser445', 'tyr630', 'tyr26', 'phe447', 'met626', 'glu629', 'phe27', 'phe449', 'cys685', 'leu686', 'endowed', 'syndrome5', 'modeller41', 'p5811', 'd601g', 'a589d', 'pcna42', 'med1', 'mbd4', 'fan1', 'angiomentin', 'brip1', 'slashes', 'dim', '446', 'arg547', 'asp870', 'strain s45', 'd870n', 'r547k', 'functional36', 'synergetic', 'interaction21', 'cth', 'previously21', '2ferc', 'sequences36', 'packs', 'residues21', 'f766a', 'c769s', 'cations', 'assays10', 'site36', 'grenlyf', 'yf', 'dyad', 'dyads', 'task s', 'avoidance36', 'meiosis', 'resolving', 'holliday', 'over37', 'analysis42', 'study47', 'degenerated', '572', 'qxxlxxfa', 'guides', 'medication', 'reg imens1', 'syndrome1', 'cancers1', 'mortality9', 'designating', 'schemes', 'cl asses10', 'variome', 'hvp', 'categorization', 'models13', 'data17', 'advice', 'lsdbs', 'opinions', 'curation', 'categorizing', 'information18', 'transparen cy', 'lsdb', 'users', 'settings18', 'genes19', 'curation18', 'microattribution', 'approach24', 'researcher', 'orcid', '458', 'standardize', 'submissions', '730', 'lodged', '932', '746', 'submitters', 'discordance', 'assay22', 'infer ences', 'database23', 'delphi', 'process25', 'hvp26', 'system10', 'inconsistencies', 'teleconferences', 'meetings', 'retrospectively', 'justifications', 'teleconference', 'items', 'reassessment', 'tier', 'flowchart', 'publication s47', 'more49', 'cmmrd', '276300', 'faced', 'paucity', 'flowcharts', 'interrupting', 'private', 'risks27', 'attained', 'scarce', 'unobtainable', 'submitter', 'downgrading', 'upgrading', 'misclassifications', '765', 'lgdup', 'lgde l', '391', 'pathogenic30', 'incentivize', 'scholarly', 'citation', 'articles33', '015', 'microattributions', '763', 'predictions34', 'physicochemical', 'roc', '481', 'bimodal', 'mapp', 'scores34', 'interruption', 'collaborative', 'counseling10', 'internationally', 'calibration', 'assays36', 'accrual', 'hemoglobinopathies24', 'formats', 'standardization38', 'mutations28', 'iterations', 'stipulating', 'contemplate', 'risk45', 'mmr46', 'communities', 'guideline', 'dated', 'clinvar', 'expert', 'stakeholders', 'centric', 'transparent', 'chair', 'continents', 'disciplines', 'governance', 'council', 'responsibility', 'invite', 'legal', 'standing', 'mutalyzer50', 'method25', 'registry16', 'analysis15', 'attending', 'antonio', 'comment', 'amended', 'clarity', 'phenotype53', 'data54', 'consistency', 'collation', 'project54', 'automatically', 'attendance', 'randomization', 'attendee', 'pertaining', 'curator', 'attendees', 'spreadsheet', 'reviewer', 'circulated', 'collated', 'elsewhere34', 'data55', 'pubmed', 'pmid', 'searchable', 'preferably', 'credit', 'publically', 'nanopublication', 'methods16', 'data16', 'senter', 'analysis27', 'phenotype16', 'variants34', 'variants13', 'stata', 'h0', 'azd8055', 'torki', 'm2327i', 'a2034v', 'rr1', 'f21081', 'rr2', 'treatment5', '4ebp', 'torki6', 'resistanc

e9', 'fkbp', 'pp242', '4jt5', 'm2327', 'torkis', 'emergent', '1fap', 'avidit y', 'bivalent', 'hyperactivate', 'c40', 'selective17', 'rapalink', 'optimum', 'moe18', 'glycol', 'azide', 'alkyne', 'cycloaddition', 'synthesize', 'shortes t', 't37', 'halves', 'fk506', 'f2039s', 'molecules19', 'hybrids', 'life20', 'lasted', 'others21', 'safely', 'intermittently', 'observed5', 'exploitatio n', 'gpcrs22', 'bitopic', 'aware', 'cdk2', 'peptidic', 'inhibitor23', 'pif', 'pocket24', 'ksr', 'mek25', 'htb', 'dme', 'f12', 'crl', '1573', 'mycoplasm a', 'bioscience', 'rad001', 'ku006', 'wy354', 'tocris', 'exposing', 'sent', 'kapa', 'nimblegen', 'seqcap', 'described26', '487x', 'mmff94x', 'interatomi c', 'am1', 'bcc', 'crosslinker', 'tethering', 'lowmodemd', 'energies', '080 1', 'montreal', 'svl', 'luminescent', 'contents', 'analyst', '40601', '9234 1', '53641', '48581', '94591', '94511', 't70', '94551', '4970s', '2972s', '22 80s', 'f1804', 'ttigfp', 'mluex', 'rtta3', 'phoenix', 'minipreps', 'improm', 'oligos', 'genewiz', 'dispensing', 'solubilised', 'microplates', 'greiner', 'hp', 'd3000', 'dispenser', 'cheng', 'prusoff', 'competitive28', 'permitte d', 'estradiol', 'innovative', 'captisol', 'width2', 'published29', 'etc', 'nihms774093f5', 'jpg', 'nihms774093f6', 'nihms774093f7', 'ch2', 'nihms774093f 8', '293h', 'nihms774093f9', 'dot', 'nihms774093f10', 'readded', 'nihms774093 f11', 'b0', '4852', 'ccrcc', 'e2033v', 'wagle', '6393', '3347', '00006559', 't01', 'im5', '0000645', 'im3', '0000614', '0004920', 'msk', 'e2014k', 'snu7 38', 'l2016r', 'b141', 'haeatopoie', '8441', 'chrcc', '0003073', 'i2017s', '2 787', 'squ', 'w2023g', 'bcp1', 'haematopoie', 'g2030v', 'b5', 'a11e', 'a2056 v', 'a25u', 'm2057i', '6598', 'q2063fs', 'h072999', 'amc', 'k2066r', 'ncih', '1792', 'q2072', '0000112', 'q2072r', 's05', '31806', 'tp', 'cscc', 'x2073', '0005266', 'r2076l', 'ee', 'a2gc', 'd2077e', '0002050', 'w2084c', 'ax', 'a0j 0', 'y2088d', 'cq', '5327', 'm2089i', '0005159', '12097f', 'a1y0', 'w21011', 'c5', 'a1bm', 'd2191h', '4370', 'r2193c', '0004376', 'f22021', 'a0yx', '481 0', 'jw', 'a5vl', 'a3hz', 'a00k', 'a6', '6141', 'f4', '6806', 'a1d4', 's12', '23181', 'cj', '5679', '7828', 'prcc', '0000208', 't02', '0005214', 'ukrv', 'a0uf', 'a0uv', 'a0vx', '5396', 'hec251', '4105', '4887', '12230v', 'snu119 6', 'biliary', 's2231w', 'kmh2', 't2232i', 'cw2', 'intest', 'g2238d', 'w2239 c', 'w2239ffs', 'ds', 'bla', '084', 'p2241s', '0296', 'dfarber', 'dk', 'a6b 6', 'r2251q', 'hec1a', 'r2254m', 'snu520', 'h2265p', '5126', 'r2266p', '523 9', 'a2272s', 'eb', 'a3y7', 'a2272v', '0005309', 'q2282p', '8363', 't2294a', '0001171', 'd2298h', 'e2311k', '0001042', 'a2325v', '0003529', 'v2330i', '36 66', 'i2333m', '5691', '12334v', 'a4gu', 'r2339', '1540', 'h2340r', 'hb', 'a4 3z', 's2342y', 'm2345v', 'escc', '148t', 'g2351e', 'fh5pj', 'h2355r', 'g2359 r', 'a54y', 'r2368q', 'hec108', '0002079', '12383f', 't2384i', 'bh', 'a0hp', 'e2388q', 'intes', 'v2389', 'cosm51966', 'v2406a', 'a7', 'a5zv', 'd2412h', 'h061142', 'd2412v', 'ek', 'a3gk', 'e2419k', 'cl1147', 'a2420p', '0003254', 'a2420v', '0004798', 'a4fc', 'a421', 'd2424n', 'st486', 'haematopoi', '12427 r', '0122', 'l2427q', 'kn', '8437', '0005355', '0005044', '0000839', 'molti 6', 'haematop', 'r2430m', 'g4', '6322', 'r2441q', 'p53s', 'saos', '290', '29 2', 'acetylated', 'r175', 'g245', 'r273', 'r282', 'accumulating', 'bh3', 'int errelation', 'precious', 'subclass', 'r175p', 'transactivated', 'p53aip1', 'p uuma', 's121f', 'harmonized', 'lipofection', 'effectene', 'pcaggs', 'pcr259', 'p53wt', 'p53mts', 'fullerton', 'ceq2000xl', 'nanograms', 'mesh', 'sorter', 'subg1', 'pepstatin', 'debris', 'electrophoretically', 'f1393', '6243', 'p21 s', 'pmdm2s', 'pbaxs', 'psigmas', 'paip1s', 'ap000920', '104403', '105010', 'pgl', 'frag1', 'p53mt', 'gradation', 'tigr', 'mev2', 'r175h', 'g245s', 'r24 8w', 'r249s', 'r273h', 'r273c', 's121c', 's121a', 's121y', 'q144r', 'p153h', 'r156c', 'i162m', 'f212y', 'h214q', 's215c', 'e221q', 'r290g', 'k291e', 'k29 1q', 'k291t', 'k292t', 'k292i', 'thrice', '053', '095', 'reevaluation', 'reev aluated', 'waf1', '064', '262', '855', '790', '664', '046', '910', '096', '03 0', '078', '331', '983', '914', '864', '373', '851', '291', '951', '612', 'sc atter', 'l1', 'transactivates', 'labor', '243', 'rras2', 'redefining', 'types 1', 'broadening', 'patterns3', 'methods1', 'ratiometric', 'types12', 'types1

5', 'principled', 'mutability', 'consortia', '694', '25th', '75th', 'percentile', 'types4', 'disseminated', 'paradigms', 'r132', 'r1446', 'leukemias17', 'reaffirm', 'basket', 'e322', 'ep300', 'd1399', 'carcinosarcomas', 'pontine', 'nasopharyngeal', 'unconventional', 'd1010', 'm582', 'met18', 'states20', 'belies', '736', 'and2a', 'incidentally', 'finer', 'mutations21', 'g12r', 'myelomas', 'therapy10', 's310f', 'r88q', 'r132h', 'growing8', 'mutations24', 'polo', 'mutations25', 'e545', 'h1047', 'cohorts27', 'augurs', 'type28', 'i24', 'i250029', 'd1067', 'h1451', 'm60', 'responses30', 'facets', 'and3a', 'gq60gk', 'converts', 'p29s', 'melanomas12', 'vitro34', 'patients35', 'a159v', 'mutation36', 'pak1', 'gtpase37', 'origins38', 'oncoproteins41', 'q72', 'pathway42', 'kinases44', 'mutable', 'therapy30', 'distinctions', 'cancer8', 'spur', 'available45', 'ncbi36', 'liftover47', 'urls', 'standardization', 'gencode', 'vep', '7748', 'vcf2maf', '424', '827', '866', '3231', '4bp', 'cghub', 'rnaseqv2', 'rsem', 'anticipating', 'nk', 'pk', 'mutagens16', 'trinucleotides', 'mt', 'ctft', 'ft', 'occurrences', 'compute', 'mc', 'cmtnt', 'cg', 'nlg', 'lg', 'gmtlg', 'overestimating', 'warmspots', '99th', 'signifies', 'exceedingly', '20th', 'g20', 'allp', 'eq', 'padded', 'equaled', 'yekutieli49', 'misattributed', 'cbiportal45', 'clia', 'center50', 'subclonally', 'originated', 'shannon', '12bp', '24bp', 'homopolymer', '10bp', 'alignability', 'crg', 'herring', 'selection4', 'unaligned', 'a159', 'nup93', 'r60', 'h28', 'reprocessed', 'peline30', 'e14k', '880', 'census', 'sources1', 'assed', 'abscn', 'seq51', 'automation', 'described52', 'rac1p29s', 'rac1q61r', 'rac1a159v', '3xflag', 'cat', '8815', 'ko', 'privates', 'tetrameric', 'giaccia', 'kastan', 'bp3', 'pig3', 'aip1', 'apaf', 'olivier', 'crook', 'rowan', 'campomenosi', 'p73', 'kaghad', 'kaelin', 'melino', 'incompetent', 'modular', 'tetramerization', 'como', 'concerned', 'inducers', 'agami', 'gong', 'heterotetramers', 'homotetramers', 'heterotypic', 'yig397', 'yph', 'inga', 'kelly', 'monti', 'cen', 'homotypic', 'ade', 'transformant', 'darker', '57k', 'constrain', 'soussi', 'r7', 'tablet he', 'r282w', '176k', 'marin', '72r', 'thereof', 'nonisogenic', 'chi2', 'analysing', 's10', 'fontemaggi', 'stiewe', 'flores', 'conditioned', 'r72', 'transdominant', 'constructing', 'appealing', 'pcyc1', 'flaman', 'adh1', 'difco', 'milano', 'ori', 'pvui', '6556', 'discrimination', 'liac', 'variantsfor', 'sgrai', 'endonucleases', 'poisson', 'ptpase', 'cowden', '158350', 'bannayan', 'zonana', 'bzs', 'ruvalcaba', 'riley', '153480', 'worthy', 'r233x', 'inthe', 'sufferers', 'phts', 'soma', 'neurosciences', 'malignancy13', 'phenotypical', 'for19', 'against2', 'c124s', 'g129e', 'allele21', 'aggravate', 'g129r', 'h118p', 'h123q', 'e157g', 'f241s', 'd252g', 'n276s', 'd326n', 'recognised', 'radiolabelled', 'vesicles', 'immunoprecipitating', 'titre', 'immunodetection', 'hypertrophy', 'behavioural', 'e15', 'flox', 'rfp', 'immunocytochemistry', 'lipofected', 'untransduced', 'workflow', 'comparator', 'categorisation', 'a39p', 'n48k', 'l108p', 'l112p', 'r130l', '0013', 'cys124', 'h93r', 'exacerbate', 'characterises', 'dimerisation', 'tetrakisphosphate', 'ptens', 'mmac1', 'tep1', 'omim158350', 'omim153480', 'lhermitte', 'duclos', 'embryonically', 'arresting', 'auxillin', 'perfect', 'ptp3', 'p4', 'd3', '417', '162', 'dorecived', 'grants', 'ministry', 'sports', 'tohoku', 'seiryomachi', 'aoba', 'ku', 'sendai', '980', '8575', '717', '8547', '8548', 'chikashi', 'idac', 'lmv', 'multilamellar', 'vesicle', 'phosphatidylserine', 'phosphatidylethanolamine', 'phosphatidylcholine', '3147', 'advantageous', 'pqe30', 's10n', 'y16c', 'g20e', 'y27s', '142r', 'h61r', 'y68h', 'c71y', 'h93y', 'c105f', 'd107y', 'l112r', 'a121p', 'c124r', 'r130g', 'r130q', 'v133i', 'm134l', 'c136y', 'y155c', 'g165r', 's170n', 's170r', 'r173c', 'r173h', 'r173p', 'y174n', 's227f', 'g251c', 'k289e', 'd331g', 'f341v', 'k342n', 'v343e', 'l345q', 'f347l', 'v369g', 't401i', 'megaprimer', 'm15', 'prep4', 'mid', 'a600', 'bacteriolysis', 'sonication', 'ted', 'nanosep', 'pall', 'filtron', 'northborough', 'glycerin', 'maehama', 'dixon', 'hcooh', 'hcoohn4', 'ag1', 'x8', 'equilibrated', 'dic8', 'biomol', 'plymouth', 'a620', 'lmvs', 'phosphoglycerides', 'koh', 'cacl2', 'trichloroacetic', 'daiichi', 'phosfig', '3148', 'phatase', 'a2g', '314

'9', 'hcxxgxxr', 'coohterminal', 'hcxxgrxxr', 'georgescu', 't348i', '342', 'preserve', 'pdz', 'dlg1', 'mast205', 'coprecipitated', '3150', 'phtss', 'asds', 'tpte', 'tpip', 'tensins', 'auxilins', 'd92n', 'asp92', 'topic', 'monoester', 'hydrolases', 'plants', 'fungi', 'voltage', 'ciona', 'intestinalis', 'clathrin', 'auxilin', 'counteracting', 'dd', 'neurodevelopment', 'buffering', 'riche r', 'hckagkgr', 'arg130', 'delimited', 'bioassays', 'conformed', 'pip3in', 'a 121v', 'g129a', 't131a', 'gly129', 'his93', 'val166', 'gln171', 'asp162', 'd9 2a', '1d5r', 'strongylocentrotus', 'purpuratus', 'dictyostelium', 'discoideum', 'largedownload', 'yph499', 'ala121', 'ala126', 'nucleocytoplasmic', 'micr oscopically', 'inlet', 'i1221', 'i122v', 'h123y', 'c124n', 'k125l', 'k125m', 'k125r', 'a126d', 'a126g', 'g127n', 'k128r', 'k128t', 'g129d', 'r130a', 'r13 0k', 't131i', 't131l', 't131s', 'divergences', 'a121e', 'i122s', 'h123d', 'k1 25e', 'a126s', 'a126v', 'g127e', 'k128n', 'g129v', 't160i', 'r161g', 'd162g', 'g165e', 'g165v', 'v166i', 't167a', 'i168f', 'p169h', 'y88c', 'f90l', 'e91a', 'd92e', 'd92g', 'd92h', 'd92v', 'h93d', 'h93q', 'n94i', 'p95l', 'p96q', 'mala chite', 'phosphatydilinositol', 'thr308', 'outlines', 'recovering', 'g44d', 'm134t', 'y176c', 't202i', 'w274l', 'a34d', 'm35r', 'g36r', 'h61d', 'y68d', 'l181p', 'v217n', 'p246l', 'r335l', 'l345v', 'hub', 'inactivator', 'complexi ng', 'consonance', 'tensin2', 'updates', 'cf', 'reproduces', 'neurobehavioral', 'heterozygotic', 'diminution', 'synapsis', 'hippocampus', 'mata', 'ammonium', 'uracile', 'microliter', '421b', '425a', 'mab', 'te2000u', 'c4742', '12e r', 'aquacosmos', 'hcimage', 'psg5', 'ycplg', 'pyes2', 'pyes3', 'jm', 'mast', 'd816', 'd835', 'd835y', 'd835v', 'd835h', 'd835e', 'd835n', 'worsen', '243 4', '2439', 'hematology', 'fms', 'tk1', 'tk2', 'accompanies', 'duplicated', 'dimerized', 'mitogenactivated', 'fdc', 'blastoma', 'nagoya', 'aichi', 'kuma moto', 'nagasaki', 'saitama', 'kyoto', 'prefectural', 'metropolitan', 'komago me', 'tochigi', 'chiba', 'okayama', 'takeda', 'hitoshi', 'kiyoi', 'tsurumai', 'showa', '8560', 'bloodjournal', 'guest', 'mastocytomas', 'hmc', 'atl', 'volunteers', 'kirin', 'brewery', 'i836', 'eco', 'rv', '17f', 'ccgccaggaaacgtgctt g', '17r', 'undigested', 'chatsworth', 'pt7blue', 'oliver', 'rosnet', 'antihu man', 'biotech', 'm7', 'ttg', '2435', 'bedford', 'chemoluminescence', 'transf ast', 'immunotech', 'marseille', 'jalsg', 'lactic', 'meyer', 'statistic', 'ab acus', 'berkeley', 'haeiii', '2436', 'hepatosplenomegaly', '0025', '047', '00 43', 'mocktransfected', 'logrank', 'looks', '3109', '632', 'inv16', 'frenchamerican', '2437', 'd816y', 'd816f', 'additively', 'karyotypic', 'cally', 'd81 4', 'd802', 'd814y', 'd814v', '2438', 'itds', 'mln518', 'd835a', 'd835del', 'i836del', '2867', '2872', 'flk', 'stk', 'rtkiii', 'pdgfrs', 'rtkiiis', 'mul tipotent', 'vav', 'mllrearranged', 'hyperleukocytosis', 'asp835', 'asp835ty r', 'asp835ala', 'brigham', 'millennium', 'prepublished', 'edition', '1182', 'ca66996', 'dk50654', 'scor', 'doris', 'duke', 'charitable', 'declared', 'je nnifer', 'francis', 'asp835glu', 'asp835gly', 'd835g', 'asp835his', 'asp835as n', 'asp835val', 'asp835deletion', 'ile836metarg', 'i836mr', 'ile836thr', 'i8 36t', '841', '840gs', 'factorindependent', 'ct53518', 'quinazoline', 'transfo rmer', 'ggattggctcgagctatcatgagt', 'ggctcgagatatgagtgattccaac', 'cttctgaggatccgac', 'turbo', 'proofreading', 'd835ef', 'ggattggctcgagatcatgagt', 'd835hf', 'ggattggctcgacatatcatgagt', 'd835nf', 'ggattggctcgaaatatcatgagt g', 'd835vf', 'ggattggctcgagttatcatgagt', 'd835delf', 'ctttggattggctcaatcatgagt', 'tgctccagactgcctggaa', 'ccggactcgatcaaatct', 'd835er', 'cactcatgat ctcctcgagcc', 'd835hr', 'cactcatgatatgtcgagcc', 'd835nr', 'cactcatgatatttcgagcc', 'd835vr', 'cactcatgataactcgagcc', 'd835delr', 'andrea', 'danafarber', 'ce lltiter96aqueousone', 'smoothed', 'originpro', 'originlab', 'northampton', 'm illiliter', 'ethylenediaminetetraacetic', 'piscataway', 'protran', '2868', 'phosphop44', 'antimouse', 'antirabbit', 'cytokineindependent', 'colorimetric', '2869', 'immunoblotted', 'devoted', 'antierk', '2870', 'quizartinib', 'ac22 0', 'f691l', 'dfgin', 'crenolanib', 'mine', 'documents', 'editorial', 'neil', 'shah', 'phd', '505', 'parnassus', '94143', 'nshah', '3303', '2390', '2392', 'prevailing', 'informing', 'sawyers', 'described5', 'plx3397', '1rjb', 'soa

p', 'patients1', 'reported5', 'mutations6', '829', '856', 'terminus16', 'dipole', 's838', '1t45', 'holo', 'm837', 'rationalizing', 'displaced', 'f830', 'p henyl', 'intermediary', 'achievable', 'misguided', 'dicer1', 'osteoprogenitor s', 'shwachman', 'diamond', 'phenocopied', 'patterning', 'osterix', 'osteoblastic', 'cells1', 'osteoprogenitor', 'ocd', 'cd45', 'cd31', 'osteogenic', 'deposition', 'osteoblasts', 'osteocalcin', 'texture', 'trabecular', 'osteoblast', 'osteoclast', '1i', 'myelodysplasia', 'leukopenia', 'erythrocyte', 'mitigating', 'hypercellular', 'immunophenotypically', 'sca1', 'cd150', 'cd48', 'micromegakaryocytes', 'hypolobulated', 'hyperchromatic', 'dysgranulopoiesis', 'criteria12', 'engraft', 'transplanting', 'longstanding', 'mep', 'duplets', 'vascularity', '2j', 'b12', 'folate', 'environmentally', 'sjl', 'chimerism', 'cytopenias', 'committed', '13a', 'demonstrated20', '14a', 'bones', 'recapitulated', 'myeloblastic', 'monocytic', '4i', 'monoblastic', 'mice12', '656', 'g sea', 'commitment', 'pathway22', 'knocking', 'intercrossing', 'lymphopenia', 'neutropenia', '18b', '18c', 'microenvironmental', 'precede', 'quiescence', 'signaling30', 'pathophysiology', 'participant', 'mice9', 'mice10', 'ptprca', 'pep3b', 'boyj', 'gy', 'congenic', 'engraftment', 'deciphering', 'pri', 'dgcr8', 'exportin', 'shuttles', 'ran', 'cutting', 'argonaute', 'risc', 'cuts', 'biogenesis', 'ribonuclease', 'drosha', 'digeorge', 'necessitating', 'kar yopherin', 'cofactors', 'bypassed', 'proceed', 'ago2', 'slicer', 'tentative', 'ties', 'mucoepidermoid', 'throat', 'anthracycline', 'pleuropulmonary', 'pp b', 'mesenchyme', 'nephroma', 'erms', 'sertoli', 'leydig', 'slct', 'gonadal', 'multinodular', 'goiter', 'mng', 'illnesses', 'intraocular', 'etiological', '823', 'unanswered', 'suspicious', 'haplo', 'degeneration', 'macular', 'degenerative', 'perplexing', 'operational', 'feed', 'overly', 'simplistic', 'lowers', 'influential', 'undecided', 'dicer', 'reinforced', 'resectable', 'hct', 'tarbp2', 'tnrc6a', 'proposition', 'obstacle', 'polyethylenimine', 'nanoparticles', 'developments', 'sentinel', '296', 'granulosa', 'regressed', 'fibroma', 'lpd', 'mcp', 'saliva', 'physicians', 'suspicion', 'draft', 'abi3730', 'beckmancoulter', 'chromatograms', 'fhcrc', 'dysgerminoma', 'sisters', 'cyst', 'myogenic', 'contralateral', 'unilateral', 'enteric', 'sarcomatous', 'hemorrhagic', 'photomicrographs', 'cords', 'basophilic', 'unintentional', 'inadequately', 'presumption', 'predominance', 'draw', 'imbalances', 'metastasizing', 'peutz', 'jeghers', 'seminomas', 'seminoma', 'pleiotropic', 'familiar', 'professionals', 'alert', 'virilization', 'seek', 'answered', 'gmgs', 'derepresse d', 'pelger', 'hspcs', 'starts', 'mpp', 'cmp', 'granulopoiesis', 'mirs', 'tunes', '146a', 'e7', 'reconstitute', 'thymus', 'eyfp', 'monocyte', 'david', 'koch', 'toe', 'biologio', 'erasmus', 'cytospins', 'dcs', 'atmosphere', 'cellgro', 'nwald', 'giemsa', 'dmlb', 'abs', 'tibias', 'crushed', 'mortar', 'sieve', 'erythrocytes', 'cd11b', 'cd45r', 'b220', 'hspc', 'sca', 'cd16', 'cmps', 'submandibular', 'granulocytes', 'cd11c', 'reanalysis', 'suppliers', 'rlt', 'messageamp', 'arna', '430', 'isogen', 'mas', 'minelute', 'cleanup', 'cards', 'pre amplification', 'targetscan', 'kolmogorov', 'smirnov', 'myelopoiesis', 'lsk', 'lox', 'mx', 'replatings', 'neutrophil', 'replated', 'asymptotic', 'micrographs', 'bilobed', 'ly6g', '784', 'messages', 'cran', 'packages', '368', 'bcl211 1', 'hmga2', 'nonregulated', 'mpps', 'sterol', 'derepression', 'depletes', 'reintroduction', '125a', 'agrees', 'hampers', 'complicates', 'silence', 'stemness', 'counteracts', 'neuroectodermal', 'goitre', 'predisposes', '1136', 'jmg', 'almetric', 'tweeted', 'pages', 'referenced', 'wikipedia', 'readers', 'm endeley', 'exported', 'unwound', 'genuinely', 'doc', 'inline', 'marsden', 'multicentre', 'icr', 'deceased', 'inception', 'contacted', 'analyser', 'bidirectional', 'histochemical', '5th', '6th', 'pnet', 'fossa', 'cousin', 'conclusively', 'grandparents', 'muscular', 'dystrophy', 'localise', 'hemizygosity', 'unproven', 'door', 'policy', 'undertaking', 'discussions', 'thymoma', 'dislodges', 'akts', 'pdk2', 'proteus', 'hemimegalencephaly', 'map2', 'n3', 's218e', 's222d', 'ev', 'book', 't81y', 'd323a', 'r23a', 'n53a', 'f55y', '178t', 'q79e', 'e191a', 't195i', 'v270a', 'v271a', 'l321a', 'd325a', 'n328a', 'pt308',

'ps473', '735e', '0002158', 'y18s', 'i19e', 'q59e', 'r200a', '1202f', '1362r', 'phosphorylation', 'nonsmall', '152r', '152', 'v270', 'v271', 'y326', 'r328', '152a', 'd323h', 'r96', 'k189', 'e198', 'q79k', 'k189n', 'polarized', 'acini', 'hollow', 'lumen', 'multiacinar', 'f351', 'r96w', 'and3a3a', 'ps6rp', 'vp16', 'behaving', 'movies', 'd324h', 'l51r', 'd320h', 'hypoglycemia', 'r371h', 'v901', 'r1011', 'q124l', 'g171r', 'inspections', 'and2c', 'necropsies', 'posttransplantation', 'assaying', '692', '929', 's8a', 's8e', 's8b', 's8c', 's8f', 's8d', 'reconstituting', 'e17', 'd323', 'multidomain', 'myelo', '005154', 'isozymes', 'isozyme', 'hm', 'phosphatidylinositide', 'tractability', '2206', '3o96', 'proteinaceous', 'proteolyzed', 'daltons', 'fluorimetry', 'crystallize', 'solve', 'nestles', '312', 'attaining', '526', 'pleased', 'ip4', 'vl3', 'discovering', 'imidazoquinoline', 'tricyclic', '2119', 'optimizing', 'conformer', 'configured', 'brafi', 'meki', 'eventual', 'pik3r2', 'e545g', 'n561d', 'pervasive', 'biopsied', 'transcriptomes', '48h', 'coinciding', 'gefitinib', 'm238', 'restricts', 'lift', 'amplifies', 'ly294002', 'repressible', 'weakest', 'theses', 'akti', 'washout', 'bkm120', 's11', 'growths', 'unleash', 'intermittent', 'achieves', 'durability', 'surmise', 'selumetinib', 'stereotypic', 'constraints', 'renewed', 'adaptively', 'innate', 'wm', 'plvx', 'protamine', 'vemurafenib', 'ser9', 'ser338', 'deparaffinization', 'rehydration', 'bleaching', 'steamer', 'aperio', 'studio', 'lips', 'durations', 'permeabilized', 'lsr', 'roi', 'trained', 'trafficking', 'purifies', 'endomembranes', 'arrayed', 'oncogenomic', 'endosome', 'knockdowns', 'sgk1', 'nf2', 'ask', 'attribution', 'igfr', 'mitigate', 'gdc', '0941', 'curious', 'endosomes', 'cp', 'byl719', 'eea1', 's13', 's14a', 's14b', 'merit', 'trafficked', 'mirrors', 'begs', 'sbc', 'sbcs', 'puc', 'macroh2a', 'resemblance', 'e3s', 'crls', 'multisubunit', 'exchangeable', 'bosu', 'kipreos', 'cardozo', 'eldridge', 'petroski', 'deshaies', 'willems', 'feldman', 'patton', 'seol', 'skowyra', 'culbins', 'eloc', 'cul2', 'schulman', 'stebbins', 'geyer', 'krek', 'pintard', 'bardwell', 'treisman', 'stogios', 'zollman', 'skplor', 'recognizable', 'propeller', '10th', 'vogel', 'mei', 'severing', 'munoz', 'daxx', 'kwon', 'hib', 'roadkill', 'scfslimb', 'struhl', 'spopmath', 'spopbtb', 'spopn', 'puc1', '390', 'proteolytically', 'puc93', 'macroh2a166', 'plasmon', 'biacore', 'traf6', 'hao', 'sbc1', 'sbc3', 'pucsbc1', 'macroh2asbc', 'cisbc2', 'sandwich', 'traf2', 'position1', 'y87', 'f102', 'y123', 'w131', 'f133', 'encircle', 'v98', 'a170', 'v1362', 'd130', 'k129', 'a170d', 'stt', 'alanines', 'd130a', 'w131a', 'katanin', 'cys94', 'tyr87', 'dow', 'mains', 'zf', 'cul5', 'elongin', 'auc', 'chew', 'monomeric', 'l186d', 'l190d', 'l193d', 'i217k', 'repel', 'solubility', 'copurifies', 'elonginc', 'boxskp2', 'mahrour', 'redefine', '258', 'gigaxonin', 'gig', 'gigbtb', 'klhl12', 'actinfilin', 'explicitly', 'angers', 'salinas', 'btbs', 'btbd12', 'copurify', 'svendsen', 'utilize', 'putting', 'encompassed', 'spopmathx', 'protomers', 'lengthen', 'flexibly', 'thrombin', 's14', 'scattering', 'saxs', 'compactness', 's15', 'gig1', 'architectures', 'selenomethionines', 'orientational', 'velocity', 'm2m3', 'pep', 'acetyl', 'pep2', 'scfbw7', 'scfcdc4', 'tang', 'polekhina', 'yin', 's16', 'cul3spop', 'cul4', 'scrimala', 'cdc4', 'orlicky', 's17', 'nedd8', 'nedd8ylated', 'duda', 'enzymologicals', 'saha', 'yamoah', 'fbw7', 'degron', 'welcker', 'clurman', 'elongating', 'cul3ntd', 'pgex4t1', 'prsfduet', 'cryoprotectants', 'als', 'nsls', 'pulldowns', 'dichroism', 'pipks', 'pipkis', 'pipkiis', 'pi5p', 'pipk', 'kis', 'pipkii', 'speckles', 'theme', 'speckle', 'sp600125', 'me2so', 'philip', 'majerus', 'kii', 'asp287', 'met365', 'thio', 'histrap', 'gstrap', 'ff', 'fluorophore', '1024', 'keck', 'viscosity', 'nickel', 'illustrations', 'ubiquitylate', 'ubiquitylated', 's326a', 'ser326', 'd278a', 'reinforcing', 'ubiquitylates', 'siah2', 'phd3', 'autocatalytic', 'stream', 'invaluable', 'transcriptomic', 'steroidogenic', 'intratumoral', 'taxanes', 'gli', 'spops', 'footprint', 'enriches', 'mdv3100', 'medivation', 'baylor', 'authentication', 'str', 'roswell', 'rppmi1640', 'lapc4', 'zoran', 'culig', 'innsbruck', 'zfn', 'malley', 'spopwt', 'y87n', 'f102c', 's119n', 'w131g', 'f1331', 'f133v', 'msigdb', 'nelson', 're

spose', 'suppl', 'y87c', 'f125v', 'rtqpcr', 'camkk2', 'abcc4', 'homer2', 'sep p1', 'fkbp5', 'arn', 'glycosides', 'nes', 'rbx', 'ubiquinated', 'dn', 'asst t', 'pest', 'pais', 'a646d', 's648n', 'cais', 'mais', 's647f', 't650a', 'stt6 48', '650aaa', 'arv7', 'contradicted', 'xenografted', 'beige', 'fed', 'prosta tes', 'androgenic', 'pcs', 'and33', 'footprints', 'extinguish', 'hint', 'anti androgens', 'elucidates', 'antiandrogen', '261', 'ferlay', 'recurs', 'accentuated', 'sinha', 'intracrine', 'tindall', 'grossmann', 'ubiq', 'uitination', 'nucleates', 'crl3', 'hern', 'ndez', 'oz', 'ubiqui', 'tin', 'kan', 'zhuang', '645asstt649', '203egsss207', 'c4', 'spopisa', 'pcmv5', 'polyubi', 'quitinate d', '1j', '1k', 'overex', 'pression', '1m', '1n', 'zegarra', 'coimmunoprecipi tation', 'coip', '565', '676', 'because645asstt649', 'egsss', 'totally', 's64 6a', 's647a', 't648a', 't649a', 'cotrans', 'fected', 's648a', 'ser647', 'thr6 48', 'v567es', 'representatives', 'krongrad', 'stood', 'mibolerone', 'appro x', 'imately', 'sumoylation', 'steen', 'rnf6', 'qi', 'dez', 'casein', 'rnber g', 'watson', 'escaping', 'escapes', 'worthwhile', 'bohrer', 'kod', 'toyobo', 'p160', 'srcs', 'ncoa', 'aib1', 'ncoa3', 'poxvirus', 'tif2', 'grip1', 'actr', 'impedes', 'posttranslationally', 'spopf133v', 'kallikrein', 'diphenyltetrazo lium', 'mtspop', 's102', 's101', 's505', 's509', 'acknowledged', 'readouts', 'epithelioma', 'msse', 'ferguson', 'fsd', 'autosomaldominant', 'regressing', 'scars1', 'marfan', 'scottish', 'omim132800', '9q22', 'candidates6', 'elude d', 'reconsi', 'dered', 'inclusive', 'chr9', 'alk5', 'proposed3', 'tle1', 'ab ca1', 'cosequenced', 'gaii', 'flowcell', 'contigs', '956', 'msse2', 'uncoveri ng', 'msse6', 'goudie1', 'mariella', 'alessandro2', 'barry', 'merriman3', 'ha ne', 'lee3', 'ildik', 'szever', 'nyi4', 'stuart', 'avery4', 'brian', 'connor 3', 'stanley', 'nelson3', 'stephanie', 'coats1', 'arlene', 'stewart1', 'lesle y', 'christie5', 'gabriella', 'pichert6', 'jean', 'friedel7', 'ian', 'hayes 8', 'nigel', 'burrows9', 'sean', 'whittaker10', 'marie', 'gerdes11', 'sigur d', 'broesby', 'olsen13', 'malcolm', 'smith14', 'chandra', 'verma15', 'decla n', 'lunny4', 'bruno', 'reversade4', 'birgitte', 'lane2', '1human', 'dentistr y', 'nursing', '2cancer', 'geffen', '4institute', 'singapore', '5division', '6clinical', 'nhs', '7unit', 'dermatologie', 'hospitalier', 'morey', 'chalo n', '8northern', 'auckland', 'zealand', '9department', 'addenbrookes', '10s t', '11department', 'rigshospitalet', 'copenhagen', '12department', '13depart ment', '14resource', 'veterinary', '15bioinformatics', '16these', 'birgit', 'imb', '366', 'undertook', 'lod', 'cys41tyr', 'asn45ser', 'gly52arg', 'pro83 leu', 'arg80x', 'asn45lyfsx30', '122g', 'scotland', '134a', '135insa', '154 g', '238c', '248c', '725g', 'trp242x', '980delc', 'pro327glnfsx8', '1059', '1 062del', 'actginscaataa', 'leu354asnfsx4', '1240c', 'arg414x', 'totals', 'nor theast', 'aug', '503', 'gly174val', 'thr200ile', 'glu218lys', 'glu228del', '1 ys232glu', 'phe234leu', 'ser241leu', 'glu245gly', 'met253ile', 'asn267his', 'met379val', 'gly374glu', 'gly353val', 'asp351gly', 'ala350glu', 'lys335gl n', 'asp400gly', 'asn478ser', 'leu486ser', 'arg487gln', 'arg487pro', 'arg487t rp', 'his315arg', 'gly312ser', 'met318arg', 'ri', 'mutations22', 'serineric h', '367', '0000001', 'immunohistochemically', 'pyrosequencing', 'bose', 'afm 070xb11', 'migration14', 'derestricted', 'risk16', 'stages18', 'keratinocytes 19', 'tgfbrii', 'agedependent', 'apoptosis20', 'mice21', 'selfhealing', 'mf s', 'loey's', 'dietz', 'lds1a', 'lds2a', 'thoracic', 'aortic', 'dissections', 'taad', 'danish', 'papule', 'rolled', 'hyperkeratotic', 'plug', 'pitted', 'scars', 'formalinfixed', 'leucocyte', 'symptomless', 'deletion24', 'duplicatio n25', 'intolerance', 'smad', 'signaling22', 'carcinogenesis27', 'unusualbipa sic', 'assembles', 'aquamarine', 'asn45', 'cys41', 'gly52', 'arg80pro83', '1 d', 'pm', 'sb', '431542', 'sb431542', 'psmad2', '369', 'heal', 'antitumorigen ic', 'cellspecific', 'orchestrate', 'signaling28', 'decoupled', 'tumorpromoti ng', 'g380r', 'achondroplasia', 'dwarfism', 'impose', 'disregulation', 'ach', 'r248c', '370c', 's371c', 'y373c', 'thanatophoric', 'td1', 'severities', 'dou bts', 'donoghue', 'pcdna', 'germantown', 'fgf1', '9007', 'impermeable', 'bs 3', 'sulfosuccinimidyl', 'suberate', 'rinse', 'endoplasmic', 'reticulum', 'qu

ant', 'unglycosylated', 'chondrocytes', 'shorten', 'chondrocyte', 'webster', 'fret', 'bilayers', 'adar', 'magnitudes', 'cited', 'endochondrial', 'ossification', 'ornitz', 'limbs', 'curved', 'thorax', 'ribs', 'bellus', 'acanthosis', 'nigricans', 'saddan', 'hch', 'tdii', 'k650m', 'k650q', 'g370c', 'l8', 'rcj', 'a391e', 'crouzon', 'p577s', 'v658a', 'r841k', 'h845y', 'g853d', 'clin', '6935', 'casting', 'fastest', 'dacarbazine', 'cd140a', 'nccn', 'nct01243346', 'peking', 'beijing', 'leibniz', 'fibroblastoid', 'adherently', 'w531', 'pspax2', 'pmd2', 'envelop', 'lenti', 'concentrator', 'wei', 'pennsylvania', 'phosphosafe', 'stocked', 'cck', 'dojindo', 'csd', 'sixteen', 'v561a', 'k646e', 'h816y', '1839p', 'q639stop', 'q579r', 'nrnasq611', 'breslow', '437', '419', 'non acral', 'pdgfras', 'p573s', 'v654a', 'd842y', 'dissolvent', 'ipilimumab', 'yevroy', 'zelboraf', 'hepatotoxicity', 'commission', 'lasts', 'portland', 'happened', 'a633t', 'a663v', 'phosphotransferase', '658', 'kitv654a', 'benzimidazole', 'd842i', 'revolutionized', 'regrowth', 'radiologically', 'salvage', 'multitargeted', 'd820y', '816', 'y721', '3391', 'y703', '9271', '2211', '2317', '719', 'hamster', 'v560d', 'biosafety', 'gist882', 'gist48', 'gist430', '250k', 'nsp', 'culturing', 'incrementally', 'a829p', 'cs4', 'calcusyn', 'biosoft', 'mathematically', 'd816h', 'd820g', 'dimh842', '845', 'confirmatory', 'n822k', 'sora', 'and22', 'd820a', 'biaryl', 'regorafenib', 'kindler', 'foot', 'bilirubin', 'lierman', 'bubnoff', 'suspend', 'placebo', 'accrued', 'fluorine', 'hypereosinophilia', 'traffic', 'e996k', 'underline', 'keeps', 'ina', 'stabilizingthe', 'mmah', 'proneural', 'remissions', 'glycosylation', 'differentn', 'heavier', 'matured', 'thewild', '80k', '73k', 'pae', 'independant', '90k', 'dampened', 'crosslinked', 'bb', 'wheat', 'agglutinin', 'handed', 'cartoon', 'fcscor', '174k', 'preddimer', 'repopulation', 'ofthe', 'obstruct', 'v649', 'theright', 'westernblot', '49k', 'lb373', 'inlb373', 'phycoerythrin', 'rplp0', 'ag01518', '52k', 'onlygain', 'v664e', 'papillomavirus', 'e5', 'v543', 'v536', 'v539e', 'populate', 'optimally', 'screenings', 'perfectlyconserved', 'csf1r', 'w349', 'predictedn', 'n353lt', 'n359lt', 'v469', 'tetraodon', 'nigrorividis', 'v469a', 'n870', 's1049ch', 'nherf', 'restrictive', 'hypermutation', 'cyto', 'diegem', 'macrogen', 'seoul', 'korea', 'described34', 'py99', 'signalway', 'eet', 'antiseraum34', 'elsewhere40', 'mab1264', 'peprotec', 'brasseur', 'brussels', 'described41', 'hepatocarcinoma', '996', 'cgtgtggactcagacaatgca', 'ggcagaggaatgttagcca', 'restimulated', 'insds', 'thermoscientific', 'grr5', 'pdrivechef1', 'ru5', '216w27', 'fscor', 'fencefsorrigh', 'fence', 'std', 'growthfactor', 'f5', 'ipi', 'imatinibresistant', 'affiliations', 'leuven', 'vib', 'raft', 'fonds', 'voor', 'wetenschappelijk', 'onderzoek', 'vlaanderen', 'debiec', 'rychter', 'maria', 'gasthuisberg', 'herestraat', '347218', '346210', 'kuleuven', '1078', '0432', 'ccr', 'clincancerres', 'amol', 'bay', 'bayer', 'amn107', 'meyers', 'infinity', 'medimmune', 'withimatinib', 'citric', 'pmcv', 'fluos', 'facscanto', 'facsdiva', 'immunoassays', 'tyr754', 'collagenase', 'deoxyuridine', 'dim842', '5751', '642', 'withthe', 'withanti', 'f50', 'weisberg', 'multitarget', 'a10', 'vsmc', 'kitd816v', 'baf', 'kills', 'f70', 'multikinase', 'kdr', 'flt', 'allylamino', 'demethoxy', 'geldanamycin', 'clients', 'chaperoning', 'accumulates', 'bcrabl', 'phosphomak', 'secondgeneration', 'antagonism', 'acted', 'steel', 'pharma', 'abl2', 'tumors3', 'mutations5', 'herndon', '1mg', 'basel', '7020', 'dm1a', 'described11', 'isobologram', 'analysis12', 'hank', 'hbss', 'dissolving', 'nmp', 'peg300', 'bath', 'vehicles', 'ncr', 'outbred', 'described13', '100mg', '50mg', '150mg', 'and1c', 'ed25', 'ed75', 'ed50', 'and55', 'c30', 'bioluminescence', 'disease14', 'infux', 'abcg2', 'literature7', 'mutant7', 'administering', 'weighed', 'cajal', 'pacemakers', 'sti', 'rubin', 'al8', 'incyte', 'osaka', 'affiliated', 'a4502', 'tyne', 'guanidine', 'bluescript', 'ks', '28s', '18s', 'described12', 'subcloning', 'gcaagagatcattggaggcc', 'cgcttcctgatattgagtgg', '1715', 'tcggggagtgatctccgtga', '1830', 'gtaccttctgcccgtgaag', '2954', '2973', 'cactcagtgtggtagagg', '3003', '3022', 'ggaaggaacccctcgaatcc', '3669', '3688', 'xm01186', '20533627', 'zeo', 'hpai', 'asuii', '1620', '2931', 'ec', '42307', 'genzyme'

e', 'techne', 'placenta', 'dupont', 'harvester', 'packard', 'meridian', 'yb5', 'b8', 'autoradiographed', '551', '557', '558', '559', 'arggluphe', '584', 'hyphens', 'reblotted', 'autonomously', 'scantiness', 'irusta', 'dimaio14', '536', 'former1', 'dimaio', 'abl20', 'spdhe566', '571r', 'er561', 'rd841', '842ki', 'di842', '843im', 'n848k', 'y849k', 'hdsn845', '848p', 'oncol', '5357', '5364', 'plateletderived', 'ascopubs', 'firmed', 'pdgframutant', 'cd117', 'kitnegative', 'northwest', 'kaiser', 'ggccctataacttaggccctt', 'tcagctgatgagttgcctga', 'tggttagctcagctggactgat', 'gggatggagagtggaggatt', 'amplimers', 'nondenaturing', 'isolates', 'sitedirected', 'bl2', 'www4', 'oba', 'generously', 'elisabeth', 'buchdunger', 'methoxyl', 'carboxanilide', 'xtt', 'referrals', 'corless', '5358', 'mesentery', 'omentum', 'lasota', 'al9', 'imhd843', '846', 'imh843', '53265', '53270', 'reamplified', 'surveyed', 'rv560', 'rvies560', '564', 'spdghe566', '571k', 'ydsrw582', 'i843', 'mhds844', '845a', '845v', 'dimhd842', '846e', '846s', '846n', '846g', 'imhds843', '847t', 'imhds8843', '847m', 'jco', '5359', 'hirota', 'al7', 'others7', 'spdhe', '566', '5360', '5361', 'rps6ka1', 'prkcq', 'oncogenically', 'swinging', 'assessable', 'outright', 'ap23848', 'mutations38', 'meantime', 'pending', 'al41', 'atpbindings', '5362', 'adjunct', 'hgg', 'higgs', 'brainstem', 'dipg', 'histopathologies', 'dipgs', 'nonbrainstem', 'pdgfb', 'memphis', 'trp53loxp', 'arog', 'guava', 'easycyte', 'cytosoft', 'intracranially', 'anesthetized', 'perfused', 'neuroopathologist', 'ellison', '5241', 'thr37', 'novared', 'ic90', 'pestle', 'mes', 'tyr572', 'ab5443', '742', 'ab5452', 'tyr720', 'ab65258', '12911', '23948', '9275', '4691', 'ser217', '9154', 'ndrg1', 'thr346', '3217', '9408', 'thr246', '2997', '2610', 'ser235', '3752', 'tyr416', '2101', 'tyr527', '2105', '2109', 'tyr705', '9459', '9452', 'a5441', 'na934', 'nxa931', 'egfrviii', 'batches', 'uhc', 'pn', 'opcs', 'oligodendrocyte', 'astroglia', 'gse46190', '2562', 'n468s', 'e10del2', 'oligodendrolioma', 'e229k', 'c235r', 'y288c', 'c290r', 'c450ins', 'a491ins', 'v544ins', 'e7del', 'e10del', 'oval', 'rectangle', '754', 'pmas', 'multityrosine', '868', 'tie', 'tunel', 'deoxyribonucleotidyl', 'parietal', 'oligoastrocytoma', '0012', 'c540ins', 'intracranial', 'circumscribed', 'masses', 'oligodendroglial', 'immunopositive', 'oligodendrocytes', 'opc', 'hc1', 'hc2', 'oligoastrocytomas', 'gliomagenic', 'cerebellar', 'neocortex', 'oligoastrocytic', 'virally', 'radially', 'patterned', 'scleroderma', 'andreae', 'hoch', 'chompret', 'deposit', 'scarring', 'afflicting', 'microvascular', 'circulate', 'autoantibody', 'baroni', 'recreate', 'bonner', 'yamakage', 'jimenez', 'activatable', 'synergizes', 'klinghoffer', 'hubbard', 'meox2', 'epiblast', 'tallquist', 'superactivation', 'relying', 'directional', 'chemotaxis', 'tyr783', 'edema', 'bostr', 'e18', 'sack', 'accomplishing', 'hoyle', 'somite', 'subepidermal', 'rib', 'vertebral', 'skeletons', 'unsegmented', 'sternum', 'unfused', 'sternal', '1953', 'weaning', 'hydronephrosis', 'distention', 'mendelsohn', 'urothelium', 'sma', 'condensed', 'ureters', 'commenced', 'pdgfa', 'pdgfc', 'hyperplasic', '4k', '4o', '4n', 'creer', 'tam', '5j', 'ics', 'pacemaker', 'icc', 'sommer', '5k', 'adhered', 'fibers', 'glomeruli', 'interstitium', '6k', '6l', 'ostman', 'pleiomorphic', 'synergized', 'aged', 'concentric', 'lumenal', 'e11', 'onward', 'santos', 'overgrowths', 'metanephric', 'ureteric', 'bud', 'spectrums', 'pasini', 'ruas', 'peters', 'dai', 'tightening', 'ponten', 'campbell', 'distler', 'abdollahi', 'sutent', 'multinationally', '670', 'prs', '0005', '0356', '0132', 'nonrandom', '0157', 'numerically', '011', 'ay', 'y823d', 'v559d', 'genotypically', 'interlesion', 'antiangiogenic', 'isotypes', 'disclosures', 'asco', 'contributors', 'employment', 'leadership', 'xin', 'darrel', 'cohen', 'baum', 'consultant', 'advisory', 'demetri', 'ownership', 'honoraria', 'robert', 'maki', 'funding', 'testimony', 'remuneration', 'conception', 'wen', 'bin', 'ou', 'jonathan', 'fletcher', 'administrative', 'diana', 'griffith', 'provision', 'cristina', 'antonescu', 'amy', 'harlow', 'ajia', 'town', 'arin', 'mckinley', 'writing', 'duensing', '708', 'mcgreevey', '1567', '1572', 'rader', 'ae', 'avery', 'wait', 'choy', 'ys', 'sl', 'blanke', '4764', '4774', 'therasse', 'arbuck', 'eisenhauer', 'ea', 'n

atl', 'inst', '4342', '4349', '4375', 'ohsu', 'knight', '3710', 'veterans', '97239', '3405', '5158', '0625', 'biron', 'bioavailable', 'imatinib sensitiv e', 'discoverx', 'kinomescan', 'elect', 'antipdgfra', '4376', 'kochi', 'eol', 'amplifi', '4q12', 'collagenasedisaggregated', '37c', 'cfs1r', '4377', 'cac y', '561rerv', '848r', '319', '006a', '001a', '02a', '003a', 'hdsn8455', '05 a', '03a', '002a', '4378', '4379', 'hypereosinophillic', 't674', '4380', 'd84 2vmutant', '510', '4381', 'creonolanib', 'hematuria', 'f1p111', 'bcrpdgfra', 'dermatofibrosarcoma', 'protuberans', 'col1a1', 'pericytes', 'antiphospho', '4382', 'nct01229644', 'nct01393912', 'fipl1', 'benzoyl', 'cdk1', 'reassesse d', 'computerized', 'positron', 'consenting', 'spectrumorange', 'cep4', 'wetz lar', 'photometrics', 'tuscon', 'quips', 'smartcapture', 'crystalline', 'rese quenced', 'aqueousone', 'fifteen', 'discontinued', 'rectum', 'awd', 'dod', 'd uodenum', 'immunophenotype', 'diploidy', 'mutationsa', 'k558n', 'wk557', '56 3', 'v559g', 'gcg', 'gat', 'aca', 'ata', 'd820e', 'tat', 'gac', 'ggc', 'immun opositivity', 'kity703', 'equivalency', 'k642e', 'cautions', 'underlining', 'basophils', 'cases7', 'described28', 'imidazol', 'trifluoromethyl', 'pyridi nyl', 'celltiter96', 'charlottesville', '508', '783', 'transplantations', 'ti tered', 'hanks', 'pyrrolidone', 'humanely', 'heparinized', 'smears', 'straine r', 'rehydrated', 'allophycocyanin', 'mac', '7aad', 'actinomycin', 'dosed', 'killing', 'wbcs', '802', 'effaced', 'splenic', 'pdfgr', 'corroboration', 'm yeloproliferation', '731', 'corroborated', 'splenocytes', 'g719x', 'l861q', 'sensitizing', 's768i', 'e709a', '1747v', 'r776h', 'a871g', 'e709v', 'dell74 7', 't51', 'r776s', 'lrea', 'rrs', 'survivals', 'pfss', 'coexists', 'inventor y', 'mycancergenome', 'oncologists', '1st', 'beth', 'israel', 'deaconess', 'b idmc2009', '000182', 'bidmc', 'radiographic', 'redcap', 'e709x', 's7681', 'ec og', 'taiwanese', 'pna', 'lna', 'consorted', 'g719a', 'g719s', 'neratinib', 'g719c', 'trough', 'ensuing', '1747s', 'troughs', 'd761y', 'e709g', 'egfrs', 'assertion', 'h870r', '1861', 'g719', 'macrodissected', 'tggttaacatccacccagatc a', 'aaaagggtgggcctgaggttca', 'tomographic', 'reimaging', 'delineates', 'radio logist', 'summing', 'pdnr', 'a750del', 'y764', 'v765inshh', 'afatinib', 'bibw 2992', 'graphically', 'py1068', '026', '347', 'african', 'adenosquamous', '74 4', '745', 'e746x', '1747p', 'vaikel', 'leu747', '1788', '1777', 'm766', '186 2', '1858', 'f723', 'disfavor', 'glu746', 'i744', 'k745inskipvai', 'k745', 'e 746instpvaik', 'palliative', 'xl647', 'rebiopsy', 'volumetric', '747', 'tcccg tcgctat', 'reviewing', 'tent', 'sizing', 'dxs', 'map2k4', 'posited', 'multiplicity', 'peroxisomal', 'proliferator', 'ppar', 'hematogenously', 'colonize', 'conditionally', 'subjection', 'hif2', 'tp53r172h', 'eicosanoids', 'mekk1', 'sorbitol', 'clasto', 'lactacytin', 'lactone', 'sb202190', 't0070907', 'cay man', 'jia', 'anderson', 'plhcx', 'zhimin', 'edward', 'yeh', 'jnk2', 'apf', 'bing', 'pparg2', '8895', 'bgn', 'vegfc', 'satb1', 'pparg', 'tabular', 'deno ting', 'gp', '393p', 'nonmetastatic', '344sq', 'bxpc3', 'h2009', 'layered', 'nitrotetrazolium', 'hemocytometer', 'countess', 'okadaic', 'jnk1', 'k55m', 'carna', 'biomedicals', 'tran35s', 'husbandry', 'mkk41', 'scgb1a1cre', 'inte rbred', 'shz', 'ribopure', 'asuragen', 'ggattacaaggatgacgacga', 'ctgccattattt gcccactt', 'l32', 'cole', 'farmer', 'gex', 'sonicator', 'pgl3', 'ppre', 'mang elsdorf', 'n234i', 's251n', 'r134q', 'q142l', 'p326l', 'v321m', 'r154w', 'a27 9t', 'canpredict', '3alo', 'w310', 'r304', 'i295fs', 'pt67', 'antitubulin', 'scgb1a1', '0186', 'demise', 'mined', 'braf1', 'underpowered', 'krasla1', 's hy', 'kr', 'r273l', 'vec', '070', 'immunocompetent', 'nonreconstituted', '02 7', '034', 'metastasisa', 'nonspliced', 'posttranscriptional', 'pgc', 'gse703 5', '343', 'rosiglitazone', 'biglycan', 'shp1', 'shp4', '8i', 'shp3', '8j', 'mitigates', 'misfolding', 'degrading', 'mkks', 'rxr', 'tyr249', 'vcam1', 'lymphangiogenesis', 'il13ra1', 'csf1', 'ccr9', 'cxcl11', 'cxcl7', 'smarca2', 'id2', 'inadl', 'crb3', 'pard6b', 'ncam1', 'snail1', '1020', 'commentary', '953', 't877a', 'demonstrable', 'declines', 'agonistic', 'approximating', 'l 701h', '876', 't877', 'coregulators', 'h12', 'conveys', 'castrated', 'ctdna', 'circulating', 'emulsion', 'beaming', 'magnetics', 'pharmacokinetics', 'pharm

'acodynamics', 'maximize', 'equivalents', 'mixing', 'priori', 'undesirable', 'bivariate', 'genital', 'penis', 'sexes', 'cigarette', 'yao', 'hpv', 'vagina', 'canal', 'beckmann', 'uteri', 'bornstein', 'editors', '1978', 'brazil', 'county', 'urban', '100', 'approximations', 'latin', 'pacific', 'columbia', 'nova', 'ontario', 'connecticut', 'iowa', 'detroit', 'mexico', 'miyagi', 'af rica', 'hungary', 'ireland', 'spain', 'zaragoza', 'geneva', 'brunswick', 'pri nce', 'newfoundland', 'alameda', 'tianjin', 'iceland', 'poland', 'trent', 'yu goslavia', 'nsw', 'victoria', 'spo', 'paulo', 'costa', 'korean', 'hispanic', 'india', 'hiroshima', 'indian', 'bas', 'republic', 'norway', 'bo', 'bronchus', 'trachea', 'o5', 'i1', 'caner', 'footnote', 'ui', 'uneven', 'industrialized', 'sound', 'country', 'heading', 'epidemiological', 'boice', 'say', 'brack en', 'sexually', '1812', 'justified', 'smoke', 'sexual', 'etiologies', 'scalable', 'pot1', 'd33e', 'phenotyping', 'collisson', '681', 'unstudied', 'previo usly implicated', '381', 'horfeome', '696', 'orfs', 'generalizable', 'hale', 'mek1dd', 'timeframe', 'compositions', 'krasd33e', 'nfe2l2g31r', 'pot1g76v', 'pik3cbe497d', 'preinjection', 'neighbors', 'barcode', 's4n', 'krasa59g', 'akt1152r', 'akt1q79k', 'nfe2l2wt', 'fam200as481n', 'akt1e17k', 'differentiates', 'mekdd', '978', 'l1000', 'luminex', '036', 'nrasq61h', 'y64d', 'idh2r172k', 'f90s', 'r233q', 'k6n', 'g127r', 'g127v', 'spopf102c', 'f133s', 'k134n', 'w131c', 'e50k', 'k101i', 'e47a', 'r658q', 'i347m', 's462y', 'r689q', 'r505c', 'r465c', 'r465h', 'akte17k', 'prosurvival', 'e62k', 'idh2r172m', 'idh1r132c', 'idh1r132s', 'idh1r132h', 'idh1r132l', 'e190k', 'p33s', 'idh2', 'g137e', 'e268d', 'a416v', 'a47v', 't331m', 'i138f', 'k134', 'e47', 'e50', 'nuanced', 'krase62k', 'pik3cba1048v', 'nfe2l2g31a', 'nfe2l2g31v', 'nfe2l2n160s', 'akt1e267g', 'akt1r370c', 'g31r', 'g31v', 'g31a', 'n160s', 'e497d', 'g76v', '4epv', 'a59g', 'latencies', 'angiosarcoma', 'y89c', 'q94e', 'y223c', 's270n', 'telomeres', 'complements', 'chasm', 'vest', 'spopk134n', 'presumable', 'arbitrarily', 'gathered', 'pik3c2g', 'pik3cg', 'pik3c2b', 'pik3c2a', 'pik3r4', 'bcl2', 'bcl3', 'bcl6', 'bcl9', 'issx', 'phylop', 'linearly', 'b2m', 'controlpriority', 'bonus', 'choose', 'pdonr223', 'discontinuity', 'destination', 'trc', 'queries', 'supplementation', 'pooling', '0089', 'prepped', 'interscapular', 'finely', 'minced', 'demultiplexing', 'demarcated', 'aatgatacggcggaccaccgagatct', 'acactttccctacacgacgtttccgatct', 'tcttggtggaaaggacga', 'caagcagaaggacggcatacgagat', 'nnnnnnnn', 'gtgactggagttcagacgtgtctttccgatct', 'taaagcagcgtatccacatagcgt', 'ampure', 'demultiplexed', 'scripts', 'perturbagens', 'tcl', 'annealed', 'sape', 'flexmap', 'perturbagen', 'lincs', 'bfp', 'hcred', 'nfkb1a', 'dnmt3a', 'delta8', 'lt1', 'schr', 'dinger', '12063', 'cst', '3965', '4060', 't9026', '12721', 'af3925', 'irdye', 'rocked', 'elute', 'purify', 'vilo', 'fwd', 'cacatccagttagaaaccagtgg', 'ggaatgtctgcgccaaaagctg', 'actb', 'caccattgcatacgccgttc', 'aggctttcggtatgtccacgt', 'humdiv', 'venn', 'venny', '785', 'expectancy', 'uncontrollably', 'shrink', 'indefinitely', 'antiapoptotic', 'cragg', '1371', 'pmed', 'english', 'spanish', 'emory', 'languages', 'anyone', 'edit', 'd761', 'supplier', 'darmstadt', 'transbronchial', 'gca', 'agt', 'cca', 'cgt', 'tct', 'pgem', 'aac', 'ctc', 'cct', 'legends', 'lexington', 'ky', 'ptyr1068', 'stressgen', 'dioc6', 'eugene', 'tko', 'bims', 'biml', 'bime1', 'hypophosphorylated', 'hypophosphorylation', 's752', 'bronchioalveolar', 'thickened', 'pleura', 'lre', 'overcame', 'intermediately', 'sudden', 'vad', 'fmk', 'bak', 'intermembrane', 'glucocorticoid', 'postulates', 'radiological', 'and6a', 'balak', 'transfектант', 't854a', 'bibw', '00299804', 'bevacizumab', 'anilinoquinazoline', 'renamed', 'prespecified', 'seldom', 'postprogression', 'halt', 't854', 'cancers62', 'couples', 'akin', 'deterioration', 'ask s', 'qualified', '797', 'arena', 'nct00525148', 'iib', 'lux', 'xl184', 'arq', 'figitumumab', '751', '871', 'gp130', 'hsp', 'abt', '737', 'pace', 'lags', 'visceral', 'multinucleotide', 'exploratory', 'procurement', 'effusions', 'tracings', '530', 'em', 'ascent', 'ethanols', 'spotlight', 'coverslipped', 'heat ed', 'enumeration', '573', '2573t', 'landmarks', 'e749', 'a750p', 'lll', 'nonneoplastic', 'aniline', 'newer', 'cys797', '762', 'e762', 'anchor', 'leptomene

nges', '732', 'cerebrospinal', 'experiencing', 'intron1', 'ssr1', '216g', '191c', 'exon19', '0036', 'maximized', 'minimizing', 'nonmalignant', 'antiapoptosis', 'supportive', 'exon21', 'exon20', 'exon18', 'contentious', 'cappuzzo', 'her3', 'takano', 'dziadziuszko', 'interchangeably', 'ethniciities', 'ethnicity', 'dubey', 'sanyo', 'resections', 'unacceptable', 'refusal', 'abbreviations', 'eastern', 'tblb', '29egfr', 'pbluescript', 'rv6', 'ssr', 'ctcctcaaaacccggagactt', 'cacttgaaccaggagcac', 'gagctagacgtccggcag', 'ctgtcgtccggctgg', 'to158', 'dnntp', 'bufferi', 'shiga', 'pop4', 'genemapper', 'ver', 'cep7', 'strata', 'mst', 'statuses', '31egfr', 'adsq', '1response', '2response', '3death', '2235', '2249', '2236', '2250', 'a750insq', '2245', 's752insi', '2237', '2255', 'a750insp', '2239', '2248', '2256', 'i759', '2254', '2276', '2573', '2155', '2156', 'g721d', '2162', 'insgvv', '2308', '2309', 'gggtcggt', '1777g', '2330', '2369', '2281', 'g721s', '2161', 'electropherograms', '051', '0024', '063', '733', '0075', '020', '0091', 'mpfs', 'survial', 'kuraya', 'interethnical', 'infrequency', 'burger', 'careful', 'hoang', 'chemonaive', 'hoangs', 'epsilon', 'exo', 'tcg', 'pold1', 'kunkel', 'erie', 'reha', 'krantz', 'lob', 'kane', 'vilar', 'gruber', 'chromophobe', 'ciriello', 'polymerases', 'palles', 'shcherbakova', 'meng', 'valle', 'zou', 'alexandrov', 'synthesizes', 'pursell', 'miyabe', 'lagging', 'mcelhinny', '3719', 'ultrahigh', 'p286r', 'v4111', 's297f', 'f367s', 'a456p', 's459f', 'eec', 'ultramutated', 'exor311h', 'exop286r', 'exod316n', 'd316', 'exod316', 'jin', 'shevelev', 'bscher', 'd275', 'e277', 'e318', 's2ab', 'korona', 'goksenin', 'd275a', 'e277a', 'p286', 'd368', 'a428t', 'misincorporated', 'fortune', 'schmitt', 'bebene', 's459', 'e275', 'mispairing', 'mispairings', 'exonucleolytic', 'schaaper', 'dunn', 'lujan', 'saturate', 'misinsertion', 'undamaged', 'deamination', 'misincorporation', 'barnes', 'lindahl', 'm13mp2', 'ssdna', 'tallied', 'komolgrov', '0019', '3555', '6814', 'a00n', 'ei', '6917', '6718', 'biases', 'execute', 'oris', 'lamb2', 'giacca', 'abdurashidova', 'vaara', 'predominates', '001635', 'keller', 'dhfr', 'looney', 'hamlin', 'altman', 'fanning', 'mechali', 'harder', 'marking', 'skews', 'agag', 'agaa', 'r213x', 'ttga', 'dominated', 'demonastrably', 'stretching', 'lmnb2', 'eukaryote', 'fire', 'synchronized', 'valenzuela', 'aparicio', 'firing', 'delving', 'deoxyribonucleotides', 'tetramer', 'accessory', 'pole2', 'pole3', 'pole4', 'ppp2r1a', '251', 'endometrioid', 'ovaries', 'tenets', 'teratoma', 'omega', 'tek', 'dalian', 'thermal', 'cycler', '9700', 'oryctolagus', 'cuniculus', 'tropicalis', 'leishmania', 'trypanosoma', 'brucei', 'plasmodium', 'falciparum', 'speculations', 'leiomyoma', 'pertuzumab', 'dm1', 'resultsidentification', 'amplificationwe', 'conducting', 'surgeons', 'z1031', '15687', '16757', 'c8', 'a135', 'g309a', 'd769h', 'v7771', 'ecd', 'powerpointfigure', 's760a', 'varscan', '4502a', '507', 'a8', 'a06z', 'mutationswe', 'multisequence', 'stars', 'p780ins', '3pp0', 'v777', 'v769', '2gs2', 'syr127063', '1n8y', '1ivo', 'v7691', 'adoptin', 'abuts', 'f856', 'f1271', 'y835f', 'v842i', 'r896c', 'ecds', 'g309', 's310', 'm253', 'g309e', 'aale', 'airway', 'signalingin', 'recombinantly', 'chemidoc', 'xrs', 'r678q', 'canertinib', 'agarto', 'spiculated', 'modelsthe', 'greulich', 'logarithm', 'lapatinibgiven', 'popuptable', 'lapatinibcharacterization', 'mutationsin', 'd769y', 'p780', 'y781insertionsp', 'i767m', 'd769', 'y772', 'i919', 'm923', '3kex', 'discussionin', 'cooccurred', 'l755w', 'a0t6', 'g1201v', '499', 'launched', 'methodscell', 'linesmcf10a', 'sarawati', 'sukumar', 'receipt', 'reauthenticated', 'hydrocortisone', 'amaral', 'systat', 'inhibitorsantibodies', 'y1289', 'tyr1173', 'py1248', 'jewish', 'pharmacy', 'medchem', 'dataleve1', '515', 'rnatalot', 'phusion', 'cagaagatccgaaagtacac', 'atacaccagttcagcaggtc', 'visualizationher2', 'agilent', 'radiometric', 'liposomes', 'linesher2', 'pcfg5', 'zeocin', 'nx', 'amphotrophic', 'biolegend', '24d2', 'cytek', 'matrigeloverlay', 'assaysix', 'bacto', 'solidification', '46r', 'stack', 'bit', 'rolling', 'watershed', 'pasted', 'unrestricted', 'circularity', 'mcmcglm', 'intercept', 'markov', 'monte', 'carlo', 'burn', 'autocorrelation', 'variances', 'covariances', 'overdispersion', 'wishart', 'implicitly', 'studiesexp

onentially', 'intercepts', 'slopes', 'lmer', 'lme4', 'specifications', 'assumptions', 'erbb1', 'her4', 'ectodomain', 'sliwkowski', 'graus', 'porta', 'pinkas', 'kramarski', 'worthylake', 'sapks', 'pak', 'jnnkk', 'slamon', 'igg1', 'hirsch', 'shigematsu', 'narrowing', 'gazdar', '776', 'her2yvma', 'beas2b', 'her2wt', 'vc', 'her2vc', 'downs', 'wounded', 'cellsa', 'confluent', 'underside', 'debnath', 'brugge', 'spheroids', 'reinitiates', 'plateau', 'transwells', 'homodimerize', 'ap1510', 'muthuswamy', 'p75', 'her2wta', 'digiovanna', 'stern', 'y1068', 'y416', 'fainter', 'precipitates', 'egfrk721r', 'egfrkd', 'y845', 'stover', 'biscardi', 'k721r', 'ewald', 'hernan', 'moasser', 'moulder', 'her2yvmaa', 'replenished', 'heregulin', 'agus', 'rusnak', 'h1781', 'her2mut', 'trypsinization', 'cys773', 'citri', 'smaill', 'submicromolar', 'steeper', 'signal inga', 'coreceptors', 'siegel', 'glockner', 'endows', 'alimandi', 'thor', 'multimerize', 'samanta', 'transgenes', 'bigenic', 'muller', 'lenferink', 'wave d', 'gillgrass', 'ueda', 'sibia', 'homodimerized', 'reinitiate', 'mcf10', 'wakeling', 'tona', 'gilmer', 'fry', 'parke', 'brent', 'polk', 'calibur', 'polycarbonate', 'wiped', 'diff', 'quik', 'dade', 'behring', 'poured', 'solidify', 'exponentially', 'dawley', 'palpation', 'biocoat', 'lsm510', 'ifa', 'axiophot', 'biotinylation', 'sulfo', 'aliquoted', 'impacting', '623', 'epha3', 'ntrk2', 'k645e', 'a775', 'g776insyvma', 'insyvma', 's310y', '5637', 'g12f', 'd845a', 'c311r', 'e321g', 'c334s', '000002', 'unamplified', 'subdomain', 'c311', 'c334', 'c299', 'c338', '2a91', 'nonreducing', 'e321', 'k369', 'r434', '1221', '1222', 'mig6', 'errfi1', 'crk', 'ccd88a', 'iqgap', 'peak1', 'ptpn12', 'an3ca', 'frameshifting', 'refute', 'nonexpressed', '310f', '310y', 'y274', 'f279', 'noncovalent', 'precedent', 'nitrosoethylurea', 'q34', 'breakthrough', 'oosterom', 'upperley', 'magnusson', 'coronary', 'polymorphonuclear', 'metamyelocytes', 'myelocytes', 'haemoglobin', 'suddenly', 'night', 'sweats', 'forever', 'ecchymosis', 'lymphadenopathies', 'aspire', 'cytochemistry', 'ph1', 'daunorubicin', 'cyclophosphamide', 'asparaginase', 'intrathecal', 'methotrexate', 'lala', 'intensification', 'mitoxantrone', 'monocytosis', 'aplasia', 'promyelocytes', 'prophylaxis', 'leucocytosis', 'meningitis', 'p23', 'reinitiated', 'meningeal', 'btl', 'chic2', 'qbiogene', 'illkirch', 'biosystem', 'hopes', 'penetration', 'juxtaposing', 'vardiman', 'juxtaposes', 'carroll', 'pardanani', '5q33', '8p11', 'golub', 'dlbl', 'vincristine', 'adriamycin', 'prednisone', 'pruritus', 'retroperitoneal', 'lymphoblasts', 'hypercellularity', 'harvest', 'nuc', 'ish', 'ideograms', 'qiamp', 'durham', '743', '925', 'interrupted', '1842', 'flt1', 'flk2', 'welsh', 'intracellularly', 'uncoupling', 'signify', 'langerhans', 'histiocytosis', 'lch', 'c121s', 'g128d', '61qkqkvg', 'map3k1', 't779fs', 't1481fs', 'e1286v', 'lcs', 'bechan', 'egeler', 'badalian', 'sahm', 'kansal', 'berres', 'implied', 'willman', 'handled', 'accorded', 'federal', 'pseudonymized', 'truextrac', 'covaris', 'oncopanel', 'ddk', 'rc218460', 'rc214358', '002755', 'pro89', 'reinserted', 'asisi', 'mlui', 'huntsville', 'bspei', 'l1481fs', 'ps100001', 'igepal', 'f3165', 'pv3314', 'precast', 'polyvinylidene', '9121s', '4370s', 'southernbiotech', 'hyblob', 'denville', 'plainfield', 's2673', 'unacceptably', 'cd1a', 'q347', 'a348del', 'f3511', 'samplesa', 't799fs', 'g61del', 'chr15', '66727450cagaaggcagaagg', 'tgg', '66729153t', '66729154g', '66729175g', 'greger', 'emery', 'chr5', '56177067tg', '56189410ta', '56180528a', 'precipitations', 'gathering', 'hairy', 'chakraborty', 'g128v', 'yan', 'siow', 'map2k7', 'lange', 'widmann', 'haroche', 'sound', 'exhaustively', 'genomewide', 'iiib', 'iiic', 'apert', 'craniosynostosis', 'syndactyly', 'feet', 'nonresectable', 'nonhypermutated', 's252w', 'p253r', 'n549', 'k659', 'tangentially', 'd626a', 'frs2', 'mfe', 'shfgfr2', 'mfe280', 'mfe296', 'pfeiffer', 'diminishes', 's267p', 'ste', 'choloroform', 'norwegian', 'inspectorate', 'normals', 'repli', 'gtm', 'hotstar', 'm13', 'snpcompare', 'polyphred', '02b', 'postprocessing', 'polydhan', 'richter', 'notched', 'averaging', 'aligning', 'thermocycling', '30s', 'unincorporated', 'deactivated', 'shrimp', 'ddntp', 'thermosequenase', 'cycled', 'hydroxypicolinic', 'spectrochip', 'spectrochips', 'bruker', 'biflex', 'spectroreader', 'fgf7', 'fg

fr2b', 'ioh36372', 'gacgtaccttatctataagaggaccgac', 'attcgtggtcggccttata',
 'k310r', 'tgccctacctcagggttctcaaggc', 'gccttgagaaccctgaggtagggca', 'a314d',
 'ggttctcaaggacccgggttaac', 'gttaaca', 'gtccttgagaacc', 'a389t', 'tcttaatcg
 cctatatgggttaacag', 'ctgttaccaccatataaggcgattaaga', 'n549k', 'cataaatcttctgg
 agcctgcacac', 'tgcaggctccagaagatttatg', 'c382r', 'gatagccatttaccgcataggggt
 c', 'gaccctatgcggtaatggctatc', 'd101y', 'cacgcctagatactccggcctc', 'gaggccgg
 agtatctaggcgtggcg', 'tggagcgtcgctcaccggccatc', 'gatggg', 'ccggtgacgcgatcgc
 tcca', 'orbigen', 'trcn0000000367', 'gccaccaaccaaataccaaatctc', 'trcn00000003
 68', 'ccgaatgaagaacacgaccaa', 'trcn0000000369', 'cccaacaataggacagtgcct', 'trc
 n0000000370', 'gccaaccttcgaacagtattc', 'gcaagctgaccctgaagttcat', '1615', 'sh
 gfp', 'readings', 'hec1b', 'ec50s', 'ipegal', '6930', 'y436', '3861', 'y196',
 '3864', '17841', '1979', 'risen', 'unrealistic', 'chang', 'kress', 'crawfor
 d', 'linzer', 'rotter', 'e1b', 'barr', 'hepatitis', 'erence', 'lived', 'ore
 n', 'caron', '1980s', 'grounds', 'mora', 'milner', 'reich', 'ndings', 'ganno
 n', 'maltzman', 'hsp70', 'nity', 'hinds', 'jenkins', 'broblasts', 'eliyahu',
 'stockton', '0950', '9232', 'onc', 'nding', 'friend', 'erythroleukaemia', 'c
 how', 'ben', 'authentic', 'tetramers', 'pab240', 'dittmer', 'ected', 'genit
 o', 'malkin', 'speci', 'homozygously', 'donehower', 'jacks', 'erentiation',
 'mdm', 'ampli', 'momand', 'oliner', 'moll', 'erentiated', 'nb', 'compromise
 s', 'una', 'immunochemical', 'pab1620', 'topologically', 'transactivator', 'j
 ang', 'herpes', 'kern', 'deiry', 'gpypypy', 'decamers', 'funk', 'tata', 'tfii
 h', '1995a', 'midgley', 'ginsberg', 'lechner', 'transrepression', 'tbp', 'tfi
 id', 'pavletich', 'pcaf', 'collins', 'scolnick', 'xpb', 'xpd', 'pab', 'uenc
 e', 'modi', 'erent', '249', 'hsu', 'bressac', 'c4a', 'bennett', '17p13', 'iso
 be', 'miller', '234', 'rm', 'anatomy', 'nterminal', '318', 'rp', 'dutta', 'ha
 upt', 'kubbutat', 'mack', 'murphy', 'walker', 'pxxp', 'chief', 'signi', 'hete
 rozygotes', 'bounded', 'acetylases', 'lsh', 'tetrahedrally', '7625', 'subili
 sin', '53bp2', 'gorina', 'antiparallel', '242', 'chelating', '1993a', 'ects',
 'arg273', 'je', 'rey', 'symmetrical', 'glu326', 'arg333', 'gly334', 'arg335',
 'gly356', 'nls', 'nls1', 'nls2', 'dang', '1990b', '363', 'pab421', 'hupp', 's
 elivanova', 'halazonetis', 'locks', 'sakaguchi', 'okorokov', 'helicases', '19
 95b', 'csb', 'diller', 'suppose', 'irreversibility', 'persists', 'telangectas
 ia', 'ser15', 'banin', 'canman', 'zhan', 'p217', 'brugarolas', 'waldman', 'sh
 err', 'arrests', 'p107', 'p130', 'okamoto', 'beach', 'zauberman', 'autoregula
 tory', 'broblast', 'agarwal', 'stewart', 'hermeking', 'rpa', 'unwinding', 'th
 ymocytes', 'clarke', 'strasser', 'downshift', 'owen', 'miyashita', 'baserga',
 'nger', 'israeli', 'pig', 'polyak', 'caelles', 'deciding', 'wt1', 'ecting',
 'cancel', 'morgenbesser', 'stott', 'restrains', 'bates', 'shuttling', 'expor
 t', 'tao', 'cak', 'mayr', 'erential', 'ck2', 'erentially', 'milne', 'ser392',
 'transferases', 'hats', 'lill', 'hat', 'dnabinding', 'acetylates', 'lys320',
 'jayaraman', 'erences', '1990a', 'tremendous', 'import', 'myb', 'erb', 'yc',
 'stratton', 'hubert', 'rogers', '725', 'wolfe', 'junior', 'faculty', 'award',
 '328', '364', 'framingham', 'nac1', 'rhodamine', 'conform', '540', 'nucleopla
 smic', 'polyoma', 'nucleoplasmin', 're1', 'k3', 'photomicrograph', 'htlv', 'p
 x', 'b23', 'sis', 'redistributes', 'rrna', 'neuro', 'nerves', 'schwannoma',
 'ganglioneuroma', 'fascicular', 'myxoid', 'storiform', 'rhabdoid', 'chondroi
 d', 'lympho', 'knockouts', 'legius', 'latres', 'kindblom', 'nakanishi', 'stra
 uss', 'sonobe', 'rasmussen', 'watanabe', 'leroy', 'shin', 'chao', 'scheithae
 r', '411', 'neurogenic', 'periodic', 'schiff', 'liposarcoma', 'indistinct',
 'fascicles', 'tains', 'neomarkers', 'tgt', 'tg', 'ttt', 'tac', 'trunk', 'le
 g', 'd281a', 'shoulder', 'elbow', 'groin', 'c176f', 'n239s', 'mediastinum',
 's241f', 'sinus', 'mandible', 'parotid', 'nose', 'orbit', 'cheek', 'dup', 'm
 axillary', 'retroperitoneum', 'vestibular', 'condensation', 'x200', 'counters
 tain', 'coexisted', 'v272m', 'r110p', 'p151s', 'p278a', 'v173m', 'h179y', 'e2
 85k', 'y220c', 'g244d', 'invariable', 'aiming', 'immunonegative', 'proposal',
 'meticulous', 'dedifferentiated', 'liposarcomas', 'hnscc', '906', 'casp8', 'p
 rocaspase', 'goals', 'swallowing', 'hearing', 'chemoradiation', 'salisbury',

'pmxs', 'iix', 'bowtie', '961', '1202', '469', '005343', '613', '591c', 'pdp k1', '226', 'e507q', 'r49c', '001664', '1127', '000546', '1183', 'prodomain', 'p20', 'hampered', 'fragmented', 'gc7', 'cancergenome', 'ded', 'caspases', 't rail', 'zymogen', 'casp3', 'casp6', 'casp7', 'maybe', '600185', 'ifdis', 'dis proportionately', 'counseled', 'mastectomy', '2to', 'mvs', 'abkevich', 'nonma mmalian', 'incorrect', 'methodssource', 'analyzedthe', 'regressions', 'caribb ean', 'probandsno', 'genotypeethnic', 'deletioneuropeana', '899', '993', '97 2', 'asee', 'optionsas', 'formview', 'mathml', 'sourceturn', 'mathjax', 'onfo r', 'straightforwardly', 'quantifiable', 'lrs', 'mutationsto', 'vanishingly', 'homozygotes', 'asview', 'onwhere', '066', 'probandswe', 'pj', 'qj', 'therefo review', 'constellations', 'rk', 'isrk', 'onlet', 'onthenview', 'onandview', 'ontherefore', 'isview', 'onthat', 'aspersonal', 'individualbrca1', 'dcis', 'ordiagnosis', 'age3', 'or7', 'cancer1', 'age2', 'or5', 'or2', 'andreporte d', 'ors', 'probandsor', 'typehistory', 'brca2personal', 'cancera', 'aassesse d', 'optionsin', 'groupings', 'overstate', 'reassurance', 'pedigreeelsewher e', 'antoniou', 'arrive', 'classto', 'byview', 'onthis', '2ln', 'partitions', 'subdivided', 'ivs', 'gtrag', 'utrs', 'ifdi', '1862', '3098', 'pufferfish', 'statacorp', 'resultswe', 'llr', 'llrs', 'n723d', '1011', '1012', 'uninforma tive', 'optionsfrequency', 'optionstable', 'causalityllrgene', 'causalitybrca 1', '431', '818', 'r1495m', '4603g', '808', '054', 'ivs19', '12g', '5181del 3', 'ivs6', 'ivs15', 'ivs17', 't1685i', 'g1738r', 't1685a', 'ivs12', '2del21i nsa', 'a1623g', '4987c', 'ivs11', 'ivs18', '24brca2', 'r2659k', '8204g', 'g27 48d', 'i2627f', 'e2663v', 'p3039p', '9345g', 'd2723g', 't2722r', 'ivs24', 'r2 336h', '7235g', 'w2626c', 'ivs21', '12653p', 'd3095e', 'r2502c', '21table', 'neutralityllrgene', 'neutralitybrca1', 'p1614s', 'e597k', 'ivs16', '20a', '17g', 'p1238l', 'p334l', 'i1275v', 'p1859r', 'r504h', 's1266t', 'i1044v', '723', 'v1247i', '11delt', '598', '799', 'i9251', 'n473s', 'r1203q', 'g890 v', 'e842g', '741', 'm1652t', '728', 'd369n', '593', 'd369del', '1225del3', 'k862e', 'r1028h', 'e143k', 'd642h', '13c', 'd1546y', '088', 'a622v', '963', '828', '10g', '803', 'i1405v', 'q804h', '729', 'k1109n', '11t', 'q155e', 'p16 371', 't1349m', '462', 'i1858l', 'n1468h', 'm297i', 'e1682k', '317', 'a280g', 'e1419q', 'f1662s', 'h1402y', 'd420y', 'm1361l', '104brca2', 'k2950n', 'r2108 h', 's1733f', 'g1529r', 'l1019v', 'a75p', 't3349a', 'r1190w', 'p1819s', 't630 i', 'g1771d', 'k1690n', 's1172l', 'q2384k', 'c554w', 'd2312v', 'ivs26', '20 c', 'g602r', 'e462g', 'n56t', '666', 'h2074n', 'r2973c', '685', 'c3198r', '54 1', 'i1929v', 'n1228d', '726', 'h1918y', 'v1306i', '979', 'y3098h', 'v2969m', 'v894i', 'i1349t', 'q1396r', 'n2113s', '749', 'r2842h', 'ivs25', 'v3079i', '5 45', '20t', 'r2888c', 't2250a', 'f1524v', '982', 'l1904v', '976', 't582p', 'd 3170g', 'd1280v', 'p168t', 'n2048i', '639', 'k2729n', '614', 'd2665g', '406', 'r3052q', 'g1194d', 'p375s', 'v2908g', 'd806h', '255', 't1354m', 'n1102y', 'y 3092c', '16c', 'c1365y', 'n2436i', 'k2411t', 'l2396f', 'k513r', '108table', 'characteristicsno', 'llrvariant', '2all', '923', '63mv', 'mv', '854', 'inva riants', '959', '103splice', '0other', '16ain', 'optionsoverall', 'df', 'disc ussionusing', 'goldgar', 'knowing', 'noncarrier', 'convincing', 'provider', 'prevalences', '37the', 'reassured', 'specialists', 'alike', 'heretofore', 'ptprt', 'homophilic', 'fniii', 'cams', 'paschos', 'wheelock', 'freiss', 'vi gnon', 'rptps', 'cam', 'brady', 'kalnay', 'pcp', 'aricescu', 'homophilicall y', 'gebbink', 'sap', 'zondag', 'tonks', 'vecchio', 'cismasiu', 'pcp2', 'burg oyne', 'flavell', 'mcardle', 'baculogold', 'graces', 'logan', 'gentamicin', 'solon', 'pvl1393', 'gaagatctatggcgaggcctcgccg', 'ccggaattcggatcctggcgctc tgagcccg', 'gcgtggaggctctgcatggatgagcaaggac', 'caatgggtttcgtcatccatgcag agaggcccacgcccattg', 'ctgcacacttggcttggttcacgatcagctccgcgt', 'tacgcggag cgtcgtaaaaccaggaccaagtgtcag', 'ccgccaggcccccaccgaggaagacgttccaggag', 'ctcctgaaacgtcttcgttggggaggccctggcgg', 'gaggagctgggtgcagactaaaatttcg ctccatcca', 'tggatggagctgaaatttagtctgaccaccagctcctc', 'gtcaccactcggttgcac cggcgccatccaccagaatt', 'aattctggtgaggcaccggtaatccgagttgtgac', 'sk18', '8 9991', '5ml', '25mg', 'flask', 'neutravidin', '55mm', '60min', 'jla20', 'mult

iimager', 'trituration', 'vials', 'gyratory', 'te', 'metamorph', 'objects', 'n30', 'n0', 'i395v', 'y412f', 'r453c', 'q479e', 's492f', 'n510k', 't605m', 'v648g', 'a707t', '1708p', 'tgcccacagaacgtggaaagtcgttagacatcagagcc', 'cttccac gttctgtggcca', 'cagtggagcccttcggcttcgcgtgaccgcgtccat', 'aagccgaaggcgtccca ctg', 'ctacaccctgcgaggcctgtgcccttcatgaccatccg', 'acaggcctcgcagggttag', 'gag cgaggagctgtggactgaggaagacgtcc', 'ccaccaccagctcgcgtc', 'tgttctctagaa ttcatccaagggggccct', 'atgaattcttagaggaaca', 'tccagtggaaacctccaaggagaccaatgg gtcatca', 'cttgggagggttccactgga', 'ccattgaatgagacagacatgacc', 'atcacagtatgct g3', 'atgtctgtctcattcaatgg', 'attattgagtgtttc', 'ggggcccgtgagctatcgaat', 'cccgaaaagcactcaataat', 'ctacagcatctacttccagacactcagcaaagccaatgg', 'tctggaa tagatgctgttag', 'agcatctacttccaggcaccaggcaaagccaatggagag', 'gggcctgaaagttagatg ct', '2v5y', '2r1', 'adhesive', 'fasciclin', 'connectin', 'dtrk', 'ark', 'neuroglian', 'capricious', 'bellost', 'hortsch', 'pulido', 'shinza', 'kameda', 'snow', 'nonadhesive', 'army', 'spodoptera', 'frugiperda', 'gouveia', 'galestin', 'inohara', 'raz', 'cam1', 'phan', 'baculoviral', 'aom', 'apposing', 'rp tp', 'fnii', 'anders', 'dephosphorylates', 'besco', 'kaiso', 'daniel', 'reynolds', 'hypotonia', 'feeding', 'psychomotor', 'paternal', 'pdest12', 'h21', 'virginia', '1a1a', 'consanguineous', 'girl', 'prematurely', 'brothers', 'preterm', 'centile', 'myopia', 'femoral', 'menstrual', 'bleed', 'curly', 'sparse', 'warts', 'mentally', 'thumbnail', '4800', 'cardiomyopathy', 'echocardiograph y', 'intraventricular', 'septum', 'mitral', 'valve', 'prolapse', 'regurgitation', 'generalised', 'extensor', 'clonic', 'normalisation', '35g', '1b1b', '46 6t', 'coincidence', 'hyperpigmented', 'revision', 'ectodermal', 'nails', 'f15 6i', 'disappearance', 'worsened', 'channels', 'syndrome2', 'educational', 'gt paseactivating', 'internally', 'dideoxynucleotide', 'pat', 'hsqc', 'heteronuclear', 'quantum', 'coherence', 'noes', 'enhancements', 'noesy', 'tocsy', 'reprint', 'proc', 'acad', 'sci', '15nh4', '2so4', 'nh4', 'cpm', 'unlabeled', 'nitrate', 'nan3', '15n', 'mhz', 'sweep', 'recycle', '5n', 'exp', '3700', '447 5', 'polyomavirus', 'chloramphenicol', '32pi', 'ucl', 'unl', 'prebound', '12 3', '1hn', 'a59', 'ras7', '16', '1275', 'protons', 'dnn', 'strings', 'connectivities', 'nucleotidebinding', 'highresolution', '031', 'sizable', 'purine', 'aml1', 'eto', 'anything', 'evi1', 'nup98', 'leukocytopenia', 'fpd', 'runt', 'rhd', 'cbf', 'd171n', 's291fsx300', 'transplants', 'collaborate', 's291fs', 'aml1b', 'pmys', 'nakamura', 'e39', 'prestimulated', 'retronectin', 'sublethal', '135cs', 'radioprotective', 'meiermethod', 'morbid', 'hemacolor', 'cytopin', 'erythroblasts', 'transplantable', 'sublethally', 'lysing', '4g2', 'carlos', 'proposals', 'nonlymphoid', 'lightcycler', 'premix', 'seconds', 'ccagat gtcacatgacagtggaaagcacta', 'ccgggttggcatgactcatattaaccatgg', 'tacctaaccctga cagctatgg', 'tcgggttggagatatcagagtgcag', 'mn1', 'gttatccatctgcatcagcatctgg', 'ggccttcaacttcatgaacagc', 'mds1', 'kurokawa', 'proviruses', 'bubble', 'aattgaaggagaggacgtgtctg', 'tcgaaggtaaggaaacggacgagagaagggagag', 'gactc', 'tccc ttctcgaaatcgtaaccgttgcgtacgagaatcgctgtcctctcc', 'ttc', 'vectorette', 'cgaatcgta accgttgcgtacgagaatcgct', 'cgagctcaataaaagagccccacaacccc', 'ataaaagagccccacaacccc tcactcgg', 'gggggttca', 'sick', 'ly5', '355', 'aml1a', 'ids', 'tsuzuki', 'plt', 'mcv', 'jolly', 'polychromasia', 'poikilocytosis', 'orthochromatic', 'karyorrhexis', 'corpuscular', 'huet', 'hypersegmented', 'pancytopenia', 'hypocellular', 'normocellular', 'transplantability', 'anisopoikilocytosis', 'erythroblast', 'bh51', 'dp12', 'uplanfl', 'cytometolic', 'biphenotypic', 'collaborated', 'proviral', 'oligoclonal', 'morishita', 'multicloning', 'collaborates', 'cotransduction', 'pmys40', 'mfg', 'hepatosplenomegaly', 'leukemia18', 'obscures', 'something', 'heuser', 'slape', 'collaborators', 'evidence30', 'shorthen', 'insertional', 'finding25', 'findings25', 'y641', 'monomethylated', 'h3k27me1', 'dimethylated', 'h3k27me2', 'h3k27me0', 'me2', 'methylate', 'tunnel', 'mef2b', 'mll3', 'eed', 'rbap48', 'aebp2', 'trimethylated', 'recirculating', 'repopulating', 'resist', '641', 'germinal', 'gcb', 'dlbcls', 'fls', 'y641f', 'demethylase', 'dimethylation', 'monomethylation', 'elisas', 'methyltransferases', 'mantle', 'extranodal', 'waldenstr', 'macroglobulinemia', 'lymph

oplasmacytic', 'y641n', 'y641h', 'iie', 'vide', 'infra', 'loses', 'preference s', 'nucleosomes', 'y641s', 'heterogeneously', 'vmax', 'h2a', 'h4', 'dhl', 'g lp', 'ehmt1', 'h3k9me2', 'set7', 'rotates', 'lysines', 'adenosyl', 'sam', 'pr edictable', 'y245', 'trimethylate', 'h3k4me1', 'h3k9', 'g9a', 'y1067', 'dimet hyltransferase', 'trimethyltransferase', 'neurospora', 'crassa', '677', 'men 1', 'carcinoid', 'lipoma', 'ependymoma', 'hyperparathyroidism', 'hpt', 'posit ionally', 'menin', 'jund', 'genotypephenotype', 'antecedents', 'jmedgenet', 'mr', 'reserve', 'gastrin', 'glucagon', 'determinations', 'csge', 'direction ally', 'gdb', 'noncarriers', 'acth', 'f11', 'f49', 'gh', '3bp', 'cyanol', 'ac rylamide', 'mbi', 'fermentas', 'zeta', 'kodak', '2265', 'cgg', '6025', 'fa', 'c21', 'c23', 'c33', 'c9', 'c53', 'c77', 'c80', '50th', 'husband', 'unfile d', 'symbols', 'hai', 'evidently', 'obligatory', 'summarise', 'phenocopies', 'modern', 'gep', 'examiners', 'doctor', 'ulcers', 'hypercalcemia', 'parathyroidectomy', 'forearm', 'chromogranin', 'lever', 'cortisol', 'dexamethasone', 'endoscopic', 'nonfunctioning', 'microadenoma', 'fulfilling', 'scheduled', 'intraoperative', 'lymphadenectomy', 'resect', 'synaptophysin', 'uicc', 'tra nsfusions', 'zollinger', 'tracer', 'ventral', 'jejunum', 'pn1', 'v0', 'pn0', 'ischemic', 'postoperative', 'uneventful', 'trusight', 'workbench', '60k', 'ncbi37', 'probemix', 'ups', 'jejunal', 'ileal', 'flushing', 'ileum', 'adren als', 'men4', 'portuguese', 'statements', 'scintigraphy', 'reminder', 'islet s', 'p18ink4c', 'cdc7', 'answering', 'insulinomas', 'breeding', 'mandated', 'fvb', '129sv', 'ubc9', 'fertilized', 'eggs', 'intraperitoneally', 'pancreat a', 'sheep', 'tritc', '10ug', 'e800', 'westbury', 'cy2', 'guinea', 'titan', 'adenoviruses', 'pulsed', 'elb', 'bl', 'montgomery', 'phosphoimaging', 'excise', 'asynchronous', 'p21cip1', 'complementing', 'leakiness', 'quicker', 'cdk n1c', 'p57kip2', '11p15', 'imprinting', 'hypomethylation', 'paternally', 'mat ernally', 'imprinted', 'featuring', 'episodes', '294', 'pth', 'radiofrequency', 'grandfather', 'arterial', 'diazoxide', 'octreotide', 'brother', 'adrenal ectomy', '3229', 'laparoscopic', 'splenectomy', 'medically', 'acromegaly', 'florid', 'cushing', 'leucocytes', '244a', '4046', '1323', '2361', '1392', '236 5', '4126', 'telomere', '11p', 'centromere', 'bisulphite', 'zymo', 'immunoper oxidase', 'h19', 'relieving', 'ultrasonography', 'pnets', '1336', 'bartsch', 'beckwith', 'wiedemann', 'babies', 'foregut', 'metallothionein', 'excising', 'blunt', 'accreditation', 'biweekly', 'housing', 'nicking', 'shaking', 'spect rophotometry', 'sakura', 'overlaid', 'nitroblue', 'hitachi', 'tracheal', 'cystostat', 'gelatin', 'micrometer', 'macintosh', 'instat', 'founders', 'mt1', 'undeleted', 'monthly', 'densely', 'neatly', 'pushed', '907', '278', 'osteop orosis', 'spared', 'pathologically', 'multistage', 'diffused', 'angioma', 'un availability', 'crabtree', 'pancreases', 'dehydration', 'embedding', 'pour', 'recherche', 'contre', 'gion', 'ligue', 'fellowship', 'laboratoire', 'tique', 'cnrs', 'lyon', 'cedex', 'univ', 'rip', '4836', 'dissociated', 'pipettin g', 'deoxynucleoside', 'fasted', 'clotting', 'ultrasensitive', 'enon', 'ecadh erin', 'pem', 'p65', 'p52', 'p50', 'trka', 'unmask', 'deliver', 'ect2', 'ost', 'dbl', 'tsukamoto', 'repertoire', 'lsc', 'g2a', 'afford', 'overlooked', 'histopaque', '1077', 'bstxi', 'dh5', 'midiprep', 'kilobases', 'bosc23', 'ea gles', 'pbabepuro', 'bsiwi', 'mariano', 'chemilluminescence', 'napo4', 'conical', '326', '674', '490', 'aml3', 'aml4', 'neurotrophin', 'tpr', 'and8', 'der egulates', '345', 'deglycosylated', 'hyperglycosylated', 'reuther', 'lambert', 'trkc', 'unreported', 'birc7', 'states1', 'rccs', 'subtype2', 'confounde d', 'collecting', 'flcn', 'sdhb', 'sdhc', 'sdhd', 'microphthalmia', 'tfe3', 'oncocytic', 'sarcomatoid', 'fischer', 'huber', '441', '537', '072', '086', 'chd3', 'crtc1', 'smit', 'succinate', 'pseudosubstrate', 'cbs', '14c', 'gsc', 'semaphorin', '12a', 'acyl', 'coenzyme', 'synthetase', '20b', 'bh1h', '21 b', 'corepressor', '25a', '21d', 'apex1', 'stability75', 'expression5', 'pbrm 1', 'kdm5c', 'malaise', 'rafs', 'paraspinal', 'discoidin', 'massarray', 'fg r', 'tnk2', 'ptk6', 'gak', 'g466vbraf', 'wtbraf', 'h226', 'h661', 'h1666', 'h 322', 'hpi', 'packer', 'plk1', 'azd0530', 'credible', 'interchangeable', 'amp

lifier', 'epitomized', 'protomer', 'superficially', 'chk2', 'sfk', 's217', 'y719', 'y754', 'hp1', 'talalay', 'facscan', 'aminoactinomycin', 'prewarmed', 'sigenome', 'nontargeting', 'kolch', 'conway', 'immunocomplex', 'unstained', '260', '3812', 'e709', 's768', 'comprehension', 'taiwan', 'bronchoscopy', 'smoked', 'discretion', 'radiography', 'procured', 'thoracentesis', 'macrodissection', 'rnalater', 'echo', 'bronchoscopic', 'cytology', 'nonadenocarcinomas', '832', '634', '092', '518', 'iiia', 'g719d', '1861r', '0months', 'iiia', 'br21', 'pallis', '1861p', 'mitsudomi', 'hata', 'kancha', 'singlets', 'doublets', 't790', 'identifications', 'v742a', 't847i', 'v769m', 'v851i', 'kosaka', 'a871e', 'notice', 'favorably', 'nssnps', 'cupsat', 'h773r', 'combating', 'stein', 'glycinerich', 'yahoo', 'market', 'shtml', 'nssnp', 'rajasekaran', 'g735s', 'servers', '005219', 'blastp', 'bioedit', 'editor', 'premise', 'fastsnp', '2j6m', 'svm', 'gromacs', 'vmd', 'spdbv', 'guex', 'gromos96', 'hess', 'cubic', 'descent', 'worked', 'equilibrate', 'swift', 'ru', 'iteration', 'ligplot', 'dockings', 'murray', 'pos', '521', '624', '668', '3941', '767', '735', '029', '768', '810', '902', '1257', '788', '399', '977', '988', '1034', '1048', '1162', 'immense', 'matthews', 'avizienyte', 'evidences', 'fluctuations', '807', 'sterical', '756', '867', 'gly719', 'glu762', 'met766', 'cys775', '2493', 'evi', 'dence', 'aee788', 'miy', 'plernl', 'migri', 'paa', 'migr1', 'glycerolphosphate', 'autokinase', '1688p', 'v851a', 'i853t', 'e866k', 'v851', 'i853', '753inss', 'n826s', 'r776c', 's784f', 'd761n', 'anilinoquinazolines', 'pyrrolopyrimidine', 'combinational', 'sparked', 'guiding', 'bystander', 'e884k', 'percutaneous', 'ajcc', 'union', 'unidimension', 'metastasized', 'ascites', 'mutationa', '2238', 'insv', '2593', '2252', '2240', '2257', 'p753', '1838p', 'v843i', '884', '344', 'drainage', 'p753inss', 'antiegfr', 'pay', 'doctors', 'cooking', 'fumes', 'k753m', '2028', 'proangiogenic', 'p70s6', '40s', 'umbilical', 'huvec', 'huvecs', 'p42', 'ink', 'millimeter', 'neomarker', 'contingency', 'pittsburgh', 'quantikine', 'vegf165', 'densitometric', 'autoradiograms', 'sucrose', 'flurouracil', 's2448', 'herceptest', 'mvd', 'chemoattractant', 'normalizing', 'stimulators', 'surprise', 'faint', 'immunostained', 'cci', '779', 'blimp1', 'lymphomagenesis', 'abramson', 'shipp', 'alizadeh', 'shaffer', '2002a', 'centroblast', 'plasmablastic', 'mediastinal', 'rosenwald', 'savage', 'oxp', 'lenz', 'iqbal', 'pasqualucci', '18q24', 'card11', 'compagno', 'cd79b', 'prdm1', 'immunoglobulins', 'shapiro', 'shelef', '2002b', 'irf4', 'angelin', '6q21', 'uniparental', 'gaidano', '6q15', 'y345h', 'graphed', 'ciita', 'monitors', 'p48r', 'p48t', 'y149d', 'sunyaev', 'sudhl2', 'u266', 'piskurich', 'c569y', 'bjab', 'documenting', 'id3', 'cd138', 'blimp1cd19ko', 'lps', 'blimp1cd19wt', 'coregulated', 'lipopolysaccharide', 'b220lo', 'klein', 'tunyaplin', 'rck8', 'oci', 'ly8', 'cd20', 'tonsil', 'shctrl', 'shbcl6', 'lymphoproliferative', 'ohinata', 'rickert', 'casola', 'blimp1c', '1ko', 'immunized', 'igm', 'pnahi', 'cd21', 'cd23', 'mz', 'lymphoproliferation', 'mzbch', 'obliteration', 'effacement', 'lpds', 'mzbchs', 'lymphoproliferations', 'transited', 'srbc', 'falini', '1250', 'balt', 'pax5', 'dalla', 'favera', 'habcl6', 'cattoretti', 'ranuncolo', 'basso', 'juszczynski', 'spib', 'schmidlin', 'leucci', 'exempt', 'deidentified', 'cfr', 'cd10', '001198', 'rl', 'mascus', 'miltenyi', 'biotec', 'lipopolysaccharides', 'pinco', 'percp', 'blimp1f1', 'blimp1cd19het', '1het', 'visibly', 'software', 'mantel', 'casitas', 'tradefoffs', 'tumorigenensis', 'interconversion', 'tkbd', 'lhr', 'p13k', 'ncbl', 'tyr371', 'outperform', 'pcbl', '2y1m', '2y1n', '4a4c', 'rectangular', 'namd', 'charmm27', 'repairpdb', 'optimizes', 'boltzmann', 'gbind', 'gcbl', 'ge2', 'gbindmut', 'gbindwt', 'evdw', 'samut', 'gfd', 'beatmusic', 'calculates', 'empirical', 'buildmodel', 'gfoldmut', 'gfoldwt', 'gfold', 'eris', 'rosetta', 'popmusic', 'webtools', 'provean', 'mutationassessor', 'inca', 'langdon', 'pc_efl', 'powerplex', 'ampfistr', 'profecction', 'incubate', 'fairlawn', 'iodoacetamide', 'tabs', 'tumbling', 'ba85', 'sanford', '22321', '3f10', 'hsc70', '7298', 'na934v', 'na931', 'omat', 'eastman', 'rochester', 'epson', 'perfectio', 'v750', 'asymmetrically', 'g415v', 'g413r', 'g415s', 's376f', 'y371d', 'l1

405p', 'kink', 'noticeably', 'm400', '1405', 'w408', 'r420', 'w408s', 'i383', 'ftp', 'pub', 'panch', 'w93', 's94', 'r420q', 'r420l', 'r420p', 'zygocity', 'undetermined', '035', 'mirrored', 'c396r', 'h398q', 'y371h', 'k382e', 'c381a', 'k382', 'e373', 'py371', 'y371', 'tan', 'm374v', 'v430m', 'p4281', 'q249e', 's80n', 'h94y', 'l399v', 'g375p', 'p395a', 'v391i', 'outperformed', 'foldable', 'tradeoff', 'coupling', 'collective', 'sustaining', 'competes', 'complicate', 'physics', 'ensembles', 'deducing', 'multidimensional', 'interim', 'uncovers', 'surveys', 'antecedent', 'subclasses', 'fragmentary', 'dcc', 'privacy', 'biospecimen', 'biospecimens', 'cnas', 'algorithms13', 'alterations3', 'park2', 'irs2', 'ptprd', 'abundances', 'htm', 'temozolomide', 'ccnu', 'prompting', 'competency', 'r1391s', 'r1513', 'e25', 'e29', 'q1966', 'v777a', 'gbm32', 'l10del', 'p17del', 'lay', 'ish2', 'd560', 'n564', 'n345', 'n345k', 'd560y', 'n564k', 's565', 'dkrmns', 'r574del', 't576del', 'w583del', 'alkyl', 'alkylating', 'juxtaposed', 'cytosines', 'cpgs', 'alkylated', 'unequivocal', 'thesis', 'relieves', '0018', 'mdm4', 'cdkn2b', '9p21', 'ensured', 'cdkn2c', 'pink3r', 'frontline', 'k535', 'e550', 'subclassified', 'pax', 'myogenesis', '1172', 'reexamined', 'figure1e', '0127', 'rh30', 'figure2c', 'intramuscular', 'figure3e', 'pelvic', 'figure3f', 'figure4a', '535', 'multiethnic', 'figure5a', 'paralog', 'transductants', 'table33', 'figure6e', 'figure7a', '0071', 'figure7b', '0079', 'figure7c', 'figure7d', 'intravenously', 'figure7e', 'figure7f', 'figure8a', 'figure8b', 'figure9b', '9d', '9e', '9f', 'multimodal', 'myoblasts', 'pax7', 'rhabdomyoblast', 'g388r', 'fgfr5', 'cytoplasmatic', 'phosphotransfer', 'doubly', '629', '760', '1092', '1101', '2324', 'protrudes', 'upward', '644', '643', 'shield', '1fgk', 'mohammadi', 'bae', '2psq', '2pvf', 'hare', 'aminoalkylamino', 'quinolinones', 'chk1', '2gdo', 'tyr653', 'unfolds', 'rfu', 'vegfr1', 'vegfr3', 'piperazine', 'quinolinone', 'contour', 'cocrystallized', 'ser446', 'apolar', 'benzimide', 'carbonic', 'phe317', 'successively', 'logical', 'tyr1235', 'disengages', 'asn535', 'disengagement', 'enforcing', '523', 'grace', 'propanesulfonic', 'tcep', 'homogenizer', '500g', 'tev', 'amicon', 'ultra', '30k', '3rd', 'pod', 'sypro', 'precipitants', 'morpholine', 'ethanesulfonic', 'ethylene', 'molrep', 'rebuilt', 'coot', 'refmac', '580', '568', '582', '583', 'phasing', 'k618a', 'h718y', 'e578g', 'v716m', 'a681t', 'p6541', '2280', '120435', '660', 'awareness', 'rq1', 'mlv', 'unconserved', '000249', 'nonfluorescent', 'steponeplus', 'stepone', 'asei', 'mun', 'mmrvariant', 'mmrmissense', 'google', 'identifiers', 'corroborates', 't117m', 'r659p', 'fulfillment', 'cosegregate', 'polish', 'cosegregates', 'undiluted', 'mismatched', 'bthe', 'raevaara', 'practically', 'unimpaired', 'cellcycle', 'parcells', 'stirewalt', 'radich', 'reindl', 'frohling', 'choudhary', 'spiekermann', 'grundler', 'rocnik', 'bagrintseva', 'schittenhelm', 'yanada', 'thiede', 'whitman', 'midostaurin', 'lestaurtinib', 'tandutinib', 'knapper', 'forrest', 'cavet', 'getz', 'wood', '581a', '662g', 'v194m', '1052g', 'd324n', '1172t', 'y364h', '1434c', 's451f', '1751g', 'v557i', '1797a', 'y572c', '1857t', 'v592g', '2293g', 'm737i', '2574g', 'g831e', '2583g', 'r834q', 'enst00000241453', 'refsnp', 'rs35602083', 'rs35958982', 'v592', 'r834', 'v592a', 'd835insalg', 'ley', 'w51', 'nontransforming', 'pmut', 'mmb2', 'pcb', '8080', 'y589', 'y591', '591f', 'trcn', 'homozygous', 'y572', 'favorability', 'hook', 'phenolic', 'e661', 'n816', 'r815', 'y842', 'd811', 'n841', 's836', 'caryboxyl', 'g831', 'deleteriously', 'karlsson', 'levinson', 'mathematical', 'discerning', 'bell', 'jeffers', 'mulloy', 'plaza', 'menacho', 'n520y', 'syampurnawati', 'nonhematopoietic', 'schnittger', 'gari', '579', 'martelli', 'platanias', '591', '831', 'deangelo', 'fiedler', 'interpatient', 'hoped', '96aqueous', 'mission', 'noncomplementary', 'scholl', 'r174q', 'allogeneic', 'trd', 'warren', 'pfalg', 'biochrom', '32dc13', 'affinipure', 'sabiosciences', 'beyotime', 'myeloperoxidase', 'lactoferrin', 'hematopathologic', 'immunophenotyping', 'cgx', 'jião', 'ear', 'bii', 'bv', 'naphthol', 'punctuate', 'tcr', 'coprecipitates', 'oligomer', 'maturational', 'stayed', 'tracked', 'myeloblast', 'promyelocyte', 'myeloblasts', 'circulation', 'thrombosis', 'capillaries', 'hemorrhages', 'b

lown', 'homodimerizes', 'enver', '1345', 'lithuanian', 'czech', 'exac', 'exemplarily', 'ethnically', 'recql', 'bonafide', 'pursuit', 'amazing', 'bears', 'danger', 'misuse', 'marketing', 'huge', 'nbn', '657del5', 'canadian', 'ih', '278730', 'photosensitivity', '601675', 'cerebrooculofacioskeletal', 'cos2', '610756', 'val77fs', 'phe568fs', 'ser746fs', '1903', '2150c', 'asp423asn', 'lithuania', 'lys751gln', 'asp312asn', 'cadd', '461', '611', '631', '678', 'hd1', 'hd2', '423', '487', 'fes', 'figs1b1b', 'phe568tyrfs', 'arg631cys', 'leu461val', 'asp513tyr', 'val536met', 'arg601gln', 'lt', '8x', 'cz', '0x', '1381c', 'rs121913016', 'rs144564120', '1704deltt', '1758', '32c', 'rs238417', 'outnumber', 'esp', 'kilometers', '1844', 'excited', 'dresden', 'prague', 'arg112his', 'hardy', 'val77alafs', '162072007', 'munich', 'mgz', 'vilnius', 'santariskiu', 'klinikos', 'brno', 'cdh1', 'rad51d', 'r455', 'samtools', 'recalculation', 'snpeff', '0e', 'ez', 'nimblegenseqcap', 'user', 'phenopedia', 'navigator16', 'solid4', 'novoalign', 'annoivar', 'revertaid', 'ttaggtacatga', 'agctcaacgtggacg', 'ttatcttagatcagagctgctgagcaatct', 'pjet1', 'clonejet', 'klaus', 'austria', 'd513y', 'dam', 'frankfurt', 'hcr', 'deploys', 'xp6be', 'kraemer', 'arg683trp', 'attractene', 'uvic', 'photoproducts', 'luciferase', 'unirradiated', 'expasy', 'ta0057', 'thermoplasma', 'acidophilum', '4a15', 'q9hm14', 'mulitseq', 'taster', 'agvd', 'summarizing', 'rating', 'homologene', 'r337h', 'g334', 'amyloid', 'l344p', 'dsred', 'fida', 'methodssite', 'mutagenesisisthe', 'gaagatctaagctt', 'accatg', '1met', 'gaggagccg', 'cgggatccgtcgac', 'gtctgagtctcagg', 'pegfp', 'c3', '359', 'ccactagtccgcggccc', 'spei', 'xmai', 'p53re', 'hcmv1', 'letters', 'hcmv', 'nhei', 'pdred2', 'm340ql344r', 'optionsconstruction', 'hcmv2', 'dsredto', 'ppromoterless', 'dsred2', 'attaat', 'tccgga', 'gcgcgc', 'actagt', 'tacgta', 'gctagc', 'bsshii', 'snabi', 'gaacatgtcccaacatgtt', 'gggcgtccgctgtcggg', 'u24170', '2303', '2321', 'ptet', 'assaynci', 'ix71', 'uplan', 'mniba2', 'dichroic', 'mwig2', 'controller', 'manager', 'immunostain', 'antibodyh1299', 'conjugated', 'analysish1299', 'mf20', 'unfold', 'resultsrelationship', 'assayh1299', 'r342p', 'a347t', 'optionswe', '1330h', 'r337c', 'r337p', 'f341c', 'l344r', 'l330r', 'f338i', 'monotonously', 'pdred', 'optionsestimation', 'p53the', 'a594', 'immunocytochemically', '4j', '2441', '2444', '2477', '888', 'optionseffect', 'p53h1299', '2k', 'f328v', 'r337l', 'l348f', 'cellsmonomeraverage', 'e326g', 'y327h', 'y327s', 'f3281', 'q331h', 'q331p', 'q331r', 'i332v', 'g334l', 'g334v', 'g334w', 'r335g', 'r335h', 'm340q', 'r342l', 'r342q', 'e343g', 'e346a', 'a347g', 'e349d', 'd352h', 'a353t', 'q354e', 'q354r', 'g356a', 'g35w', 'e358d', 'e358k', 'mf', 'equipment', 'ishioka', 'nanomolars', 'sold', 'heterogeneities', 'counterbalancing', 'stressors', 'apoptogenesis', 'viz', 'p53td', 'entailed', 'leu344', 'unstressed', 'heightening', 'extinction', '1280', 'g356w', '6800', 'y327d', 'holder', 'effluent', 'igor', 'wavemetrics', 'jasco', 'spectropolarimeter', 'ellipticity', 'mateu', 'fersht', 'hutm', '3sak', 'modeler', 'l330p', 'l342p', 'contemporaneously', '1330a', 'l348s', 'chromatographed', 'catastrophically', 'cooling', 'phe328', 'leu330', 'arg337', 'phe338', 'phe341', 'leu348', 'devastated', 'met340', 't329i', 'p53tds', 'e339k', 'e339q', 'q354k', 'i322v', 'thermally', 'unfold', 'tds', 'disassociation', 'intranuclear', 'ws', 'neocarzinostatin', 'normoxic', 'normoxia', 'acerbate', 'lys351', 'crm1', 'pirh2', 'ile332', 'glu339', 'glu346', 'mistakes', 'intraprotein', 'complementarity', '1450', 'fibrils', 'gln331', 'sufficed', 'r47s', 'r72p', 'v217m', 'g360a', 'p47s', 'subtly', '309g', 'lowering', 'snp309', '19q', 'oligodendrogliomas', 'cic', '19q13', 'capicua', 'ependymomas', 'chemosensitive', 'clearing', 'uniqueness', 'codeleted', 'globally', 'fubp1', 'gro', 'calgary', 'buffy', 'coat', 'burlington', 'mississauga', 'bcc', 'haenszel', 'wgss', 'dibayes', 'disqualified', 'exomic', 'snvmix2', 'pileup', 'chr19', '467', 'unannotated', 'bloods', '100x', '3212', '2696', 'arg132his', 'arg172his', 'sox11', 'dach1', 'ptprz1', 'chr1', '78205319', '78193604', 'q96rk0', '2315', 'arg201trp', 'arg215trp', '47483653deltcc', '1515', 'disproportionally', 'v54', 'ex5', 'gliomatosis', 'cerebri', 'tics', 'symmetric', 'tic', 'bt54', 'bt88', 'oligodendroglioma'

tous', '30x', '47483801g', 'arg1515his', 'abyss', '1608', '2517', 'mrm', '154', '188', '628', '883', '887', 'chemosensitivity', 'imaginal', 'granular', 'medulloblastomas', 'spinocerebellar', 'scai', 'q35', 'dux4', 'etv5and', 'arg215', 'arg201', 'arg202', 'cicarg505trp', 'arg1515', 'whatever', 'necessitate', 'concentrate', 'distill', 'atxn2', 'atxn1', 'pave', '2209', 'lookup', 'viewport', 'iss', '1014', 'dacomitinib', '1016', '1520', 'designer', 'desorption', 'ionization', 'flight', 'fw1', 'tcttcacctggaaggggtcca', 'tgccacacctccactccgtcta', 'cattcatcgcttacccgt', 'gtatagggtagccgtttag', 'patternslepidic', 'a763', 'c775', 'd770n', 'gif', 'base64', 'r0lgodlhaqabaiaaaaaap', 'yh5baeaaaaa laaaaaaabaaeaaaibraa7', 'mct', 'molcanther', 'carousel', 'factoring', 'highwire', '42302', 'markup', '42236', 'iframe', 'postprocessors', '2chighwire', 'embed', 'nonmucinous', '775', 'h773', 'v774insnph', '42193', 'y764insffea', 'v774', 'c775insnhv', 'd770', 'n771insgt', 'd770insasv', 'taxane', 'multimodalit y', 'anecdotally', '42272', '716', 'pitfall', 'intronexon', '42251', 'pvai', '42290', 'a767', 'v769dupasv', 'd770dupavd', 'distally', '771', 's768instla', 'a767insai', 'blanket', 'unmet', 'elderly', 'astreneca', 'wilmington', 'chemona', 'exsmokers', 'succeeding', 'nonplatinum', 'platinums', 'premetrexed', 'vinorelbine', 'till', 'nonresponder', '939', 'included18', 'egrf', 'worst', '044', 'nonerlotinib', 'addario', 'discontinuation', 'contraindication', 'getting', 'echoes', 'ntroduction', 'radiographically', 'n771insnpg', 'n771inssvd', 'v774insh', 'd770dupasv', 'flparp', 'clparp', 'e762inseafq', 'd770dupsvd', 'p772', 'h773insh', 'deactivation', 'approximates', 'nonmeasurable', '0175', 'bid007', 's9a', 's9b', 's9c', 'sirna1', 'sirna2', 'pd168393', 'inward', 's10a', 'superimposes', 'insnpg', 'c797', 'pivots', 'outward', 'pivot', 'fqa', 'register', 'e762q', 'insffea', 'y764insfqq', 'facto', 'i759a', 'insfqqa', 'nonsensitizing', 'notwithstanding', 'wedge', 'sensitization', 'essence', 'proximate', 'counterscreening', 'pf00299804', 'update', 'leuarggluala', 'europeans', 'randomised', 'gly719ala', 'delglu746', 'ala750', 'ser752insval', 'delleu747', 'thr751', 'ser752', 'pro753insser', 'insx', 'aminoacid', 'subsis t', 'homodimerise', 'heterodimerise', 'epiregulin', 'gly719x', 'gly719ser', 'thr790met', 'favours', 'intertwined', 'inheritably', 'disfavours', 'push', 'aminoacids', 'tyr764', 'ala767', 'orienting', '7066', 'sensitising', '54egfr', 'asp761', 'glu762insglualaphegl', 'ala763', 'tyr764inspheglnluala', 'val1765inshishis', 'ala767insalaile', 'val1769dupalaserval', 'ser768insthrleual', 'ser768', 'asp770dupervalasp', 'val1769insvalalaser', 'val1769insalatrprth', 'val1769', 'asp770insalaserval', 'asp770insglyval', 'asp770inscysval', 'asp770insaspasval', 'asp770insglyserval', 'asp770insglyvalval', 'asp770insmeta laservalasp', 'asp770', 'asn771insservalasp', 'asn771insasnprogly', 'asn771insalaprotrp', 'asn771insasp', 'asn771insasp gly', 'asn771insgly', 'asn771insglyleu', 'asn771insasn', 'asn771insasnprohis', 'asn771inservalpro', 'asn771inservalgln', 'asn771insmetalathrpro', 'delasp770insglytyr', 'asn771', 'pro772in shis', 'pro772insasn', 'his773dupasnprohis', 'delasn771insglytyr', 'delasn771insglyphe', 'pro772', 'his773insproarg', 'his773instyrasnpro', 'his773insx', 'his773insaspprohis', 'his773insaspasnpro', 'his773insglaval', 'his773insthr prohis', 'his773insasn', 'his773insval', 's27', 'his773', 'val774insasnprohis', 'val774inshis', 'val774insprohis', 'val774insglyasnprohis', 'val774duphis val', 'val774insgly', 'val774insglyhis', 's21', 's26', 's30', 'val774', 'cys775inshisval', 'webappendix', 'hcc4011', 'ic50values', 'conc', '060', 'mw', 'extrapolations', 'ci1003', 'rudimentary', 'ipass', 'tribute', 'asn770', 'val765', 'ala767insaia', 'his773tyr', 'sequist', 'delasn770insglytyr', 'pf00299084', 'tyr790met', 'ala775instyrvalmetal', 'temsirolimus', 'nct00838539', 'see king', 'enrol', 'nct01090011', 'v765m', 'admission', 'expired', '387785', 'ce llar', 'r776g', 'unresponsiveness', '085', 'customization', 'descendants', 'taipei', 'cytologically', 'intolerable', 'sonography', 'lobectomy', 'globin', '005228', 'unidimensional', 'nonprogressive', 'rays', 'nonsmoker', 'nonmutant', 'backward', 'ordinal', 'increment', 'dela767', 'a763v', 'observational', 'fragmentation', 'concurred', 'oils', 'ncslc', 'kinki', 'chu', 'acttcacagcc

ctgcgtaaac', 'atgggacaggcactgattgt', 'transillumination', 'singed', '0996', '0917', 'nph', '2320', '2328', 'aaccccccac', 'd7', 'asv', '2316', 'gccagcggt', 'svd', '2311', 'gcgtggaca', '2312', '773insertion', '6783', '2666', '4236', '9999', '7025', '7186', 'gehan', '8593', 'calboplatin', 'ftegafur', '7215', 'quitted', 'n771', 'n771ins', 'baf250a', 'dualistic', 'organizes', 'endometriosis', 'mesodermal', 'chromosomally', 'protocol13', 'anonymously', 'endometriod', 'tw', 'kk', 'e545a', '1634', 'kt', 'n1044k', '17t', '21t', 'ucc', '1t', '3t', 'chromatogram', '0072', 'gynecologic', 'study1', 'b56', 'pr61', 'carboxyterminus', 'janssens', 'goris', 'lechward', 'virshup', 'pr65', 'pr55', 'pr72', 'pr93', 'pr110', 'arino', 'hemmings', 'calin', 'colella', 'ruediger', '2001a', 'takagi', 'koma', 'groves', 'ridges', 'xing', 'bending', 'bryant', 'kloeker', 'tolstykh', 'lcmt', 'homogeneity', 'mclr', 'macroseeding', 'semet', '992', '252', '439', '0083', 'fp', 'interhelical', 'burial', '2072', 'grasp', 'nicholls', 'kraulis', 'superhelical', 'curvature', 'convex', 'concave', 'tilted', '4300', 'laterally', 'overlay', 'dali', 'holm', 'sander', 'contini', 'graham', 'speciessecondary', '18490281', '30795206', 'gamma1', '1418819', '31083279', '31083295', 'rts1p', '532526', 'chizosaccharomyces', 'cpb', '19075706', 'ceb', '17557914', 'dmb', '60677747', 'atb', '2244898', '651', 'gln125', 'his339', 'phe340', 'trp382', 'buttressed', 'asp131', 'ser298', 'proto305', 'tyr307', 'phe308', 'lys256', 'tyr292', 'lys295', 'asp63', 'glu64', 'glu101', 'strengthened', '2001b', 'e64d', 'e64g', 'leu309', 'arg268', 'asp123', 'sit', 'trp140', 'phe141', 'arg183', 'glu214', 'glu100', 'trp257', 'holoenzymes', 'p65s', 'l101p', 'k343e', 'd504g', 'v545a', '189p', 'k331e', 'd492g', 'v533a', 'pro53', 'leu89', 'lys331', 'y456a', 'y495a', 'kamibayashi', 'withstand', 'val533', 'val545', 'twist', 'swings', 'pro406', 'interior', 'pivotting', 'bend', '1256', 'shugoshin', 'kitajima', 'riedel', 'mimicry', 'bayliss', 'swing', 'shoe', 'b55', 'wd40', 'alzheimer', 'pathogen', 'horseshoe', 'activities3', 'approx65', 'approx36', 'pp2b', 'calcineurin', 'pp4', 'pp6', 'tpdy307f1309', 'dimer7', 'complex7', 'pathogenesis12', 'b56gamma', 'alpha', 'abeta', 'suppressors19', 'drugs20', 'calpha', 'b56gamma1', 'ellipsoidal', 'b56bold', 'agr', 'bladed', 'latches', 'protruding', 'mumby', 'pr48', 'pr130', 'benecib', 'goedert', 'kins', 'sontag', 'polymerize', 'neurofibrillary', 'tangles', 'spillantini', 'weingarten', 'witman', 'alonso', 'drewes', 'ptau', 'vigorous', 'remedy', '4r0n', 'ser396', 'ikehara', 'experimenting', 'multiwavelength', 'convention', 'radiating', 'torus', 'wilson', 'grabs', 'leu87', 'val126', 'tyr127', '4270', 'interdigitate', 'pro131', 'tyr157', 'phe54', 'tyr60', 'arg21', 'arg257', 'asp218', 'blades', 'diverging', 'intersubunit', '274', 'e27r', 'k48e', 'd197k', 'k345e', 'phe84', 'leu90', 'glu93', 'glu94', 'lys95', 'tyr178', 'his179', 'ala93', 'ala94', 'ala95', 'ala178', 'ala179', 'lys345', 'wille', 'leulliot', 'magnusdottir', 'tap42', 'alpha4', 'pme', 'lcmt1', 'obstacles', 'solving', 'tough', 'pser396', 'gentry', 'longin', 'seleno', 'ronin', 'diffracting', 'vapor', 'peg35000', 'i4', 'selenium', 'x29', 'denzo', 'scalepack', 'otwinowski', 'phaser', 'ta6br12', 'shelx', 'sheldrick', 'noncrystallographic', 'phased', 'calc', 'interpretable', 'brunger', 'presume', 'goldberg', 'biosource', 'cancers3', 'activity25', 'cystadenocarcinoma', 'skov3', 'ser63', 'ser73', 'pp2', 'tumors18', 'studies3', 'gallo', 'mutation7', 'usc', 'gypsy', 'subtyping', '1946', '1949', 'd2', 'monoubiquitination', 'diepoxybutane', 'deb', 'noma', 'barcelona', 'vall', 'hebron', 'madrid', 'universitario', 'fe', 'joan', 'heine', '2588', 'smhc', 'fis', 'funds', 'contract', 'microbiology', 'monoubiquitinated', 'analysis17', 'afro', 'mitomycin', 'mmc', 'pd20', 'nonubiquitinated', 'cd28', 'fancc', 'fang', 'unspecific', '1947', 'ivs4', '14th', 'century', 'arrived', '15th', '1425', 'roma', 'millions', 'romania', 'irish', 'homozygote', 'portugal', 'wellknown', 'uoc', 'fancg', 'fancf', 'skeleton', 'ebv', 'fance', '1815', 'xhol', 'toronto', 'hal', 'broxmeyer', 'trireagent', 'antisera', 'prosci', 'grompe', 'hitrap', 'hydroxymethyl', 'aminomethane', 'antiubiquitin', 'electroblotting', 'powdered', 'pmmp', 'faa', 'faf', 'continuity', 'ho', 'morgan', 'navigator', 'etude', 'polymorphisme',

'humain', 'ceph', '1347', 'cephb', 'genotyper', 'bifunctional', 'banded', '95', 'freedom', 'fancd1', 'externally', 'ubiquitinylated', 'multimeric', 'jy', 'volunteer', 'immunodepleted', 'dep', '058', 'horizontally', 'exchanger', '2p22', 'unremarkable', 'bipartite', 'nlss', 'conservatively', 'monosomic', 'garcia', 'higuera', 'al13', 'coimmunoprecipitate', 'hematologically', 'bmf', 'xie', 'nonleukemic', 'chaperones', 'fancb', 'reverted', 'circumstance', 'revertant', 'stigmata', 'speak', 'meetie', 'atresia', 'soulier', 'aplastic', 'slx4', 'kupfer', 'kashiyama', 'whenever', 'deoxynucleotide', 'triphosphates', 'carboxyfluorescein', 'disfavouring', 'exchanges', 'unpubl', 'morocca', 'brazilian', '3788', '3790del', 'saharan', 'blacks', 'col7a1', 'dystrophic', 'settled', '11th', 'arab', 'lyonnet', 'facilities', 'mustard', 'phytohemagglutinin', 'autologous', 'nonclonal', 'commun', 'cytom', 'matologie', 'rue', 'publishing', '655', 'pha', 'trade', '0165', 'hypotonic', 'glacial', 'permeabilize', '2150', 'argon', 'florescence', 'reacted', 'dyskeratosis', 'congenital', 'usable', '1751', 's10038', '0037', 'ayadi', 'ahmed', 'ali', 'culaire', 'pasteur', '1002', 'rns', 'habib', 'faculte', 'medecine', 'biotechnologie', '227650', 'microcephaly', 'joenje', 'timmers', 'anaemia', '1992a', '1992b', 'winter', '2000a', '2000b', 'taniguchi', 'howlett', 'tachibana', '16q24', 'pronk', 'turkey', 'microinsertions', 'wijker', '890', '2172', 'insg', 'allogenic', 'cousins', 'genethon', 'dib', '1u', '94c', '72c', 'radioactively', 'savino', 'quiagen', 'prostar', 'genehunter', '00001', 'multipoint', 'parametric', 'tdt', 'commands', 'makers', 'roman', 'numeral', 'arabic', 'numerals', 'intermarker', '8728', '1057', 'inherit', 'ancestors', 'disequilibrium', 'b3', 'cctg', 'krawczak', 'senopathy', 'hypermutable', 'sib', 'adair', 'steensma', 'graft', 'grafted', 'telangiectasia', 'prolymphocytic', 'al25', 'banks', 'analysis31', 'mrec', 'mul', 'mum', 'typed', 'al33', 'mug', 'seminested', '2r', 'cleaned', 'beta2', 'microglobulin', 'hemopoietic', 'ncn', 'previously35', 'unsatisfactory', '67k', 'cancer36', 'thymines', '0048', 'went', 'relapsing', 'silencers', 'al22', 'comet', 'ted2', 'meristems', 'p53175p', 'wooster', 'neuhausen', 'caretaker', 'inviable', 'hakem', '1998b', '150i', '199r', '202s', '144p', '163n', '168y', '234c', '248w', '181c', '219h', '158h', '240r', '215c', 'pcb6', 'identified', '234h', 'parker', 'hpv16e7', 'ej', '173l', 'expressed', 'corollary', '175p', '181l', 'transactivatioan', 'wagner', 'possession', 'tantalising', '283h', '1988a', 'paul', '175h', '273c', 'stanchina', 'zindy', 'haematoxylin', 'bag', 'palter1p53', '16e7', 'pcmvbeta', 'dynex', 'sx21', 'fredersdorf', 'formol', 'fixing', 'do1', 'facsort', 'compard', 'acetyltransferases', '1165', 'destabilisation', 'grunstein', 'acetylate', 'eckner', 'gayther', 'ecacc', 'dnazol', '3p24', 'trishcl', 'engine', 'tetrad', 'mj', 'peltier', 'hct15', 'heteroduplex', 'abi377', 'abi3100', 'pa1', 'ch1', '1795', 'rubinstein', 'taybi', 'petrij', 'kung', 'gcn5', '2415', 'occurring', 'oshima', 'bromodomain', 'nonhistone', 'siha', 'mv1lu', 'mink', 'hacat', 'ex15', 'hincii', 'goodman', 'toyo', 'perry', 'smads', 'multimerized', 'caga', '22a', '3tp', '33a', '33b', '55a', '55b', 'nucleosomal', 'syt', 'rewiring', 'plasmablast', 'bcrs', 'anergy', 'nfat', 'healy', 'goodnow', 'endless', 'cleary', 'borche', 'chiorazzi', 'ferrarini', 'stevenson', 'ngo', 'terminates', 'carma1', 'cbm', 'bcl10', 'malt1', 'thome', 'rawlings', 'gerondakis', 'siebenlist', 'cyster', 'hen', 'egg', 'lysozyme', 'ighel', 'cd40', 'g123s', 'mut3', 'l2321i', 'mut10', 'l251p', 'mut2', 'd387v', 'mut6', 'rag1', 'preparing', 'rela', 'comparably', 'nontransgenic', 'igma', 'allotypic', 'nontransduced', 'auto fluorescence', 'baff', 'enders', 'thien', 'fsc', 'fairfax', 'helper', 'martins', 'calame', 'overturn', 'autoantibodies', 'cd19low', 'b220low', 'plasmablasts', 'multiplier', 'neutralization', 'measurably', 'ohashi', 'broke', 'tfh', 'oscillations', 'bclxl', 'bfl1', 'slows', 'mccully', 'pomerantz', 'enlist', 'vdj', 'wardemann', 'shlomchik', 'microbial', 'autoimmunity', 'mombaerts', 'ml5', 'prdm1flox', 'kallies', 'mb1', 'hobeika', 'littermate', 'australian', 'phenomics', 'curtin', 'cared', 'pfx', 'acr', 'biomolecular', 'minato', 'ecotropic', 'fgk4', 'bioxcell', 'il4', 'dotap', 'facsaria', 'adoptive', 'bottome'

d', 'fortessa', 'ra3', '6b2', 'cd25', 'pc61', '1d3', 'cd4', 'rm4', 'qdot', 'c d44', 'im7', 'hyhel19', '3c5', 'carbonate', 'thermomax', 'inova', 'ago1', 'tas 3', 'chp1', 'rits', 'silences', 'centromeres', 'piwi', 'derepress', 'tgs', 'r its1', 'transcripts5', 'pombe4', 'organisms6', 'metazoan', 'identified12', 'g w182', 'tnrc6b', 'bodies13', 'isoforms17', 'metazoa18', 'derepresses', 'paz', 'overhang', 'sirnas22', 'sirna27', 'activity30', 'thisminimal', 'kiaa1093', 'vivo17', 'thd', 'subfamily34', 'pirna', 'pathway35', 'detectably', 'repress ion33', 'macroh2a1', 'mh2a1', 'deltantd', 'piwi1', 'piwi2', 'deltac4', 'oligo peptide', 'suffices', 'deltac', 'w1', 'w2', 'wdouble', 'thanf', 'thana', 'w1d ouble', 'w2double', 'mof', 'component41', 'bodies16', 'wg', 'repeats41', 'aw g', 'gwg', 'competed', 'family34', 'bodies17', 'well142', 'bodies43', 'metazoa 6', 'involved6', 'reported17', 'function17', 'leptomycin', 'lmb', 'structures 27', 'tyr625', 'lys660', 'tyr529', 'lys533', 'gln545', 'val797', 'arg812', 'a rchaeoglobus', 'fulgidus', 'strand27', 'tyr123', 'stacks', 'complex27', 'y529 l', 'y123l', 'acids28', 'piwi28', 'ionic', 'lys127', 'gln137', 'autoradiograph', '2bgg', 'arg385', 'ile400', 'piwi27', 'l209a', 'w2441', 'k127a', 'q137a', 'r385a', 'i400a', 'a280', 'hooks', 'thermophilic', 'leu209', 'trp244', 'provo cative', 'translation48', 'trnas', 'oligomer27', 'defects4', 'deltatas3', 'me chanism5', 'otrir', 'reaper', 'repression54', 'pairing50', 'mirna54', 'tgs5 5', 'hints', 'sensed', 'holes', 'abasic', 'guanines', 'triplets', 'thermodyn a mically', 'purported', 'rrrcwwgyyy', 'c238', 'c242', 'h179', 'c124', 'c135', 'c277', 'workers', 'imagine', 'electronically', 'oxidize', 'ems', 'alkylation', 'cyclotron', 'ethylmaleimide', 'isotopically', 'qtrap', 'm1331', 'v203 a', 'n268d', 'tbe', 'hybond', 'semidry', 'phosphorimaging', 'storm', 'fla', '9000', 'solar', 'uvb', 'centrifugal', 'carboxyethyl', 'phosphine', 'vacuo', 'lysyl', 'endopeptidase', 'formic', 'centrifuging', 'elu', 'fmol', 'abscie x', 'nanolc', 'imported', 'n239y', 'emsas', 'autoradiogram', 'desalted', 'lc l', 'floating', 'conveying', 'extraordinarily', 'instrumentation', 'tucked', 'rearrange', 'funnel', 'triad', 'conforming', 'chapter', 'nemesis', 'condemn ing', 'quarantining', 'buttress', 'excitement', 'poisons', 'impelling', 'dest ruct', 'tissue2', 'rewarding', 'tantalizing', 'therapy6', 'masks', 'degradati on10', 'destroyed', 'overexpressed11', 'mice13', 'atr', 'ckii', '3sigma', 'p5 3aip', 'astonishingly', 'tnp', 'pathway18', 'entail', 'cription', 'complex2 0', 'coloured', 'domain65', 's100b', 'betabeta', 'shades', 'molmol69', 'comp5 3', 'divides', 'sequences22', 'pupupuc', 'pyrimidine23', 'v143a', 'questionab le24', 'state26', 'denaturation27', 'range28', 'enthalpy', 'extrapolation', 'mol28', 'irreversibly', 'binding28', 'deltag', 'classes28', 'fingerprint', 'temperature29', 'design30', 'selection31', 'domain32', 't284r', 'h168r', 'm ol33', 'v143a31', 'nmr32', 'treats', 'mouse35', 'human20', 'misfolded', 'domi nance36', 'militate', 'revive', 'therapy38', 'hybridize', 'constructs40', 'st ructure42', 'tetramer43', 'nullify', 'effect45', 'vivo44', 'way32', 'reversed 32', 'equilibria', 'tether', 'affinity42', 'macromolecule', 'suffice', 'rasti nejad', 'studies6', 'microtitre', 'denatures', 'convenient', 'thermostabilit y', 'propyl', 'polycyclic', 'substituent', 'ionizable', 'discontinuous', 'gra sp70', 'clefts', '257042', '31398', 'v173a', 'r175s', 'treatment6', 'leukaemi a47', 'facet', 'taxol', 'concentrations49', 'pitfalls', 'deleting', 'site50', 'cornerstone', 'strides', 'degradation51', 'therapy39', 'onyx', 'cells54', 'p ifithrin', 'pft', 'methylcytosine', '5mc', 'malignance', 'fingers', 'dsbh', 'nog', 'postreplicative', 'bird', 'dnmt3', 'dnmt1', 'goll', 'bestor', 'hydro xymethylcytosine', '5hmc', 'bacteriophages', 'wyatt', 'globisch', 'kriaucionis', 'heintz', 'szwagierczak', 'tahiliani', 'jbp1', 'dioxygenase', 'hydroxylat es', 'cliffe', 'jbp', 'tet1', 'tet3', 'iyer', 'formylcytosine', '5fc', 'carbo xylcytosine', '5cac', 'oxidizes', 'oxoglutarate', 'og', 'tdg', 'demethylatio n', 'branco', 'cimmino', 'shi', 'dawlaty', 'lorsbach', 'ono', 'lian', 'roll', 'unstructured', '1129', '1156', 'balls', 'mc6', 'kilodalton', 'commassie', 'h iload', '1099', 'oxalylglycine', '1481', '1843', '1131', '1143', '1925', '146 4', 'c11', 'a12', 'wraps', 'flipped', '1288', '1312', 'invisible', '1463', 'e

xterior', 'g6', 'h1382', 'd1384', 'h1881', 'keto', 'bidentate', 'o2', 'schofield', 'r1261', 'h1416', 'r1896', 's1898', 'hxd', 'h1382y', 'd1384v', 'hydrolytes', 'dsdna', '2structure', 'hexagons', 'flip1', 'flip2', 'abh2', 'r1896m', 'r1261g', 's1898f', 'joins', 'protrude', 'cys3his', 'rosenstr', 'ci2', '1y4d', 'radisky', 'bpn', 'zn1', 'zn3', 'corners', '1159', 'h1380', 'crick', 'g7', 'r1262', 's1290', 'g8', 't9', 'k1299', 's1303', 'w1291', 'm1293', 'y1294', 'y1295', 'r1302', 'g5', 'mc7', 'phosphorous', 'flipping', 'angstrom', '1291', '1296', 'ggsggs', 'w1291r', 'k1299e', 's1303n', 's1290a', 'y1295a', 'r1262a', 'alkb', 'h1904', 'n1387', 'endocyclic', 'n4', 'y1902', 'hydroxylation', '5mcpx', 'm1293a', 'y1294a', 'n1387a', 'h1904r', 'y1902a', '1125', '18000', 'nucleosides', 'duplexes', '5mcpg', '5mcpa', '5mcpc', 'jbp2', 'trypanosome', 'hydroxymethyluracil', 'hmu', 'd218', 'd396', 'r287', 'r463', 'cpc', 'ab3', 'alkbh3', 'fto', 'methyladenine', 'mea', 'mec', 'methylthymine', 'methyladenosine', 'm6a', 'sedgwick', '3buc', '2iuw', 'sundheim', '3lfm', 'ahb3', 'demethylates', 'f124', 'h171', 'taverna', 'predications', 'chg', 'chh', 'cxxc', 'cxxc4', 'sin3a', 'blaschke', 'metabolites', 'gainof', 'dosages', 'eur', 'wilkins', 'gastroenterology', 'hepatology', 'endocrinology', 'ttingen', 'giuliano', 'uni', '1097', 'meg', 'prohibited', 'q556', 'k558del', 'peritoneum', 'k550', 'w557', 'sympysis', 'w557r', 'k558', 'v559del', 'a502', 'y503dup', 'w557g', 'ifosfamide', 'epirubicin', 'communicate', 'motivate', 'team', 'restarted', 'reestablished', 'l576p', 'kit1576p', 'mitogenesis', '10nras', 'anus', 'hiph', 'taqdna', 'metaphor', 'biowhittaker', 'blasted', 'ccagagtgcataatgactg', 'acccaaaaaggatgacatgga', 'ex13', 'atcagttgccagtttgct', 'tttataatctaggattgcc', 'tgtgaacatcattcaaggcgat', 'caggactgtcaagcagagaatgg', 'ex12', 'tccagtcatgtgctgcttc', 'gcaagggaaaagggagtct', 'ex18f', 'accatggatcagccagtct', 'ex18r', 'tgaaggaggatgaggcgtacc', 'ldr', 'mutations13', 'braf13', 'com12', '2v', 'com13', 'com61', '2l', '3ha', '3hb', 'ex1', 'ex2', '3180g20', '722f21', 'metasystems', 'grinding', 'powergen', 'omni', 'marietta', 'immun', 'biomax', 'gary', 'braunshweig', 'hygromycin', 'genpulser', 'kitv559d', 'wtkit', 'gaagtacatggaaaggatgtgaggagata', 'tatccctcaacatccttccactgtacttc', 'atagacccaacacaacctccttatgtacaaa', 'tttgtgatcataaggagggtgtgtgggtctat', '35482', 'sprycel', 'tasigna', 'annexinv', 'dickison', 'magna', 'phosphoryated', 'tyr20', 'anorectal', 'intraepidermal', 'pdgra', 'mmunohistochemistry', '1fish', '2western', '3fish', 'cod', 'fluorescence', '253', 'f3v559d', 'melanoma6', 'paralleling', 'lines18', 'immunoexpression', 'payne9', 'v560g', 'jmr', 'physiopathologica1', 'inspired', 'juxta', 'kid', 'd810', 'f811', 'g812', 'extremity', 'emplace', 'swap', 'transformations', 'gastro', 'v559', 'v560', 'deregulatory', '1t46', '693', 'neglected', 'kitv560g', 'kitv560d', 'kitd816y', 'kitd816n', '9v7', 'kitd816h', 'maestro', 'kitwt', 'amber', 'parm99', 'hydrated', 'hydration', 'h816', 'macromodel', 'relax', 'berendsen', 'thermostat', 'restraining', 'langevin', 'piston', 'velocities', 'pmemd', 'ptraj', 'popped', '2nd', 'visited', 'lone', 'pcbi', '1003749', 'e002', 'issued', 'dha', 'djms', 'ajms', 'gor4', 'sopma', 'simpa96', 'predator', 'jnet', 'eigenvectors', '3n', 'collectivity', 'e005', 'e007', 'jpgare', 'communications', 'moneta', 'idss', 'cps', 'lfa', 'covariance', 'matrices', 'autocorrelations', 'eigenvalues', 'formalism', 'uncorrelated', 'pcut', '040', 'mdmat', 'dcut', 'commute', 'gephi', 'replicas', 'fluctuating', '825', '885', 'displacements', '817', '3g0e', '3g0f', 'cartoons', 'd792', 'y823', 'aas', 'sorma', 'dissimilarity', 'and7', 'tiny', 'uncharged', 'provokes', 'carboxylic', 'g560', 'y570', 'n787', 'i789', 'y568', 'f848', 'y846', 'v570', 'v569', 'e561', 'computation', 'norms', 'd792e', 'e011', 'v816', 'cts', 'comnet', 'communicating', 'composing', 'lowly', 'comnets', '965', '857', 'c787', '1783', 'm780', 'h790', 'sockets', 'cpwt', 'socket', 'f782', 'd851', 'e861', 'unconstrained', 'communicated', 'cascading', 'ptb', 'spatio', 'scales', 'cdk5', 'adenylate', 'd816n', 'reversibility', 'disturbing', 'departure', 'd802v', 'sensibility', 'nussinov', 'breathers', 'relocation', 'adapts', 'perpetual', 'decoupling', 'centriole', 'fergus', 'ouch', 'southwest', '55905', '284', '3623', 'farrugia', '3523', '3531', '158

7', 'mislocalize', '8http', 'r2659t', 'atggaaaggctggaaaagaa', 'aagaatccaagtggata', 'centrin', '20h5', 'centrioles', 'sce1', 'pcbasce', 'flagbrca2', 'gf pexpressing', 'human9', '9http', 'nocom', 'approximated', 'gaussian', 'cutpoints', 'n319t', 'r2520q', 'l2647p', 'd2723a', 'r2784w', 'l2865v', 'a2643g', 'n372h', 'k3326x', 'd2723h', 'r3052w', 'r496c', 'exon skipping', 'hdr', 'cutpoint', 'verifying', 'discriminates', 'discriminated', 'multifunctionality', 'non sensemediated', 'overinterpretation', 'hydroxyglutarate', '2hg', 'onco', 'mardis', 'kang', 'r172', 'ketoglutarate', 'r172k', 'ehrlich', 'colman', 'decarboxylating', 'ceccarelli', 'mtbstfa', 'derivatization', 'derivatized', 'enantiomer', 'chirality', 'stereoisomers', 'esterification', 'butanolic', 'anhydride', 'kamerling', 'enantiomers', 'esters', 'intrasmple', 'deberardinis', 'interconvert', 'nadhd', 'rydstrom', 'idh3', 'denton', 'ochoa', '1948', 'siebert', '1957', 'consume', 'isotopic', 'reductive', 'carboxylation', 'aconitate', 'r132c', 'r132g', 'r140q', 'r140', 'r100', 'neomorph', 'hartmann', 'carboxylating', 'inborn', 'kolker', 'nadp', 'anaplerotic', 'latini', 'hematopoetic', 'mutagens', 'maker', 'xenotransplantation', '703185', '091', 'fcoll', 'mncs', 'tgttgagatggacgccta', 'ggtgtactcagagccttcgc', 'ctgcctttgtggcctaag', 'attctggtgaaagatggcg', '1lwd', 'bc009244', 'potb7', 'bc012846', 'sport6', 'g515a', 'g395a', 'sf188', 'prolico', 'threo', 'od340', 'ab55271', 'na931v', 'nadoc', 'sc49996', 'idh3a', 'ab58641', 'sc1616', 'chilled', 'redissolving', 'deionized', 'butyldimethylsilyl trifluoroacetamide', 'regis', 'derivatize', '7890a', '5ms', '5975c', 'splitless', 'ramped', 'butyldimethylsilanol', 'tetra', 'tbd ms', 'isotopomer', 'isotopes', 'weckwerth', 'evaporation', 'dryness', 'tributylamine', 'luo', 'electrospray', 'kv', 'quadrupole', 'bajad', 'exactive', 'orbitrap', 'operated', 'learners', 'performances', 'unconfirmed', 't725m', 'insmuch', 'harmless', 'realize', 'landscapes', 'attributes', 'kinbase', 'prokino', '1451', 'emboss', 'uniref90', 'preprocessing', 'names', 'billion', 'extant', 'ancient', 'mapgaps', 'hmmer', 'sampler', 'oner', 'weka', 'splits', 'disjoint', 'evaluators', 'arithmetic', 'j48', 'forest', 'dtbn', 'lwl', 'knn', 'unseen', 'imbalanced', 'learner', 'harmonic', 'recall', 'fn', 'voting', 'vote', 'votes', '1003545', 'e003', 'e004', 'accuracies', 'decimal', 'carpenter', 'du becco', 'bioexpress', 'hams', 'py1173', 'py845', 'py1045', 'py1086', '498', 'e006', 'e008', 'indexes', 'performs', '989', 'vouches', 'leak', 'euclidean', 'fg1', 'g724s', '1858q', 'e746k', '1jiu', 'camk1a', 'camk2b', 'ck2a', 'ck2b', 'tesk1', 'tesk2', 'tousled', 'tlk1', 'tlk2', 't725', 'vaik', '6178', 'y1086', 'y1045', 'y1173', 'causativeness', 'natal', 'myocytes', 'e63k', 'alleviated', 'cgap', 'contingent', 'archetype', 'heptahelical', 'arguably', 'disputed', 'vulval', 'mesoderm', 'relayed', 'nominally', 'cytokinesis', 'ordinarily', 'angiotensin', 'adrenergic', 'atrial', 'natriuretic', 'myocardium', 'diastole', 'adaptations', 'atria', 'relations', 'hsv', 'sequestering', 'modulatory', 'envisioned', 'feig', 'psv', 'sp6', '705', '1047', 'pzipneo', 'pxp2', 'nahco3', 'myocyte', 'sarcomeric', 'transferrin', 'chloroquine', '934', 'agitation', 'luciferin', 'acylation', 'supposition', 'unmasks', 'substantiates', 'unprocessed', 'inference', 'squelching', 'muscarinic', 'trivial', '3y1', 'processes', 'provisional', 'p190', 'intermediaries', 'provoked', 'outweigh', 'contrabutory', 'passively', 'rebinds', 'productively', 'p34q', 'i36m', 'v152g', 'nacl2', 'f234', 'grds', '2a2a', 'jmml', 'enumerate', '2c2c', '2b2b', '2d2d', 'proteinsa', 'fates', 'tuveson', 'rebinding', 'mx1', 'defects1', 'approx50', 'ras3', 'hydrolyzing', 'gaps5', 'levels9', 'jmmls', 'mutations11', 'specificities1', 'jmml8', 'approx35', 'jmmls11', 'vitro14', 'buccal', 'val14', 'thr58', 'asp153', 'alpha5', 'figurefull', '51k', 'tablefull', 'expected17', '36k', 'data14', 't581', '89k', 'data19', 'shown20', 'extensively5', 'promoter21', 'inactivation22', 'activities6', 'rates24', 'not26', 'd61g', 'animals29', 'tissues29', 'n308s', 'n308d', 'jmml2', 'ras14', 'approx45', 'erlangen', 'freiburg', 'universities', 'publish', 'described2', 'transgenomics', 'elsewhere9', 'c57b1', 'inhalation', 'described16', 'stemspan', 'sfem', 'env', 'lipofectamine2000', 'facsvantage', 'previously30', 'hcdc4', 'cpd', 'cullin1', 'cpd'

s', 'srebp1', 'mowiol', 'canto', 'diva', 'peqlab', 'hexanucleotide', 'conform s', 'ni2', 'nucleoplasm', 'uncouples', 'balances', 'myelogenic', 'neuropathology', 'bonn', 'roessle', 'andtumor', 'delbrueck', 'berlinbuch', 'eva', 'warde lmann', '2120', '53011', '5353', '5030', 'ukb', '1211', '1743', 'irresectable', 'extragastrointestinal', 'computerassisted', 'histomorphologic', 'responses', 'nda', 'recurrences', 'dodx', 'ctype', 'boverall', 'xthe', 'perforation', 'regressive', 'pericardium', 'desmin', 'dianova', 'gelectrophoresis', 'weiterstadt', 'hsu63834', 'twentytwo', 'kitreceptor', '11mutation', '1745', 'synchronously', 't670e', 's709f', 'f44', '1746', '0164', '62005', 'gistcase s', 'v560del', 'a504', 'y505ins', 'n567', 'p573del', 'q556h', 'w557t', 'd579del', 'd816e', 'l576del', 'g565del', 'd572del', 'm552', 'w557del', 'e554a', 'p551', 'e554del', 'w558del', 'e562del', 'e561del', '1747', '11led', 'exon13', 'exon14', '1748', 'proofs', 'phosphatidyle', 'inositole', 'broudy1', 'smcd', 'vivo15', 'gnnk', 'rmil', 'rhscf', 'mbq', '17806', 'stat5a', 'deltyd', 'rg', 'a502y503dup', 'aydup', 'delv559v560', 'delvv', 'r634w', 'n822k7', 'e839k', 'line10', 'd820g4', 'v530i', 'internalized', 'v559v560', 'wcls', 'v560g19', 'tyd', 'nonrandomly', 'receptortyrosine', 'markerfor', 'imatinibinduced', 'affairs', 'oncologic', 'hitchcock', 'dartmouth', 'lebanon', 'hanover', 'turku', 'finland', 'fund', 'lester', 'regina', 'glossary', 'flicts', '0732', '183x', '2429', 'csti571', 'b2222', 'ofimatinib', 'deteriorating', 'protocolspecified', 'reviewboardofeachparticipatinginstitution', 'andwritteninformedconsent', 'directedmutagenesiswas', 'dnawere', 'conditionsin', 'dnamolecules', 'imatinibmesylatewas', 'purchasedfrom', 'akitolentiviralshort', 'obtainedfromwilliam', 'hahn', 'ccggccataaggttcggttctgtactcgagt', 'acagaaacgaaaccttatggaaaa', 'aattcaaaaa', 'ccataaggttcggttctgtactcgagtagacagaaacccat', 'agei', '1puro', 'cotransfecting', 'pmd', 'postinfection', '21the', 'eachlysatewas', 'byimmunoblottingfor', 'stan', 'korsmeyer', 'lystates', 'k8', 'ttggatatttgtc', 'caggaactgagc', 'agagaacagctccaaagaaaaatcc', 'k9', 'atgctctgtttctgtactgcc', 'cagagcctaaacatcccccta', 'k11', 'ccctttctccccacagaaacccat', 'acacggctttacccat', 'agacaataaaaggcagcttgacgac', 'k12', 'ctgcacaatggccttcaa', 'caaaaa', 'gcacaactgcaaac', 'k13', 'catcagttgccagttgtgc', 'k14', 'ctcacccctttctaaaccc', 'ttctt', 'cccatgaactgcctgtcaac', 'gaccatgatgtcccttct', 'tctctgtaaagtactcttgg', 'ttg', 'agtgatctgcctgcaagttcacattag', 'gctctaaatgctctttctc', 'tgtattcacagaga', 'cttggc', 'ggatttacattatgaaagtacacagg', 'k18', 'ttctgttcaattttgtttagctc', 'gca', 'ggacaccaatgaaactt', 'k19', 'gcaactgtgtctcaggaagc', 'ccctcaacatctgggttct', 'k20', 'aaaacaagctgagggcattg', 'tgggagaaggggattctat', 'k21', 'tggtcggtttagg', 'gactgct', 'aaccatcatgaaagccaaag', 'ccatctcctaacggcttttgc', 'caggacatgggtttccat', 'p17', 'acctgatgatgttcctgctgc', 'cgtccacactccactactg', 'tgaaggaggatgaggctgac', 'tgctgtggatcatcagtgagt', 'ttccctccagtggtgtttca', 'caatgcactgagcgtttgt', 'ctccccctagaccacagac', '4765', 'phosphokin', 'kitmonoclonal', 'pdgframutations', 'anidenticalmethodologywas', 'exceptfor', '24as', 'ofmay', 'themedia', 'n', 'treatmentfailurefor', 'treatmentfailuresweredividedinto', 'twogroups', 'patientswhenever', 'withkitexon', 'preimatinib', 'al21', '4766', 'treatment failure', '400mg', '600mg', 'stabilization', 'twentynine', 'debulking', 'sityone', 'originalmutation', 'whichwas', 'knownfor', 'wkvve557', '921', 'pmye 551', 'iab', '497', '558evqwk', 'kv558', '559n', '615', 'vv559', 'wkvv557', '560c', 'myevqw552', 'n822h', 'c809g', 'kpmeyevqwk550', '778', 'delyevqwk553', 'evqwk554', 'delkpmeyevqwk550', 't6701', '560f', 'evqwkvveeingnnyyvi554', '237', 'k558np', '416', 'kpmeyevqwk550', '557fl', '881', '4767', 'threikitexon', '9mutant', 'werefound', 'additionalkitkinasemutations', '638', 'n822y', '378', 'exams', '4768', 'baselinekitv560d', 'tumorwasfunctionally', 'substitutionfound', 'resistancewas', 'triplemutant', 'toimatinib', 'usingkitexon', 'loopweremutatedinimatinib', 'rangingfrom', 'n882k', 'alleleic', 'c809', 'de 1554', '4769', 'ofd820gwithkitexon', 'resultedinmoderateimatinib', '14wasmode rately', 'ic50of', 'mutantkitisoform', 'cellswasimatinib', 'ic90of', 'activationin', 'kitknockdown', '4770', '4771', 'themolecularmechanisms', 'resistance in', 'differencesin', 'ofigists', '4772', 'differencein', 'thefrequency', 'tha

nkitexon', 'kitexon', '11mutant', 'themost', 'onkit', 'signalingfor', 'dispersal', 'y1003f', 'endosomal', 'monoubiquitin', 'morphogenic', 'degradative', 'multivesicular', 'vacuole', 'endocytic', 'uba', 'uev', 'escrt', 'endophilin', 'babco', 'richmond', 'p4d1', 'psen217', 'psen473', 'irdye800', 'gilbertsville', 'vande', 'woude', 'andel', 'rapids', 't47d', 'faststart', 'rotor', 'corbett', 'sydney', 'accacagtccatgccatcac', 'tccaccaccctgttgctgta', 'cd1', 'madin', 'darby', 'fort', '975', 'rinsing', 'acidified', 'soybean', 'vineland', 'saponin', 'immu', 'shandon', 'linda', 'v76', 'mbd', 'tkb', 'p170', 'uncleaved', 'p145', 'posttreatment', 'uo126', 'internalize', '1003', 'e0', 'syk', '1235', 'nivelle', 'treier', 'expended', 'hamid', 'vincent', 'ollendorff', 'nofat', '6400', 'phosphorylatable', 'uncouple', 'script', 'projected', 'ron', 'orthologues', 'plexins', 'semaphorins', 'unsuspected', 'versatility', 'rodrigues', 'yoon', 'levkowitz', 'miyake', 'joazeiro', 'waterman', 'fournier', '70z', 'py1003', '1356', '1390', 'weidner', 'y1356', 'rong', 'impart', 'basally', 'vigna', 's985', 'gandino', 'villa', 'moruzzi', 'eguchi', 'wai', 'syriann', 'v824v', 'superiority', 'modality', 'vendors', 'elected', 's727', 'camarillo', '2m', 'gtype', 'arcturus', 'utp', 'upgrade', 'gcos', 'biosun1', 'complab', 'friendswood', '20ng', 'h23', '4q32', 'vascularized', 'osteosarcomas', 'fli', 'rhabdomyosarcomas', 'hmec', 'systematical', 'pik3cas', 'mda157', 'mda231', 'mda361', 'bt20', 'sfihe', 'fungizone', 'gentamycin', 'ethanolamine', 'sum159', 'sum149', '1954', 'quikchange1', 'plenti6', 'clonaseii', 'producer', 'virapower', 'blasticidin', 'invivogen', 'equalized', 'thr389', 'thr24', 'thr32', 'sfih', 'mediums', 'solidified', 'iodonitrotetrazolium', 'q60k', 'k111n', 'c420r', 'p539r', 'e542k', 'h701p', 'withdrew', 'sfhe', 'sfh', 'wtpik3ca', 'phpho', '3000cell', '9100', 'comparasion', 'morpho', 'pfoxo1', 'bacteria', 'novelties', 'gymnopoulos', 'cef', 'eot', 'nanogram', 'q60r', 'p110c', 'rept', 'isakoff', 'shutting', 'happening', 'pik3a', 'appointed', 'leukemiant', 'eln', 'panelists', 'weigh', 'elude', 'identified1', 'spent', 'narrower', 'exist18', 'moving', 'prescribed', 'k247r', 'y320c', 'italic', 'preexist', 'cyr', 'anyway', 'worthless', 'subcategories', 'chances', 'operationally', 'milestones', 'milestone', 'pcyr', 'warning', 'advise', 't315a', 'f3171', 'y253h', 'f359v', 'eln33', 'gimema', 'party', 'studies57', 'wherever', 'prescrenn', 'diagnosis29', 'ccyr67', 'unmutated', 'matters', 'm244v', 'g250e', 'e355g', 'advised', 'conjecture', 'bystanders', 'excretion', '1248v', 'e255v', 'e279k', 'f311l', 'v379i', 'l384m', 'l387m', 'f486s', 'warned', 'patients78', 'soundly', 'advising', 'map2k2', 'melanoma1', 'agents5', 'nras6', 'mek15', 'amplification8', 'loss9', 'cot10', 'egfr11', 'loss12', 'transcripts13', 'regulation14', 'dysregulation6', 'hgf16', 'priorities', '200x', '92x', 'val600', 'lys601delinsglu', 'hoxd8', 'phosphorylation18', 'mek2c125s', 'vitro7', 'mek2v35m', 'mek2146f', 'mek2n126d', 'engendered', 'vrt11e', 'mek1v60e', 'mek1g128v', 'mek1v154i', 'mek1p124s', 'mek1p124l', 'mek1g276w', 'mek1f53y', 'inhibitors5', 'previously7', 'phosphorylation', 'lineage20', 'mitfr217', 'skmel19', 'syndrome21', 'case22', 'vastly', 'undersample', 'clear12', 'models12', 'dabrafenib26', 'ptenk128t', 'ptenh93d', 'pteny86fs', 'pik3cah1047r', 'subunit27', 'phenotype28', 'nrasq60h', 'ptenr233', 'gdc0941', 'ptenwt', 'ptennull', 'a2058', 'unclear5', 'melanoma33', 'factor35', 'cancers36', 'previously9', 'longitudinally', 'nationwide', 'decog', 'unite', 'salient', 'genes38', 'inhibitors39', 'agents40', 'carcinogenesis38', 'form22', 'underestimates', 'schedule42', 'secured', 'essen', '4961', 'trios', 'citrisolv', 'crosslinks', '51306', 'platform43', 'described43', 'palindromic', 'forked', 'immobilization', 'spring', 'vortexing', '40ng', 'picogreen', '60ng', '50ng', 'minijanus', '20pm', 'flowcells', 'multiplexing', 'rta', '76bp', 'aligns', 'manages', 'genepattern44', 'contest45', 'fingerprints', 'mutect46', 'method47', 'algorithm48', 'oncogenesis7', 'pcw57', 'doench', 'plix', '41395', '23555', 'described20', 'm36', 'pdonor223', 'mr217', 'plx', 'vsvg', 'treme', 'shoter', '1000x', 'bradford', 'ontogeny', 'mpms', 'advancing', 'lfs', 'r248q', 'al10', 'varley', 'al11', 'p1521', 'r158h', 'catchments', 'lifelong', 'accs', 'manchester', 'nonadrena

l', 'arar0332', 'submit', 'relational', 'catchment', 'tp53r337h', 'prkar1a', '375g', '327', 'c229r', 'scarcity', 'pex4', 'pin3', 'mdm2snp309', 'presentat ion30', 'hac1532', 'mutbs', 'to3c', 'tp53r273h', 'struck', 'eeles', 'lf1s', 'gonzalez', 'y163c', 'h193p', 'c275x', 'e285v', 'g334r', 'f134y', 'n235d', 'multiply', 'al35', 'referrals', 'al40', 'raymond', 'tp539', 'gone', 'tp534 2', 'carriage', 'backseat', 'germane', 'sao', 'paran', 'endemic', 'transmissi ble', '1966', 'curitiba', 'willingness', 'endocrinologists', 'mellitus', 'hyp othyroidism', 'reared', 'amplitaq', 'placental', 'nusieve', 'fragmenting', 'hy bridizing', 'agtc', 'hhai', 'deparaffined', 'do7', '337h', 'deoxynucleotidyl transferase', 'creatine', 'd17s1832', 'centimorgans', 'centirays', 'd17s786', 'shgc', 'pentanucleotide', 'girls', 'boys', 'masculine', 'weakness', 'hyperfu nction', 'aunts', 'aunt', 'virilizing', 'ideogram', '44a', 'supraphysiologica l', 'brazilians', 'qidong', 'agricultural', 'pesticides', 'industrial', 'oxid ant', 'fasay', 'tiling', '5p15', 'trio', '11q22', 'iap', 'spotted', 'stringen tly', 'karyogram', 'characterizes', '8185', 'i2005', '1513', 'int', 'tech', '9p', '11q', 'reliability', '957', '9q', '7p11', '8p23', 'cub', '3q23', '7q2 1', '7q35', '16q23', '7p', 'cmyc', 'mmp1', 'birc2', 'goodness', 'birc3', 'per oxisome', '4q', 'triploid', '4p', '16p', '18p', '20p', 'britain', 'sons', 'an euploid', '12p11', 'q23', '20q12', 'stk15', 'picking', 'hyperdiploid', 'igrov', 'caov', 'jama', 'skov', 'ov90', 'mdah2774', 'ovcar', 'hypodiploid', 'pac s', 'cosmids', 'fluorochromes', 'paralogues', 'rejecting', 'correcting', '2s d', 'smoothing', 'kmeans', 'polyploid', 'illustrative', 'wilk', '15q21', '17p 12', 'golden', 'rp5', '2q37', 'rp1', '17q25', '3p22', '4q31', '6q25', 'rp3', 'ctb', '3q21', '3p14', '3q13', '4q24', '5q23', '6p22', '7p22', '9q21', 'harb ours', 'xp11', 'contradistinction', '9p24', 'p24', '9q33', 'q33', '15q26', '1 p34', '1p22', '5q31', '6p12', '7q31', 'q31', '8q23', '11q12', '16p12', '20p1 2', '1p32', '1p13', '7q11', '9p23', '10q11', '11q24', 'q25', 'impression', 'fgf3', 'fgf4', 'ww', 'wwox', 'prudent', 'underpinning', 'kallioniemi', 'intake', 'toxins', 'cirrhosis', 'aguilar', 'murakami', 'retransplantation', 'hepat oblasts', 'bipotential', 'feeder', 'zender', 'explanted', 'intrahepatic', 'alkaloid', 'hccs', 'pseudoglandular', 'fetoprotein', 'orthotopically', 'manipul ated', '2mb', 'representational', 'egan', 'alexander', 'peng', 'mmps', 'yap', 'ciap2', 'hepatocarcinogenesis', 'oncogenomics', 'bashyam', 'mmp7', 'mmp13', 'mmp10', 'silke', 'cointroduced', 'synonyms', 'yap1', 'tead', 'lats', 'sh1', 'sensitizes', 'dampening', 'nonapoptotic', 'wts', 'hippo', 'merlin', 'hamarat oglu', 'lats2', 'mcclatchey', 'giovannini', 'lats1', 'haber', '20q13', 'znf21 7', 'immortalizing', 'prefoldin', 'reexamination', 'nedd9', 'garber', 'depinh o', 'rdi', 'afp', 'immunolabeling', '85k', 'freezes', 'rnaeasy', 'taz', 'rege nerative', 'mst2', 'sav1', 'gpqr', 'homeodomain', 'lgl', 'larvae', 'ankyrin', 'kh', 'ptpn14', 'scribble', 'sik', 'zo', 'zona', 'occludens', '426', 'tbx5', 'adherens', 'par1', 'gpcrs', 'lysophosphatidic', 'sphingosine', 'appreciate d', 'ajuba', 'kinases2', 'follicle', 'overgrown', 'fruitflies', 'fruitfly', 'cuticle', 'biologists', 'cancers7', 'cancer7', 'mfc10a', 'mammospheres', 'm ammosphere', 'cases137', 'ctgf', 'events7', 'tumours15', 'cancers18', 'hepato megaly', 'dvl', 'paneth', 'unintended', 'dss', 'pathogens', 'enterocytes', 'cells141', 'facultative', 'myocardial', 'cardiomyocytes', 'regenerate', 'cardi omyocyte', 'proliferation147', 'manipulating', 'leveraged', 'aiding', 'fan', 'adrenaline', 'dopamine', 'botulinum', 'porphyrin', 'delaying', 'demethylase s', 'sirt1', 'cgmps', 'morphogen', 'cgnp', 'egl', 'igl', 'ruiz', 'reifenberge r', 'kenney', 'irs1', 'runx2', 'smad7', 'p53bp2', 'chick', 'eradicated', 'nor thcott', 'kool', 'smoothened', 'cyclopamine', 'isotype', 'berman', 'corcoran', 'wechsler', 'reya', 'gage', 'gli1', 'gli2', 'buttitta', 'smo', 'n2a', 'co activation', 'patched', 'newman', 'cd15', 'resume', 'neurobiology', 'camarg o', 'excitingly', 'proliferative', 'jag1', 'dll1', 'notch4', 'telomerase', 'mao', 'garnis', 'roth', 'andersson', 'shimizu', 'albertson', 'ingham', 'hans son', 'dendrogram', 'log2ratio', 'wolff', '201k', 'maroon', 'mouth', 'tongue', '2q14', '9p13', '11p13', 'underneath', 'lettering', 'opposition', 'morpho

gens', 'ihh', 'desert', 'dhh', 'ptch', 'gli3', 'taipale', 'enters', 'weng', 'aster', 'hes1', 'continually', 'outermost', 'incubations', 'fluorometry', 'printed', 'pinkel', 'reject', 'permutation', 'demand', 'gusb', 'ccttactgcct cttgcttc', 'neuroradiologic', 'gadolinium', 'autosomes', '12q14', 'pter', '10q25', 'esc', '7p15', '2p16', '10q21', '15q22', 'dp1', 'tfdp1', 'e2f1', 'pisan i', 'habits', 'escs', 'mcc', 'meltzer', '18q21', 'eppert', 'bunkyo', 'ims', 'bicarbonate', 'fluorochrome', 'highlevel', 'dig', 'a32p', 'prehybridized', 'overviews', 'xq', '7q34', 'eightfold', 'scgs', 'balsara', 'land', 'ernst', 'tgfb', 'rectangles', 'netic', 'q27', 'converges', 'htr', 'tanner', 'cng', 'btak', 'bcl1', 'protooncogenes', 'histiocytomas', 'merkel', 'helin', 'chell appan', 'thangue', 'erythroleukemic', '7q36', '22q13', 'databank', 'brb', 'superior', 'biocarta', 'hppt', 'qiaprep', 'dapk1', '1q31', 'twist1', 'deacetylase', 'hdac9', 'supra', 'genechips', 'ependymal', 'hey2', 'lfng', 'igf2', 'itgb1', 'cdc2', 'tcf', 'mlf1', 'thy1', 'cd90', 'net1', 'suffers', 'sulfotransferase', 'estrogens', 'therapeutical', 'relapses', 'tumor1', 'tenascin', 'ospedale', 'torino', 'clinica', 'pediatrica', 'trimming', 'radboud', 'nijmeg en', '1261', 'genepix', 'lowess', 'hggateway', '22q12', 'simon', 'sul', 'enz o', 'fuller', 'predeveloped', 'altschul', '1394', 'substantia', 'nigra', 'entre', 'db', 'neuropathol', '2938', '4912', 'powergene', 'msp', 'nat', 'nicholson', 'ross', 'grill', 'pang', 'jc', 'hui', '2133', '2141', 'reardon', 'heritage', 'motile', 'cilia', 'kawano', 'nagashima', 'acta', 'merchant', 'cd24', 'ornithine', 'decarboxylase', 'odc1', 'kleihues', 'bw', 'nrnxn2', 'dynamin', 'snrpn', 'chgb', 'synuclein', 'protocadherin', 'ribosylation', 'glutamyl', 'homeo', 'korshunov', '1727', 'clu', 'dra', 'rj', 'bentley', 'cdkn2d', 'overexpressions', 'map3k5', 'chibon', 'leiomyosarcomas', 'lms', 'rot', 'vassilev', 'modena', 'yokoyama', 'fernandez', 'law', 'nationale', 'sante', 'respecte d', 'fnclcc', 'coindre', 'uds', 'wycombe', '946', 'massy', 'orsay', 'swabs', '8400', 'kingston', 'orifice', 'myxofibrosarcoma', 'thigh', 'scapula', '00003', 'probesets', '276', 'probeset', 'chr3', 'hypergeometric', 'permeable', 'epidemic', 'gynecology', 'hysterectomy', 'federation', 'obstetrics', 'figo', 'zhong', 'microtome', 'affixed', 'cooker', 'tetrahydrochloride', 'doppler', '032', '041', 'significances', 'itgb2', 'sfb', 'histotypes', 'cytoreductive', 'mps', 'peter', 'maccallum', 'u133a', 'ctct', 'ca125', 'pvalue', 'ccp', 'bootstrap', 'internationale', '257', 'mycl1', 'radioresistant', 'radiosensitive', 'cancerrelated', '5803', 'xiap', 'ski', 'style', 'punch', 'iva', 'ivb', 'accelerator', 'portals', 'boosts', 'intracavitory', 'sessions', '241d13', 'nontumorous', 'myometrial', 'icn', 'logarithmically', 'dewaxed', 'foamy', 'observer', '4862', '0031', 'yrs', '4863', '6p21', '0021', '0055', '0026', 'survivin', 'redistribute', 'edmonson', '081', 'hbv', 'etiology', 'hpc', 'engrafted', 'thread', 'hepatectomy', 'queen', 'mary', 'pokfulam', 'described21', 'calif', 'washings', 'ethylenediamine', 'tetraacetic', 'electrotransferred', 'categoric', '056', 'parenchymal', 'ligandindependent', 'allergy', 'immunology', 'groningen', '3969', 'scholars', 'cem', '33p', 'described19', 'p815', 'mastocytoma', 'butterfield', 'miami', 'paramagnetic', '077', 'serumfree', 'pd173955', '1pkg', '1m52', 'prefers', 'cladribine', 'temporarily', 'cytoreduction', 'fraught', 'incurable', 'f522c', 'mutation24', 'promiscuity', 'chemotype', 'radiologic', 'csti571b', '2222', 'nonassessable', 'criteria24', 'radiology', 'transaminases', 'unsuitable', 'reversibly', 'deli843', '687', '0067', 'withheld', 'margaret', 'mehren', 'heikki', 'joensuu', 'sandra', 'silberman', 'dimitrijevic', 'deate', 'kiese', 'burton', 'eisenberg', '0022', '0065', '0866', '0558', '0289', 'dimh', 'hinged', 'sinonasal', 'k1', '07e', 'walt her', 'allergic', 'axel', 'martinsried', 'plxsn', 'm5gne', 'tyr568', 'tyr703', 'antiphosphothreonine', 'antiphosphoserine', 'py20', 'aebsf', 'leaflet', 'marseilles', 'annex', 'hmc1', 'interchange', '07ee', 'mapk2', 'lombardo', 'disconnect', 'cddp', 'abutting', 'p1241', 'q56p', 'bypassing', 'diarylamine', 'masse', 'mutatgenesis', '3eqc', 'arylamine', 'i99t', 'f1291', 'v211d', '1215p', 'i103n', 'k104n', 'i111n', 'h119p', 'e120d', 'p124', 'p124s', 'p124

q', 't55p', 'd67n', 'g328r', 'e368k', 'pertinent', 'postrelapse', 'm307', 'wm3482', 'wm3457', 'wm3506', 'measureable', 'rafv600e', 'elaboration', 'expressly', 'interfacing', 'scp', 'database9', 'noninherited', 'landesbioscience', 'chimpanzee', 'cow', 'clawed', 'finger30', 'diseaseassociated', 'serendipitous', '1293', '1559', 'vallejo', 'al46', 'actseq2', 'ctagctgggtggcatatg', 'actseq3', 'cggggttttcagtatctacg', 'padbrca1', 'pact2', 'manufactuer', 'untargeted', 'l40', 'his3d200', 'lexaop', 'procedures48', 'restreaking', 'streak', 'pipettor', '1651', 'lk1600', '01ee', 's1613c', 'previouslytested', 'v1653m', '1783', 'disagreements', 'mirkovic', 'd1739e', '1739', 't1720i', 'v1804a', 'v1804d49', 'sectors', 'clockwise', 'k1601e', 'm1783', 'exercised', 'l1600e', 'k1606e', 'a1608v', 'a1615d', 'm1628v', 'a164it', 'intrahelical', 'v1810g', 'n1819y', '457', 'carbohydrate', 'sialic', 'mannose', 'matthijs', 'isomerase', '1536', 'candida', 'albicans', 'crt', '413', 'lebanese', 'admitted', 'apnea', 'stool', 'pseudomonas', 'aeruginosa', 'readmitted', 'hypoalbuminemia', 'oxaloacetate', 'transaminase', 'iga', 'prothrombin', 'xi', 'proteinuria', 'uenta', '1537', 'substance', 'sweat', 'endoscopy', 'villus', 'stellate', 'naso gastric', 'collateral', 'angiomas', 'gastrostomy', 'multisystem', 'oligosaccharide', 'reyes', 'alimentary', 'glut2', 'isthat', 'treatable', '1538', 'symptomatology', 'quebec', 'pelletier', 'glucosyltransferase', 'bacteriophage', '533', 'figueroa', 'heatmaps', 'substantively', 'nonamplified', 'styi', '2p13', '2og', 'dioxygenases', 'genes7', 'apobec', 'reprogrammed', 'nhl', 'lymphocytosis', '1785', 'nssnvs', 'genes29', 'gna13', 'myd88', 'dlcbl', 'hist1h1e', 'epigenetically', 'lymphopoiesis', 'igd', 'ebf1', 'cd79a', 'hydroxymethylation', 'h3k4me2', 'h3k27ac', 'ingenuity', 'ipa', 'gna14', 'renewing', 'mcl', 'nhls', 'compounded', 'demethylating', 'aza', 'deoxycytidine', 'infusions', 'dj', 'reported28', 'plotting', 'sectioning', '6059', 'epitomics', 'hpa1i', 'amplifiable', 'mspi', 'femur', 'potassium', 'ack', 'rb6', '8c5', 'cd71', 'cd8', 'resuspension', 'aria', 'treestar', 'mm9', 'mm10', 'burrows', 'wheeler', 'aligner', 'realignment', 'recalibration', 'toolkit', 'callers', 'flushed', 'nucleated', 'unfractionated', 'm3434', 'flt3l', 'ovation', 'nugen', 'genespring', 'gx', 'polya', 'cuffdiff', 'mev', 'ncounter', 'tss', 'ccgg', 'cohesive', 'errbs', 'hiseq2500', 'ser139', 'igk', 'blinding', 'vexing', 'ssa', 'nonpathogenic', 'sweet', 'quandary', 'distress', 'dorval', 'petrucelli', 'dijk', '567', 'cimba', 'lindor', 'plon', 'reclassification', 'millot', 'carvalho', 'cotta', 'ramusino', 'kais', 'ransburgh', 'vallee', 'jasin', 'parvin', 'duarte', 'bennard o', 'newyork', 'nonhomologous', 'rejoined', 'moynahan', 'diagrammed', 'cotransfect', 'transfект', 'timeline', 'depleting', 'functioned', '0014', 'nondepleted', 'r170q', 'v191d', 's153r', '1775', 'expresser', 'benchmarking', 'c39y', 'vega', 'benchmarks', 'y179c', 'sake', 'prespliced', 'misclassify', 'centrosomal', 'experiences', 'ruffner', 'discordancies', 'centrosomes', 'ku80', 'beass', 'a775insyvma', 'g776vc', 'g776lc', 'p780insgsp', 'g660d', 'h2170', 'adi', 'porton', 'sykkinase', 'pidt', 'pcmvir', 'copying', '3rcd', '4g5j', 'stride', 'tyr1221', 'asp214', 'fujifilm', 'cycletest', 'facscaliber', 'yokohama', 'dunnett', 'a775ins', 'm774', 'r784', 'reports19', '471', 'administer', 'perera', 'concentrations38', 'g776c', 'v777inscg', '1915', 'takezawa', 'blimp', '1970', 'hydrochloric', 'thiocyanate', 'oven', 'buffalo', 'vm', 'cep6', 'tonsils', 'pleasanton', '407', '30b', 'negativity', '074', '0095', 'cd30', '0062', '0028', 'mum1', 'degs', 'hlf', 'xbp1', 'plasmacytic', 'transportation', 'batf', 'cd22', 'bcl11a', 'blk', 'apobec3b', 'vdr', 'prf1', 'lancl2', 'irf1', 'rbpj', 'ikzf1', 'mandelbaum', 'jointly', 'gadd45a', 'bf1', 'riz', 'ly3', 'editing', 'conversions', 'bcell', '4100', 'prdm', 'kruppel', 'groucho', 'deacetylates', 'prdm2', 'orchestrates', 'riz1', 'nebraska', 'wayne', 'na2', '1490', 'e4', 'dnasei', 'twentieth', 'bioresource', 'ithaca', 'otsu', 'amido', 'uga', 'znf', 'aataaa', 'nontranscribed', '057', '362', 'zf1', '4097', 'g3a', 'hdacs', 'prdm5', 'primordial', 'hypermutations', 'apob', 'lipoprotein', 'deaminase', 'ptpn6', 'tsg', 'centrocytes', 'plasmacytoid', 'digests', 'jem', 'dc1', 'ly10', 'polypeptides', 'paragraph', 'triplication', 'irf', 't50i', 'g60e',

'i24n', '151100', 'pandit', 'shoc2', 'cordeddu', 'razzaque', 'oliveira', 'an astasaki', 'jopling', 'epiboly', 'gastrulation', '71t', 'webbing', 'pectus', 'excavatum', 'cryptorchism', '1ab', 'heterozygously', '1ad', 'gln22', 'burg t', 'ile24', 'repulsing', 'organismal', 'fertilization', 'hpf', 'lapse', 'axe s', 'su5402', 'solnica', 'krezel', 'krox20', 'rhombomeres', 'somites', 'myo d', 'dlx3', 'hgg1', 'den', 'hertog', 'eekelen', 'tracking', 'paraxial', 'mese ndodermal', 'intercalation', 'mediolateral', 'concha', 'adams', 'cartilage', '9sox10', 'ba2', 'carney', 'ceratohyal', 'bluntness', 'laws', 'ensg000002132 81', 'cgccaattaaccctgattac', 'agagac', 'aggatcaggtcagc', 'atagcattgca', 'ttcc ctgtg', 'cacaaagatcatcctt', 'tcagag', 'ccactgtacccagcctaatac', 'aaactcttgaccaa atgctg', 'cctccaaattgcccataac', 'atttg', 'gattgtgtccgtt', 'rooij', 'rasophat hies', 'westerfield', 'liz', 'capped', 'mmassage', 'mmachine', 'inadvertent', 'thisse', 'dechorionated', 'rsbweb', 'ij', 'dpf', 'sits', 'tells', 'fortunate ly', 'medlineplus', 'a172', 'sf268', 'sf295', 'u118', '50k', 'xba', 'hind', 'wuxi', 'pharmatech', 'scbt', 'appliedbiosystems', 'spectrodesigner', 'spect rotyper', 'robot', 'illinois', 'cavenee', 'postplating', 'r108', 'e330k', 'r3 241', 't263p', 'a289v', 'g598v', 'compartmentalization', 'sublines', 'reacp5 3', 'organoids', 'hgsoc', 'shrinks', 'hgsocs', 'vazquez', 'vousden', 'joerge r', 'seidman', 'dies', 'vaughan', 'kurman', 'leitao', 'silva', 'levy', 'jucke r', 'rescuer', 'zipperdb', 'goldschmidt', 'ghosh', 'ltiitle', 'tiitle', 'micr ocrystals', 'sawaya', 'zippers', 'p53252', 'sievers', 'lendel', 'leaver', 'fa y', 'ltrittle', 'substoichiometric', 'i254r', 'fuchs', 'raines', 'rpi', '20h', 'puncta', 'godl', 'janzen', 'ovcar3', 'enlargements', 'yo', '16h', 'qv', 'op h', 'nec', 'necroptosis', 'phopsho', 's608', 's3i', 's3j', 'utfib', 'rance', '3l', 'utfibs', 'aggregating', 's3o', 'untrasfected', 'insoluble', 'cavities', 'pili', 'envelopes', 'detroit562', 's4k', 'rnaseq', '2400', 's5e', 'gadd4 5b', 'noxa', 'dram1', 'gluconeogenesis', 'mevalonate', '45h', 'shrank', 'doub led', 'pank', 's6i', 's6g', 'piknotic', 'grafts', 'implants', 'quadruplicate s', 'intraperitoneal', 's6m', 'bullock', 'bykov', 'issaeva', 'wilcken', 'apr', 'lehmann', 'khoo', 'oncosuppressive', 'ventura', 'hashimoto', 'srebp', 'alend ronate', 'lovastatin', 'fluvastatin', 'zoledronic', 'martirosyan', 'kruiswij k', 'vossio', 'rufini', 'topkhane', 'naora', 'entails', 'siddik', 'lasagna', 'reeves', 'rieber', 'strasberg', 'dilemma', 'fuses', 'fourths', 'pdr', 'cba s', 'gcuccucucacucuucagu', '798', 'ay273801', 'uaauuuuggacauaaggaguccucc', 'dr13', 'osu', 'fluorescing', '1292', '1527', 'c27a', 'decrements', 'i31m', 'as terisked', 't37k', 'radio', 'jdp', 'dsbr', 'scully', '1997c', 'sharan', 'impe ded', 'tomlinson', 'nbs1', 'frameshifted', 'lightly', 'stippled', 'ms110', 'sg11', 'sd118', 'cortez', 'mb435s', 'shao', 'uncloned', 'ecor', 'hua', 'cocult ure', 'cocultivated', 'mabs', 'influenza', 'untagged', 'gowen', 'chapman', 'verma', 'postirradiation', 'acl4', 'apai', 'klenow', 'mfei', 'religated', 'mc g', 'cocultured', 'resorted', 'vantage', '137cs', 'badie', 'kiltie', 'plugs', 'oxaliplatin', 'stereotactic', 'radiosurgery', 'gustave', 'roussy', 'edouard', '94805', 'villejuif', 'invest', 'deteriorated', 'lymphangitic', 'carcinomatosis', 'limb', 'blurred', 'lasting', 'monovalent', 'endpoint', 'dozen', 'pa ediatric', 'schedule', 'stratum', '2q', 'amendments', 'enrolment', 'haemorrhage', 'karnofsky', 'aminotransferase', 'uln', 'qtc', 'lactation', 'guardians', 'assent', 'minimise', 'capsule', 'keith', 'dlt', 'stopping', 'mineral', 'electrolyte', '12h', 'trapezoidal', 'mibg', 'quantitate', 'fluorodeoxyglucose', 'cru', 'sums', 'escalated', 'cv', 'multiagent', 'vinblastine', 'deltoid', '123i', 'aspirates', 'accruing', 'prospect', 'v600d', 'brafv600r', 'brafv600d', 'rg7204', 'az628', '0879', 'm1a', 'm1c', 'imvi', 'mkn74', 'microtiter', 'xlfit', 'dilute', 'schering', 'plough', 'ellipsoid', 't0', 'predefined', 'jande l', 'brafk601e', 'nonresponsive', 'rasq61k', 'corn', 'inhibitions', 'postdos e', 'peaking', '740', '2986', '1368', 'postimplantation', 'amorphous', 'stasis', 'escalations', 'members1', 'tumours1', 'epistatic', 'dependencies', 'kins se4', 'mek5', 'skmel103', 'skmel31', '59k', 'nci60', 'lines8', 'partitioned', 'analysis9', 'hypothemycin', 'resorcylic', 'cleavage12', 'colourgrams', 'fon

t', '104k', 'function8', 'surpassed', 'complex14', 'cytoidal', '39k', 'capitated', 'suppression15', 'skmel30', 'advantage1', 'activity16', 'fashion18', 'expression24', 'activation26', 'skmel', 'accumed', 'phenylmethylsulphonylfluoride', 'caging', 'humane', 'euthanasia', 'calliper', 'paraformaldahyde', '796', 'dtp', 'matlab', 'ncs', 'nan', 'data1', 'group28', 't1796a', 'colo', 'c1786g', '1596v', 'h2087', 'g1403c', 'h1395', 'ov18t', 'cancertable', 'stc', 'stcs', 'subcategory', 'histiocytoma', 'cancers4', 'proopiomelanocortin', 'melanocortin', 'melanogenesis', 'radiation8', 'erk9', 'ape', 'motifs10', 'v599d', 'e585', 'f594', 'g595', '1596', 'araf1', 'v599', 'g463', 'g465', 'g468', 'kinases11', 'orientate', 'catalysis10', 'activity12', '83k', 'substrates15', 'g468abraf', 'v599ebraf', 'g463vbraf', '1596vbraf', 'dually', '54k', 'molecules18', 'proposed2', 'mutantstable', 'braf10', 't598', 's601', 'braf17', 's598', 'explicable', 'k600', 'kras2', 'hx62', 'f5941', 'culture19', 'prompts', 'proliferationtable', 'types21', 'kinase15', 's445', 'd448', 'leukaemia22', 'cumbría', 'nccgp', 'analyser23', 'previously24', 'microm', 'yt', 'd4', 'pefm6', 'microg', 'experimenters', 'described15', 'v599eb', 'cobb', 'rapp', 'avruch', 'cutler', 'mason', 'rajagopalan', 'singer', 'brose', 'thr598', 'ser601', 'val599', 'yuen', 'wtb', 'lyons', 'subphysiological', 'e585kb', 'v599db', 'g468eb', '1596vb', 'cancerthe', '646kb', 'g465ab', 'n580sb', 'd593vb', 'kdb', 'lys482', '372kb', 'uninjected', 'g465eb', 'g465vb', 'g595rb', 'wtc', 'kdc', 'welbrook', 'client', 'p50cdc37', 'blagosklonny', 'stancato', 'michaud', 'macnicol', 'gln432', 'lys600', 'gln611', 'bilobal', 'interfacial', 'lowinger', 'pyridyl', 'trp530', 'phe582', 'phe594', 'lipophilic', 'leu513', 'thr528', 'glu500', 'asp593', 'birb', 'nagar', 'pargellis', 'schindler', 'cys531', 'ribbons', 'asn580', 'glu585', 'marine', '596kb', 'stereoview', 'pd173855', 'disposition', 'orients', 'gly595', 'gly463', 'val1470', 'leu596', 'gly465', 'phe467', 'arg574', 'asp575', '384kb', 'asp293', 'lys298', 'glu299', 'ile301', 'phosphothreonine', 'asp275', 'arg274', 'glu599', 'glu600', 'glu300', 'lys569', 'gly300', 'thr309', 'arg575', 'lys506', 'f467cb', 'g468ab', 'gly468', 'wedged', 'pthr598', 'asp448', 'asp447', 'arg505', 'concluding', 'remarks', 'kitayama', 'analogies', 'k482mb', '176kb', 'rkip', 'farrar', 'coregulatory', 'retinoids', 'nrs', 'afs', 'af1', 'af2', 'corepressors', 'lxxll', 'holocomplexes', 'kazusa', 'pbk', '2tk', 'blunted', 'pmmtv', 'pap', 'e19', '5372', 'coimmunoprecipitations', 'cationic', 'plyss', 'bc100', 'mmdithiothreitol', 'lumat', 'andf', 'mmp', 'henylmethylsulfonyl', 'swollen', 'dialysis', 'mmtris', 'deoxycholic', 'eluates', 'mmmgcl2', 'proteins2', 'isolating', 'andd', 'probasin', 'cdk8', 'med12', 'hdm2', 'nutlins', 'dr5', 'med13', 'credence', 'specialization', 'expectedly', 'phosphoacceptor', 'btg2', 'sjsa', 'espinosa', 'paused', 'ser5', 'ser2', 'gtfs', 'tfiia', 'tfiib', 'oscillates', 'tfiif', 'preinitiation', 'pic', 'co-regulator', 'reproducing', 'rko', '4001', '2965', 'desire', 'hydroxyurea', 'elongate', 'reinitiation', 'gtf', 'med23', 'cdk9', 'drb', 'antimycotic', 'hoffman', 'pgl2', 'leiomyomas', 'fibroadenomas', 'med24', 'finnish', 'xq13', 'interactome', 'rex', 'alanyl', 'pog44', 'protean', 'ab9485', 'brightness', 'tactin', 'iba', 'collision', 'cid', 'proteome', 'thermofisher', 'ppm', 'pcs2', 'p3xflag', 'aarhus', 'consents', 'director', 'district', 'hus', 'tampere', 'fch', 'dg', 'purifications', 'turunen', 'bore', 'pccl3', 'thyrotropin', 'tsh', 'ngan', 'sar', 'notes', 'lang', 'stacey', 'yau', '1q22', 'atrophic', 'nonpalpable', 'encapsulated', 'microcarcinoma', 'caveats', 'checkups', 'anticoagulant', 'xed', 'verbal', 'mgcl', 'coon', 'independency', 'minoo', 'prc', 'gibbstown', 'phenylmethanesulfonyl', 'ser727', 'mastermix', 'akira', 'wildbad', 'supershift', '2027', 'twosided', 'perineural', 'laryngeal', 'gbq', 'radioiodine', 'swelling', 'pt1', 'multifocality', 'remnant', '924', '826', '20s', 't1n0m0', 'bilateral', 'daughters', 'nonconsanguineous', 'nontoxic', 'outpatient', '1960s', 'iodine', 'locoregional', 'ignored', 'immunosorbent', 'quantifi', 'quanti', 'densitometer', 'transactivational', 'inheritable', 'c363', 'supershifted', 'claim', 'scientists', 'tumori', 'zonula', 'estab', 'carolina', 'chapel', 'gested', 'rylation', 'aj', 'epi', 'faulty', 'cadherins', 'ox', 'c'

atenins', 'quence', 'plakoglobin', 'desmosomes', 'mysterious', '1667', 'sat', 'wingless', 'cy', 'hy', 'nin', 'tein', 'profuse', 'interfamilial', 'mutat', 'colonoscopy', 'epidermoid', 'desmoid', '843', '1061', '1309', 'humu', 'inwiley', 'interscience', 'bodmer', 'gastroenterological', 'i1307k', '951c', '721c', '15b', '601c', 'cosmid', 'digoxigenin', 'olschwang', 'dunnen', '1790', 'finished', 'myh', 'tassan', 'dgge', '2p12', '2p11', 'elder', 'excludes', 'theapc', 'eurogentec', 'lindgren', 'barber', '1526', 'andreutti', 'zaugg', 'rosa', '3736', 'bj', 'hutter', 'mcr', 'genders', 'tubular', 'rim', 'pouch', 'constipation', 'polypectomy', 'obstructing', 'antrum', 'fundus', 'gastritis', 'pylori', 'iranian', 'tubulovillous', 'rivera', 'nonexistent', 'submicroscopic', 'naples', 'fluorophores', 'bulletin', 'semiautomated', 'table3', 'amounted', '1580', 'oncoseq', 'taxonomy', 'traction', 'tacc1', 'tacc3', 'tacc', 'enrolls', 'deliberate', '1036', 'q1573', 'c736', 'bicc1', '1039', 'r267w', '1051', 'aff3', 'g199e', '1081', 'slc45a3', 'glinsky', '2053', 'kiaa1967', 'ofd1', 'baiap2l1', 'bag4', 'erlin2', 'lis1', 'sphf', 'colled', 'cit', 'seo', 'rounding', 'hme', 'j82', '1197', 'k652e', 'pazopanib', 'rt4', 'disciplinary', 'deliberation', 'poles', '99a', 'persistently', 'crizotinib', 'hum00046018', 'qiazol', 'tissuelyser', 'truseq', 'mag', 'indexed', 'qiaex', 'chastity', 'basecall', 'dynal', 'fastqc', 'boxplot', 'q20', 'nomination', 'eland', 'mate', 'nominations', 'realigned', 'reconfirm', 'hugo', 'blosum', 'dbsnp135', 'v5668', 'flexmix', 'scatterplot', '90th', 'ht1197', 'incucyte', 'pcdh510b', 'fugen6', 'dynabeads', 'scientific', 'dharmafect1', 'targetplus', '003133', '018664', 'b17', 'kenneth', 'pienta', 'xylazine', 'ketamine', '1xpbs', 'ucuca', 'laid', 'therapy1', 'targeting2', 'patients5', 'response9', 'leukemias10', 'lymphomas11', 'mf12', 'paved', 'ruxolitinib', 'trials18', 'selectivity3', 'streamlining', 'oncogene22', 'v61724', 'treatments24', 'glean', 'fedratinib', 'azd1480', 'cyt', 'bypasses', 'domains25', 'domains26', 'model27', '2j0j', '2j0l', '2j0m', '2src', '1opj', '1opl', '2g2h', '2g2i', 'patients28', '1opk', '4fvr', 'd976n', 'y931c', 'r971g', '1983f', 'jak230', 'crosslinkers', 'm181r', 'e1006k', 'e1046k', 'n1063p', 'abl28', 'whitlock', 'witte', 'r683t', 'y931f', 'y44c', 'n99s', 'h110r', 'r277k', 'v366g', 'g417p', 'm483r', 'mutations33', 'g652e', 'l511p', 'tf556ps', 'v582a', 'e596qs', 'l680p', 'e684y', 'n731', 'l551p', 'e684q', 'e596q', 'n731y', 'autodimerization26', 'jh126', 'l902m', 'e1080k', 'd840n', 'n924t', '1927f', 'q853p', 'n981s', 'l1001w', 'y966e', 'n986y', 'e987d', 'swissdock35', '911', 'leu932', '933', 'e930g', '1932f', 'v9111', 'g935r', 'p933a', 'e985k', 'n981i', 'd994n', '931', 'loops36', 'conformation73', 'residues37', 'azd1840', 'resistance38', 'semilogarithmic', 'tg10138', 'resistance39', '12d', '12e', 'autodock', '940', '3tjd', '16a', 'fluctuate', '931showing', 'incb01842', 'resins', 'app', 'd544h', 'reported41', '16e', 'microscale', 'thermophoresis', '8307', '804', 'i1018f', 'l1026f', 'f1019l', 'w1020c', 'y1021h', '8k', 's1025c', 's1039f', 'g1041r', 'v1042f', 'i1074f', 'v1075f', 'ifwy', 'w1038c', 'y1045w', 'p1058f', 'f1061w', 'i1079l', 'l1082r', 'sink', '1018', '1026', 'sulfonamide', 'i1018w', '1045', 'leu1086', 'phe1061', '1058', '1019', 'val1075', 'weekend', 'bioluminescence', 'it42', 'dependence39', 'kinases39', 'immunoblotting', 'pdgfr32', 'erloitinb', 'socs47', 'interaction33', 'counterselection', 'ono1238048', 'ono12380', 'cytokines18', 'used50', 'cmax', 'kg51', 'dasatinib52', 'ap2416353', 'ap2453454', '203655', 'gnf2', 'gnf556', 'variants56', 'formidable', 'targeting53', 'counterselective', 'g469a', 'k601n', 'plasmacytomas', 'pct', 'opm2', 'kms11', 'h929', 'plasmacytoma', 'cryopreserved', '2044', 'rajakulendran', 'avraham', 'dougherty', 'dusp', 'pratillas', 'lito', 'montero', 'conde', 'paik', 'aminoterminal', 'berghof', 'preusser', 'brafs', 'pcraf', '1597v', '1597r', 'inductions', '1597q', 'g469v', 'g464v', 'g466e', 'd287h', 'heidorn', 'r1881', 'drosten', 'poulikakos', 'r509h', 'r401h', '401h', 'k601', 'esrp1', 'r401a', 'singular', 'nazaria', 'hatzivassiliou', 'mel239', 'inhibiter', 'lgx818', 'bgb659', 'vemurafenib', 'hyperactivates', '1hr', 'kia1549', 'jvm', 'maxima', 'lavoie', 'lgx', 'lanes7', 'coexpress', 'unassociated', 'ought', 'chuvas', 'elongins', 'polyubiquitin'

uitylates', 'hre', 'anaerobic', 'respectively5', 'haemangioblastoma', 'ang', '598c', 'russian', 'rbcs', 'categorised', 'paraneoplastic', 'phd2', 'ecs', 'polycythemic', 'c162f', 'fibrillar', 'polyubiquitylation', 'cish', 'ubiquitylating', 'f119s', 'l128f', 'autonomy', 'ca9', 'igfbp2', 'ryk', 'cdc53', 'soc s2', 'heterocomplex', 'hinders', 'shvhl', 'plantar', 'footpads', 'vbc', 'over production', 'receptor35', 'thrombotic', 'bosutinib', 'transferring', 'redirected', 'cope', 'mnc', 'lymphoprep', 'oslo', 'transcriotor', '486', 'wales', 'jm109', 'biopolymer', 'accelrys', 'enact', '4th', 'tackle', 'justifies', 'b af180', 'pbaf', 'hsnf5', 'ini1', 'baf47', 'brg1', 'baf200', 'hcc1187', 'sum5 2', 'sum190', 'ethier', 'schwartzentruber', '5ug', '10ng', 'bx', '1475', 'bg 1', 'lymphoblast', 'subtractive', 'mung', 'bean', 'misonix', 'salmon', '8r', '7786', 'hoescht', '2261', 'lsrii', '74104', 'mx3000p', 'mxa', 'pires', 'sic ontrol', 'signalsilence', 'bx41', 'reintroduced', 'assists', 'fx1', 'pollack', 'varela', 'xia', 'bromodomains', 'bds', 'rsc', 'solomon', 'mescs', 'subtelomeric', 'shcontrol', 'misregulation', 'fortuitously', 'm538', 'micronuclei', 'misseggregation', 'anaphase', 'kinetochores', 'xue', 'brough', 'mesc', 'cytocell', 'u20s', 'm538i', 'anaphases', 'indentified', 'differentiation4', 'smarcb1', 'tumours8', 'types10', 'pfam', '727', 'gall', 'panc', 'sleeping', 'beauty', 'described17', 'intraductal', 'panin', 'previously2', 'parsimonious', 'snu', 'arid5b', 'lrec', 'approvals', 'described1', 'polybromo', 'unveiled', 'embarked', 'a498', 'achn', 'lgc', 'rcc4', 'maxwell', 'thrich', 'beecher', 'prairie', 'mertz', 'ccrccs', 'cc1', 'a301', 'nanozoomer', 'shizuoka', 'a531 6', 'pt4', 'nonexpressing', 'pbmr1', 'development31', 'remodel', 'development 1', 'iswi', 'nurd', 'chd', 'remodels', 'nucleosome', 'bap', 'brahma', 'brm', 'swi2', 'snf5', 'baf155', 'baf170', 'arid1b', 'brd7', 'ejection', 'catalysing', 'baf57', 'hdac1', 'rts', 'schwannomatosis', 'cancers34', 'gene44', 'orthologue', 'death28', 'methylase', 'arid', 'adipogenesis', 'osteogenesis', 'haem atopoiesis', 'peri', 'oct4', 'nanog', 'ips', 'sugar', 'ecdysone', 'oestrogen', 'rock1', 'antagonistic', 'pcg', 'hox', 'cancer103', 'e2fs', 'assemblies', 'rxra', 'hh', 'contractility', 'bundles', 'cancer108', 'perspectives', 'jmjd2c', 'kdm4c', 'kdm4b', 'myst3', 'cdc25c', '25c', 'cdc6', 'licensing', 'hoxa1 0', 'ras2', 'threat', 'thr383', 'linux', 'txt', 'perl', 'dun1', 'clustal', 'udson', 'thr68', '2947', 'ser216', '2032', 'pretxo', 'stomatitis', 'extractor', 'gotaq', 'sasi', 'hs01', 'sic', 'il132', 'tq', 'fha', 'autophosphorylates', 'n145w', 'br3', 'degrades', 'liver5', 'cright', 'arrowt', '99k', 'assay1 9', 'proportionately', 'enolase', 'km12c', 'photograph', 'numb', 'v354e', 'rr m2', 'srsf1', 'ASF', 'secretase', 'cb17sc', 'dorsal', 'rg6', 'chrX', 'rbm5', 'rrm', 'rrms', 'ocre', 'outlook', 'qki', 'shc002', 'sic001', 'k758', 'lev', 'retrotranscribed', 'fw', 'p32', 'cm2', 'phosphorimager', 'repairing', 'rbm', 'maniatis', 'graveley', 'interplays', 'barash', 'witten', 'ule', 'padgett', 'blanco', 'licatalosi', 'mukherjee', 'yeo', 'rbp', 'rappsilber', 'presplieosomal', 'behzadnia', 'parologue', 'sutherland', 'bonnal', 'tarp', 'talipes', 'equinovarus', 'septal', 'robin', 'vena', 'imielinski', 'clip', 'oe', 'hafner', 'mappable', 'manifesting', '619', '6396', 'mortazavi', 'puf60', 'skipped', 'strengths', 'cdf', 'afflicted', 'lcls', 'joubert', 'heterogenous', 'ciliopathy', 'vermis', 'syndromic', 'hypogonadism', 'weiss', 'eldest', 'bioanalyses', 'koac', 'hexamer', 'sybrgreen', 'mms', '25778', 'ein', 'pir', 'iin', 'loess', 'h37', 'rbm6', 'angeloni', 'edamatsu', 'rintala', 'ramaswamy', 'wellington', 'mourtada', 'maarabouni', 'triplet', 'ranbp', 'anguaa', 'song', 'fushimi', 'deckert', 'iggs', 'genic', 'kaida', 'denser', 'nig', 'solexa', '079', 'rbms', 'mers', 'zing', 'rnacompete', 'heinz', 'ucauc', 'aguua', 'uc', 'cucugaa', 'cucu', 'repertoires', '1294', 'hjay', 'hnrrnp', 'huelga', 'tdp', 'tollerv', 'hnrrnpu', 'tia1', 'tial1', 'mbnl', 'hnrrnpa1', 'martinez', 'contreras', 'resemblances', 'darnell', 'llorian', 'misquitta', 'orengo', 'imr9', 'imr9 0', 'prrs', 'dfp', 'noncancer', '1062', 'westhoff', 'maraver', 'precedents', 'hey1', 'presenilin', 'prrl', 'aons', 'silencer', 'aon', 'iss2', 'nonphysiological', 'iss1', 'hyperubiquitination', 'srsf3', 'srsf4', 'fox2', 'fet', 'hoe

ll', 'valc', 'rcel', 'unlikelyto', 'pal', 'karni', 'esrp', 'warzecha', 'pec e', 'nicd', 'mcgill', 'mcglade', 'toriya', 'verdi', 'purow', 'smn2', 'smn', 'palate', 'oligonucleotide', 'rbmnter', '150v', 'paronetto', 'bargmann', 'gullick', 'myotubes', 'basilico', 'moscatelli', 'mansukhani', 'mansour', 'yamaguchi', 'werner', '1138', 'rousseau', 'muenke', 'schell', '1397', 'myoblast', 'prk', 'leder', 'anking', '1398', 'r3', '378c', '1399', '200ng', '125i', 'frankly', '1401', 'myotube', 'eld', '1402', 'uorescence', '1403', 'endochondral', 'osteoclasts', 'iwamoto', 'orderly', 'beare', '7350', 'fgf10', 'fgf16', 'fgf23', 'ibrahimi', 'mesenchymally', 'fgfr2c', 'proteoglycans', 'chondrodysplasia', 'passos', 'bueno', 'wilkie', 'celllines', 'epithelial', '075259', '000132', 'mismatchhad', 'a315t', 'a315s', 'histotype', 'violating', 'glandular', 'sangha', 'moller', 'fgf9', 's373c', 'y376c', 'disulphide', 's372c', 'y375c', 'cutis', 'gyrata', 'witha', 'fgfr3c', 'chondrohyperplasia', 'illegitimately', 'c383r', 'm392r', 'monsonego', 'ornan', 'k660e', 'ivs10', 'vt', 'burgar', 'gcaagt', 'burset', '1492', 'rhijn', 'obel', 'biologics', 'cytotoxics', 'rita', 'jhu', 'hn31', 'pharyngeal', 'pci', 'vitamins', 'g245d', 's19', 'shp53', 'radiosensitizer', 'chemoresistant', 'implements', 'hdaci', 'butyrate', 'nab', 'trichostatin', 'tsa', 'zbp', 'ppel', 'p53r273h', 'hep3b', 'viabilities', 'delocalization', 'recruitments', 'promotive', 'reactivated', 'c2h2', 'unlock', 'krasg12c', 'insusceptible', 'triphasphatases', 'ars853', 'cmpds', 'cmpd', 'puzzling', 'disable', 'h358', 'hrasa59g', 'y64a', 'y40a', 'n116h', 'y32s', 'augmentation', 's7f', 'h1972', 'h2030', 'h2122', 's7i', 'exits', 'mucus', 'innumerable', 'preinvasive', '004448', 'peritumoral', 'generalizability', 'q787r', 'attainable', 'transmits', 'leptomeningeal', 'e109', 'e120', '1860', 'genejuice', 'settle', 'y869', 'bracketed', 'y1197', 'y1110', 'antiactin', 'phenylmethylsulfonylfluoride', 'densitometrically', 'subsided', 'prestimation', '1086', 'corroborative', 'gxgx', '719gsgafg724', 'choong', 'e884', 'r958', '874', 'oriental', '21841502', 'deletional', 'borne', 'astral', 'zene ca', 'misclassifying', 'upto', 'hypoxanthine', 'deaminated', 'hiden', 'deoxinucleotide', 'yr', 'adk', 'pra', 'undiff', 'v769insvas', '771insvdsdnp', '1862v', 'sdb', 'station', '1831r', 'g709a', 'i747', '755insat', 'k746', 'h8351', '709a', 'monochemotherapy', 'hijiya', 'acetabular', 'carboplatinum', 'okami', 'ilie', 'pn3', 'cisplatinum', 'tsao', 'ludovini', 'v786m', 'd770inssvq', 'n771insg', 'n771insd', 'r831c', 'data28', 'omo', 'lobar', 'vinflunine', 'knife', '858r', 'pretty', 'tyrosin', 'prognosticators', 'tumors24', 'auspices', 'chromogens', 'premixed', 'definable', 'splitting', 'apple', '29k', 'hbl100', 'fibrocystic', 'fibroadenoma', 'subdivision', 'lobules', 'x400', '16k', 'timepoint', 'brighter', 'characters', 'cytological', 'inf', 'immunoreaction', 'schmidt', 'hprc', 'oncogenetic', 'cytomorphologic', 'greco', 'renzo', 'rosen', 'ponentzetto', 'pulverized', 'braun', 'aylesbury', 'embl', 'dipartimento', 'oncologista', 'menopausal', 'iic', 'gb', 'etal', 'cystadenoma', 'teratomas', 'gtl16', 'perez', 'mesothelium', '16th', 'ceases', '661', 'mullerian', 'overtly', 'rice', 'mok', 'giancotti', 'ruoslahti', '1991a', 'bottaro', 'comoglio', 'birchmeier', 'olivero', 'liang', 'ueki', 'kinoshita', 'noguchi', 'clonetics', 'providers', 'tre', 'berlin', 'loosened', 'nh4cl', 'hybridoma', 'lap', 'herman', 'bruce', 'perfusion', 'retrograde', 'gauze', 'apoptag', 'bishop', '1991b', 'autopsies', 'lobule', 'ferrell', 'disorganization', 'trabeculae', 'tumorous', 'reappeared', 'calcification', 'scar', 'okuda', 'overactive', 'danilkovich', 'miagkova', 'refractoriness', 'cease', 'felsher', 'unleashed', 'kramer', 'desmosomal', 'tal', 'paraffinembedded', 'takeichi', 'fukuoka', 'xm', 'polyvinyl', 'droplets', 'ju', 'polymerized', 'penetrated', 'ram', 'ix', 'intensifying', 'pcdna4', 'xg', 'jim', 'intensely', 'ctl', 'nondetectable', 'condense', 'adhesions', 'bcatenin', 'ai', 'immunode', 'amin', 'wilhelm', '1993b', 'fuhrman', 'p6', 'subregion', 'nonchimeric', 'yacs', 'minichromosome', 'rehybridized', 'tetrasomy', 'superimposable', 'autoradiographs', '757', 'insure', 'ensure', 'zbar', 'lerman', 'transduces', 'xenografting', 'r905q', 'remitted', '2713c', 'r905w', 'r905g', 'rupture', 'tsc10', '15tsc1', '16tsc1', 'complex21',

'antagonizing', 'trips', 'universit', 'montr', 'woods', 'ophthalmological', 'ultrasounds', 'echocardiograms', 'electroencephalograms', 'tsc25', 'dermato logical', 'rotterdam', 'p8', 'neurology', 'totallab', 'attacks', 'consciousness', 'flickering', 'vibration', 'diurnal', 'sleep', 'carbamazepine', 'phenytoin', 'gingival', 'noncompliance', 'parietooccipital', 'hmms', 'attended', 'sawmill', '1556a', '2415c', 'h732y', 'disfiguring', 'hmm', 'whomet', 'colleague s17', '1826m', 'review35', '4642g', '4684g', 'appendices', 'depigmented', 'reputation', 'lennox', 'gastaut', 'cosegregating', 'ameliorates', 'amphipathic', 'r905', 'mood', '3106t', 's1036p', 'g1556s', 'tsc237', 'widens', 'repetition', 'neurologists', 'save', 'advisable', 'staadt', 'aicda', 'globulin', 'spike', 'p100', 'subclassify', '1128', 'ser32', 'neri', 'smz1', 'rossi', 'mahony', 'fabre', '95051', 'tiled', 'hotstartaq', 'gemini', 'spectrofluorometer', 'unreliable', 'fasta', '958', 'sema', 't1010i', '000077', 'a148t', '004985', '000245', 'n375s', 'g691s', 'r1581', 'v2721', 'americans', 'mexican', 'x00588', 'ectodomains', 'fc1', 'pires2', 'su11274', 'plotter', 'attesting', 'r988c', 'h289', 'h1437', 'ipt', 'monomorphic', '3912', '7677', 'p848l', '1377', 'r280k', 'p210', 'e168d', 'attest', 'l211w', 'alveoli', 'covariables', 'culprit', 'r1443x', 'universite', 'certificate', '2836', 'innovation', 'pq', 'u43746', 'eso', 'rec', 'thy', 'pheo', 'ov', 'sto', 'topped', 'mutyh', 'pers', 'demography', '3418', 'austrian', 'nazionale', 'dei', '3x', 'methanesulfonate', 'f7425', 'rescaled', 'integer', 'midpoint', 'discriminant', 'alignvgd', 'maxent', 'finder', 'hsf', 'dss1', 'ob1', 'equates', '987', 'misinterpretation', 'hbc', 's10689', 'authorized', 'masson', 'partement', 'pathologie', 'cellulaire', '1650', '1a4', '517', 'affirms', 'province', 'mutationpositive', 'les', 'que', 'bec', 'genemarker', 'normalizes', 'designations', 'genebank', '3007', '2244', '3768', '1217', '1443', '1701', '1846', '1742', '3034del4', '1064', '1529', '6503', '2092', '1388', '1522', '1232', '1519', '1245', '1420', '1435', '1436', '1484', '1487', '1488', '1495', '1498', '1504', 'larynx', 'mbr', 'hamper', 'bwh', 'pph', 'kitcomplementary', 'gto', 'rpe', 'myenteric', 'nonidentical', 'kitin', 'lys550', 'val559', 'fma3', 'dashes', 'kitmutants', 'kitgrep', 'rmscf', 'multipotential', 'heterotrimers', 'heterotrimer', 'homotrimer', 'morphogenetic', 'bmp', 'activins', 'attisano', 'wrana', 'blobe', 'caestcker', 'derynck', 'massague', 'wotton', 'lagna', 'mh1', 'mh2', 'abdollah', 'souchelnytskyi', 'homotrimers', 'kawabata', 'chacko', 'smad1', 'foxh1', 'homomeric', 'homotrimerization', 'heteromer', 'calorimetric', 'leu484', 'crystallizable', 'arrange', 'ihkl', '3726', 'bollag', 'tsang', 'bashir', 'ambit', 'sankyo', '679', 'f691', 'interlobe', 'ether', '1691', 'abridged', 'pyridine', 'mv4', 'rs4', '10nm', 'molm14', 'n676s', 'nct01349049', 'polyclonality', 'd839h', 'recreation', '6497', '3844', 'malignancy1', 'kmt2a', 'h3k36', 'h3k36me3', 'tumors7', 'cd56', 'pbmc5', 'h3k79', 'sri', '10d', 'af9', 'mlt3', 'h3k79me2', 'and5', 'lscs', 'flurry', 'tumors33', 'offering', 'branched', 'var', 'trithorax', 'trimethylates', 'compaction', 'mcm7', 'deubiquitylase', '3p25', 'cochran', 'armitage', 'ledgf', 'logfc', 'micrococcal', 'mcm', 'exited', 'adaptability', 'and77', 'confinement', 'evoke', 'dt40', 'g34v', 'truncal', 'diversification', 'lys27', 'lys36', '543', 'h3f3a', 'hemispheres', 'g34r', 'hemispheric', 'impediment', 'k27m', 'thalassemia', 'pericentric', 'neuropathologists', 'titanium', 'epiquik', 'epigentek', '9050', '1791', 'g34', 'humanmethylation450', '904', 'consensusclusterplus', 'csm3', '0017', 'k27', 'k36', 'mirroring', 'extractions', '450k', 'ptm', 'nsd1', 'nsd2', 'nsd3', 'ash11', 'setd3', 'oncometabolite', 'malfunction', 'mispaired', 'irreplaceable', 'innocent', 't33p', 'l187p', 'a272v', 'd603n', 'a636p', 'c697f', 'a834t', 'v923e', 'statusb', 'g162r', 'g164r', 'e749k', 'hba1', 'd3s1611', 'fe11', 'abiprism', 'tes', 'lovo', 'gtbp', 'tub', 'described14', 'bsai', 'linearize', 'uncorrected', 'cybernetics', 'pongo', 'pygmaeus', 'macaca', 'cryptococcus', 'aspergillus', 'fumigatus', 'parsimony', 'classifies', 'mispairs', '3193', 'heteroduplexes', '1833', '1360', 'prokaryotes', '625', 'inflexible', 'locates', 'mangold', 'step1', 'step2', 'step3', 'discrepant', '600678', '000179', '600259', '0

00535', 'umar', 'diagnosing', 'hampel', 'kariola', 'ollila', 'nystr', 'laht i', 'acii', 'ambiguities', 'mispair', 'deduce', 'jcvi', 'sidow', 'mendel', 's tandofrd', 'sidowlab', 'tavtigian', 'acknowledges', 'compose', 'realistic', 'slightest', 'unanimously', 'obtainable', 't45', 'i47delinscf', 'e71del', 'i 330del', 'p578', 'e632del', 'e633', 'e663del', 'l550p', 'p6481', 'p648s', '65 41', 'h329p', 'v612del', 'k616del', 'k618t', 'y646c', 'r687w', 'e460a', 'deductions', 't44m', 'n127s', 'v722i', 'i145m', 'g322d', 'm688v', 'e886g', 'pathogenicities', 'e1193k', 'g566r', 'cyr', 'heinen', 'r1281', 'k728t', 'g881delinks', 'p623l', 'r1095h', 'l1354q', 's144i', 'p1087t', 'p1087r', 'hendriks', 'want', 'a1889v', 'guerrette', 'kondo', 'r659q', 'kohonen', 'corish', 'such s', 'nonpathogenicity', 'f123i', 'l225li', 'bimp3', 'guk', 'neutralizes', 'ta k1', 'k63', 'rests', 'f123', 'l225', 'f130', 'l232', 'lucas', 'jurkat', 'cs k', 'subsaturating', 'sihcard11', '500ul', '30min', '1ug', '30ul', '100ug', 'preclearing', '10ul', '5273', 'ib13026', 'titrating', 's657', 'disengage', '1800', 'k31', 'costimulatory', 'downmodulate', 'paracaspase', 'disassemble s', 'autoinhibit', 'genemorph', 'ezclone', 'ggatccactagtaacggccgcc', 'gcagtt gatgacctcggtcatcag', 'reassayed', 'iplb', 'pebb', 'pebg', 'tggtaagaagctgacga ttc', 'ttcaagaga', 'shnt', 'caacaagatgaagagcacaa', 'ctcgag', 'shm1', 'cctcac taccagtggttcaa', 'sht6', 'ccacgaagagataatggat', 'pmdl', 'rre', '1664', '722 1', '7606', '5z', 'oxozeaenol', '499610', 'pclip3a', 'sin', 'ionomycin', 'ion o', 'p8340', 'p9424', '5611', '8330', '6077', '8032', 'eric', 's40110', 'biologicals', 'mcat', 'ubc', 'mslc7a1', '17224', '3189', 'psuper', 'psrgn', 'oligoengine', 'sihnt', 'ttctccgaacgtgtcacgt', 'l138', 'andc', 'c49y', 'v119e', 'h 129y', 't117a', 'andbb', 'titrations', 'anddd', 'f97y', 'g123d', 'g126d', 'h3 11', 'm32k', 't112i', 'e122d', 'k50m', 'n62y', 'e93d', 'r113q', 'r114i', 'i11 8v', 'y98f', 'toc', 'e121g', 'l127i', 'f130y', 'ande', 'unconjugated', 'c49', 'n62', 'e93', 'y98', 't112', 'e122', 'g123', 'g126', 'l127', 'l244p', 'superinfected', 'shc11', 'licenses', 'c49s', 'f115l', 'f115i', 't117p', 'f130v', 'incur', 'thr18', 'ala19dup', 'gly23asp', 'arg24gln', 'gly35ala', 'gly35val', 'ala57val', 'ala60val', 'ala60arg', 'leu65dup', 'gly67arg', 'gly67', 'asn71del', 'glu69gly', 'asp74tyr', 'thr77pro', 'arg80pro', 'pro81thr', 'arg87trp', 'leu97arg', 'arg99pro', 'leu113leu', 'pro114ser', 'cmm', 'skinned', 'pigmentary', 'sunlight', 'banal', 'armstrong', 'goldstein', 'molven', 'sponsor', 'marne', 'incacance', 'ropole', 'lanome', 'homme', 'porc', 'pre2005', 'lncc', 'fd1', 'anna', 'gutierrez', 'arroyo', 'wolfson', 'thorn', 'middlesex', 'mortime r', 'w1t', '3aa', 'franc', 'oise', 'avril', 'cochin', 'pavillon', 'tarnier', 'assas', '75006', 'colleague', 'agne', 'brigitte', 'paillerets', 'camille', 'desmoulin', 'igr', 'sunburn', 'chaudru', 'velden', 'begg', 'poθ', 'drayton', 'sharpless', 'imperceptible', 'rizos', 'laud', 'harland', 'prb', 'cip', 'kip', 'brookes', 'p19ink4d', 'byeon', 'russo', 'brotherton', 'venkataramani', 'loophelix', 'lilischkis', 'biodesktop', 'arg80', 'met53ile', 'gly101tr p', 'phylogenetic', 'degreer3', 'fm', 'mpm', 'necker', 'bice', 'opossum', 'multalin', 'prodes', 'toulouse', 'inra', 'af044336', '031550', '004927', '52324 0', '524145', 'sus', 'scrofa', 'aj316067', 'mesocricetus', 'auratus', 'af2925 67', 'monodelphis', 'domestica', 'af064808', 'ay138246', 'xiphophorus', 'maculatus', 'p13cdkn2x', 'u69273', 'helleri', 'af132500', 'takifugu', 'rubripes', 'p13cdkn2a', 'aj250231', 'felis', 'catus', 'baa33540', 'aag01087', 'bork', 'emblicheidelberg', 'submit2', 'panther', 'snps3d', 'umd', 'roud', 'pgems', 'fidelis', 'evry', 'parry', 'coprecipitate', 'arg24pro', 'val59gly', 'ala68leu', 'asn71ser', 'leu16pro', 'pro48thr', 'ser56ilo', 'leu62pro', 'asn71lys', 'val 126asp', 'hdfs', 'arg87pro', 'snp3d', 'cosegregated', 'pdb1dc2', 'g35', 'a5 7', 'pdb1bi7', 'g23', 'r24', 't77', 'r87', 'gly23', 'ala60', 'asp74', 'pro8 1', 'arg87', 'leu97', 'pro114', 'gly35', 'thr77', 'arg24', 'ala57', 'glu69', 'arg99', 'pollock', 'ciotti', 'concordances', 'z4', 'z3', 'reporta', 'contextb', 'exonc', 'mutationd', 'changee', 'segregationf', 'reports', 'reportsh', '3063', '57dup', 'caldas', '1353', '68g4a', '10710', '71g4a', '104g4c', 'ghiorzo', '1772y', '104g4t', '1468', '19696', '170c4t', 'quesnel', '1338y', '179c

4t', 'gazzeri', '15149', '179delinscg', 'gly74ala', 'gretarsdottir', '11384', '194dup', 'ala79dup', '3359', '199g4c', 'arg81pro', '2729', '213del', 'arg82', '86del', '3018m', '10249', '18750', '2fm', '206a4g', '10704', '220g4t', 'arg88leu', '12670', '229a4c', 'his91pro', 'soufir', '16086', '239g4c', 'gamielien', '2847y', '2919y', '241c4a', 'thr95asn', '10411', '14725', '2mpm', '259c4t', 'pro101leu', '290t4g', '12477', '296g4c', 'castellano', '3181', '15138z', '12826z', '14302z', '339g4c', '340c4t', '11131', 'p114s', 'ala128leu', 'monnerat', 'kannengiesser', 'nfm', 'nmpm', 'n1pa', 'billiary', 'yakobson', 'reymond', 'ranade', 'della', 'r30', 'z60', 'hdf', 'respecting', 'masbakk', 'auroy', '34g4t', 'coinheritance', 'kefford', 'ank1', 'cpk', 'residuesb', 'polyphenc', 'siftc', 'pantherc', 'snps3dc', 'pmutc', 'predictorc', 'bindingf', 'testg', 'conclusionh', 'concordancei', '47993', '3172', 'rldh', '03473', '5322', '10969', '1192', 'gly35vla', '90368', '4036', 'avptse', '10789', '3474', '02285', '4555', '70107', '7864', '60066', '7928', 'edns', '409', '16053', '6874', '82893', '2892', '5802', '0686', 'rvslt', '16882', '15187', '532', '13593', '9304', '63195', '2659', 'rqgek', '813', '42579', '2080', '57943', '2231', '65297', '1096', 'gqt', '21928', '7330', 'aimpact', 'ballowed', 'cgrantham', 'christophe', 'montpellier', 'din', 'eas', 'fcdk4', 'gcell', 'hfunctional', 'iconcordance', 'collate', 'untested', 'backed', 'credentials', 'orlow', 'ab6546', 'mab88057', '554132', 'hs68', '2ha', 'serizawa', 'jnk3', 'lukas', 'medema', '33ins9', '32dup', 's8trip', 'spm', 'oligomerize', 'additions', 'detergents', 'p1141', 'g101w', 'd84h', 'mts1', 'kamb', 'meyerson', 'matsushime', 'ller', 'whyte', 'kouzarides', 'dyson', 'mts2', 'hannon', 'quelle', 'greenwald', 'cairns', 'wainwright', 'hopkin', 'cordon', 'cardo', 'hartwell', 'nonpeptide', 'g104t', 'errore2d', 'phagemid', 'v35g', 'correctness', 'curlis', 'described', 'b834', 'doherty', 'tryptone', 'paste', 'stirred', 'solubilization', 'preequilibrated', 'spu', 'excessively', 'sephadex', 'sephacryl', 'stirring', 'centriprep', 'lyophilized', 'mepsadwlata', 'm9', '15nh4cl', '13c6', 'rifampicin', 'baleja', 'muchmore', 'ink4b', 'shpaer', 'complimentary', 'pet17b', 'rz', '1032', 'r408', 'phage', 'g35v', 'd84n', 'd84y', 'gtd', 'c72a', 'a127g', 'a127s', 'gcy', 'pga', 'pgad424', '1441', 'bsphi', 'pas', 'plam5', 'lamine', 'cg1945', 'y187', 'diploids', 'onpg', 'mrf', 'isoelectric', 'phastsystem', 'kratos', 'kompac', 'maldi3', '500c', 'd2o', 'dmx', 'hncacb', 'ccbca', 'spectrometers', 'pochapsky', 'kay', 'cosy', 'tmix', 'presaturation', 'piotto', 'sklenar', 'bodenhausen', 'ruben', 'mlev17', 'hnca', 'grzesiek', 'witekkind', 'mueller', 'muhandiran', 'ikura', 'hcch', 'dipsi', 'shaka', 'muhandiram', 'vuister', 'no3', 'hno3', 'srinvasan', 'lichter', 'silicon', 'indy', 'indigo', 'workstations', 'xwin', 'felix', '950', 'biosym', 'softwares', 'extraaneous', 'renatured', 'detergentsmethanol', 'dioxane', 'octylglucosidewere', 'widths', 'p16a', '1945', 'refolded', 'dialysate', 'impurity', 'renaturatin', 'disaggregation', 'refolding', 'approximation', 'mutantsa', 'refraction', 'cantor', 'schimmel', 'his98', 'tyr129', 'indole', 'w15', 'w110', 'leucines', 'homonuclear', '15nh', 'capital', 'downfield', 'upfield', 'spera', 'glu27', 'ala36', 'arg58', 'ala68', 'val82', 'phe90', 'ala102', 'asp116', 'his123', 'ala133', 'leu32', 'ala34', 'ala100', 'ala118', 'glu120', 'glu35', 'leu64', 'asp84', 'ala86', 'gly89', 'val115', 'cys72', 'arg103', 'asp108', 'prolines', 'ile145', 'asp156', 'spectum', 'elaborated', 'trimers', 'floppy', 'trp15', 'trp110', 'invertebrate', 'ankyrins', 'yochem', 'matsuno', 'cdc10', 'aves', 'akr1p', 'kao', 'vaccinia', 'gillard', 'prospects', 'ample', 'exchanging', 'collaborator', 'noel', 'unaggregated', 'haemopoiesis', 'reiter', 'keersmaecker', 'prolific', 'vizmanos', 'delaval', 'hidalgo', 'normoblastic', 'wdr48', 'bin2', 'reinterpreted', 'erythematous', 'hydroxycarbamide', '230314', '10005', 'wd', 'samiri', '872', 'thornton', '002078', 'golgin', '991', 'grip', 'n5', '016187', '25bp', 'singe', 'prendergast', 'internalisation', 'bin1', 'blanks', 'pancytopenic', 'fungal', 'commencing', 'rabep1', 'hip1', 'sptbn1', 'dic', 'brodsky', 'olabisi', 'tohru', 'osada', 'ohno', 'katabuchi', 'kanazawa', 'hiroyuki', 'kamata', 'insurance', 'kudo', 'hiroshi', 'nihon', 'nippo

n', 'rosai', 'kitasato', 'tanaka', 'jikei', 'oita', 'tsujimoto', 'teramoto', 'gunma', 'seirei', 'hasegawa', 'fujii', 'fujiwara', 'kawasaki', 'yajima', 'yasuda', 'tsutomu', 'municipal', 'wada', 'aida', '3144', 'questionnaires', 'sitespecific', 'alf', 'ova', '2195', '2508', '2731', '3226', '3493', '1214', '1373', '1778', '1447', '3026', '0076', '3232', '3667', '4427', 's1613g', '4956', '1613', '203g', 'h372n', '1342', '1593', '2578', '3624', '1132', '4035', 'marriage', 'eat', 'hedenfalk', '0077', 'ent', 'obese', 'gonadotropin', 'urgent', 'southeast', 'fgf19', 'kmt2b', 'm114', 'neuveut', 'sawey', 'svs', 'basis', 'hbsag', 'seropositive', 'hcv', 'capturing', '550k', 'invalidated', '5015', '4482', '3739', '809', 'midrange', 'totoki', 'sriv', 'il6st', 'rebouissou', 'axin1', '4314', 'banet', 'poon', 'rejto', 'youn', 'lrp1b', 'eps15', 'slc10a1', 'cacna2d4', 'adcyl2', 'fam5c', 'col11a1', 'defenses', 'aflatoxins', 'gozas', 'miyoshi', 's33', 's37', 'axin2', 'vinyl', 'weihrauch', 'arid2', 'wiegand', 'kmt2c', 'echoing', 'ush2a', 'usher', 'ush', 'deaf', 'blindness', 'reinehrs', 'gpr98', 'pcdh15', 'myo7a', 'quintas', 'cardama', 's703i', 's729c', 'q644h', 'v645f', 'hornakova', 'hek293ft', '1910p', 'n242s', 'g902e', 'il6', 'pilati', '3eyg', 'flex', 'saharinen', 'silvennoinen', 'a634d', 'nonactivating', 'quint', '911453', 'purandare', '13p', '21p', '1967', 'elk4', 'ptk2', 'rptor', 'irak1', 'ptk2b', 'focally', 'lgr', 'liberzon', 'whittaker', 'fatima', 'wnt9a', 'fzd6', 'dact1', 'kypta', 'damages', 'inflicted', 'mishra', 'lgr5', 'karin', 'underappreciated', 'il7', 'il20', 'il6r', 'il21r', 'nault', 'adcyl3', 'instable', 'executor', 'tnfrsf10a', 'tradd', 'casp9', 'dff', 'hoshida', 'mesa', '0332991', 'yom', 'audeh', 'tutt', 'clonogenicity', 'morbicentric', 'eludicate', 'y589f', 'classs', 'naoe', '2801', 'succession', 'mds12', 'elongates', 'pmosblue', 'bucks', '21m13', '1334', 'muni', 'caacaattgggtttgttcctc', 'catgatatctcgaggccatccaaag', 'pcdhf3', '2503', 'cctcagataatgagttcttctacgttatttcag', 'ctgaaatcaacgtagaagaactcattatctgagg', 'cctcagataatgagttacttctcgatttcag', 'ctgaaatcaacgaagaagtactcattatctgagg', 'y591f', 'na2moo4', 'maleic', 'hrplabeled', 'plusagarose', '589yfyvdfreyey599', 'y597', 'y599', 'upsteam', 'mt6', 'mt5', '1335', '608e', 'unessential', 'y579', 'yfy', 'ye', 'nakao', 'yokota', 'horiike', 'al12', '2131', 'highmolecular', 'neuromuscular', 'd593', 'k602', 'f605', 'tandemly', 'leukemiatype', 'dnstat5', 'dnras', 'stat5pathways', '3907', '3914', 'flt31', 'f14', 'samples18', 'hosts', 'nster', 'forschungsgemeinschaft', '119804', '229838', 'wolfgang', 'berdel', 'universita', 'tsklinik', 'albertschweitzer', 'strasse', '48129', 'muenster', 'pepro', 'antiphosphostat5', 'antiphosphoerk2', 'antiphosphoakt', 'felicitas', 'rosenthal', 'aml5', '1562', 'sspi', 'bsmbi', 'nt1567', '2077', 'substitutionally', 'lemischka', '59ltr', 'momsv', 'pmam', 'bsd', 'kaken', 'ovine', 'y694f', 'popali', 'popi3cat', 'lac', 'plam', 'lacr', 'poprsvi', 'rass17n', 'uninduced', 'thiogalactose', 'ppur', 'zncl', 'navo3', '3908', 'mizuki', 'stat5y694f', 'pappenheim', 'nt1624', '2134', 'hmwt', '3909', 'evidencing', 'itd1', 'itd2', 'moab', '3910', '3911', 'aml18', 'ligandactivated', 'terk', 'antitotal', 'activatesakt', 'coactivated', 'tkd1', 'itd627e', 'controversially', 'gale', 'offs', 'autophosphorylate', 'mln', 'scheduling', 'amlsg', 'hsct', 'transfusion', 'pharmacodynamic', 'increments', 'amendment', 'qt', 'myelosuppression', 'haploidentical', 'allografting', 'yid', 'heidel', 'hampering', 'su5614', 'n676', 'nct01522469', 'n676d', 'f691i', 'g697r', 'y842c', 'rewired', 'unraveling', 'birnbaum3', 'series4', 'shurin', 'al6', 'blackfan', 'cir', 'calvin', 'sawyers39', 'flt5', 'gttaaccatgccggcggtggc', 'gttaactacgaatctcgacc', 'gggtttgtctcatttcattgtca', 'aaagcacctgatccttagtaccc', '1801', '604f', 'caggggaaagctgtaaag', 'w51b2', 'gatcatattcatattctctg', '1781', '2982', 'w51f2', 'tcagagaatataatgat', 'tatcttaggtgacactatag', 'polylinker', 'w73', 'w78', 't6', 'w73b2', 'ccatattcatattctctgaaa', 'w73f2', 'gctctacgttattcagag', 'w78b2', 'ccttcatttcattctgaaa', 'w78f2', 'gttggtaggtgaccggc', 't6b2', 'ccaaactctaaatttctttgg', 't6f2', 'ttacgttattcagagaatatg', 'posb2', 'taaatttctttggaaactccatttg', 'posf2', 'tgctcctcagataatgagttactc', 'neo42', 'mscv2', 'pik6', 'genesys', 'redwood', 'weekold', 'asphyxiation', 'rbcl', 'm

'icroisolator', 'autoclaved', 'bannockburn', 'permanganate', 'metabisulfite', 'ferric', 'ammonical', 'metasulfite', 'thiosulfate', 'cd32', 'biotinconjugated', 'caltag', 'epics', 'altra', 'granzyme', 'radprime', 'previously44', 'npo s', 'mut117', 'flt347', 'wrightgiemsa', 'emaciation', 'ruffled', 'fur', 'r2 8', 'doublepositive', 'provirus', 'leaching', 'magnifi', 'n78', 'peconjugated', 'mhy11', 'respectively14', '3569', 'flt1td', 'mastocytoses', 'phenotype7 1', 'ltrs', 'sti57173', 'csc', 'birc5', 'adcc', '1789', 'foxo3', 'gelcount', 'optronix', 'profiler', 'dorsum', 'counteracted', 'pegfr', 'il8', 'aldehyd e', 'and6b', 'renew', '6b6b', 'ameliorated', 'her1', 'nashville', 'scraping', 'dounce', 'strokes', 'igg2a', 'avanti', 'oxalate', 'kobe', 'spectrumqua', 's canalytics', 'redirect', 'redirection', 'multimerization', 'mkn', 'afx', 'cyt ostasis', 'impeding', 'ado', 'emtansine', 'cep17', 'h878y', 'nephron', 'synov ial', 'mammography', 'depolymerizing', 'afatanib', 'q3', 'orr', 'q2', 'thioether', 'phagocytosis', 'delivers', 'baso', 'interventional', 'randomizing', 'arcila', 'g776l', '1869r', 'going', 'n857s', 'lobuloalveolar', '1209', 'sphi', 'neunt', 'mcmaster', 'postcoitum', 'pseudopregnant', 'virgin', 'protections', 'riboprobe', 'grb', 'interbreeding', 'leaky', 'lactating', 'digitized', 'term inating', 'tm1', 'adj', 'tm2', 'tm3', 'tm5', 'tm4', 'tm6', 'tm7', 'negated', 'thank', 'grateful', 'monica', 'hc11', 'pads', 'hynes', '1959', 'medina', 'comm', 'rapidity', 'cfp', 'pyrrolo', 'fabbro', 'edwards', 'wnt7b', 'lengthy', 'mentioning', 'olayioye', 'c712r', 'bilaterally', 'dihydrofolate', 'weintraub', 'vairo', 'brg', 'hbrm', 'elf', '1996a', 'adipocyte', 'neuman', 'saos2', '928', 'teto', 'kossa', 'schwarz', 'claudio', 'retinoblastomas', '661w', 'ex 4', 'dryja', 'undermine', 'ceased', 'decaprio', 'comm', 'innervation', 'neec', 'mmmt', 'impute', 'inclusivity', 'm326i', 'a727t', 'a727v', 'thr180', 'try r182', 'stathmin', 'caveolin', 'intermingled', 'characterizemutations', 'e439 del', 'r574fs', 'e160', 'r348', 'r503w', 'f268i', 'a171v', 'v727t', 'g1049r', 'efe184', 'thatbinding', '1693', 'proteosome', 'rock', 'pkn', 'leung', 'nakagawa', 'ishizaki', 'amano', 'mlc', 'kimura', 'rhotekin', 'rhophilin', 'citron', 'reid', 'mdia2', 'mdia', 'matsui', 'narumiya', 'lpa', 'olson', 'khosravi', 'qiu', 'willumsen', 'joneson', '1996b', 'westwick', 's190', 'rescreened', 'f39a', '1351', 'epitopetagged', 'twohybrid', 'sahai', '942', '1053', '1327', 'g14v', 'pkn1', 'filamentous', 'insets', 'sre', '1354', 'filaments', '539', 'y11', '1355', 'nraf', 'y42', 'ordinates', 'sdc25', 'oxidase', '1357', 'thicker', 'spindles', 'disagreed', 'rok', 'b11', 'd11', '1358', 'uehata', 'mbs', 'lightchain', 'n19', 'lebowitz', '1359', '5171', 'stratech', 'chris', 'optim em', 'gibcobrl', 'retrosternal', 'ck', 'ae1', 'ck19', 'cd5', 'ck7', 'asthenia', 'y553n', 'ast', 'tot', 'bil', 'symptom', 'thymomas', 'bel', 'h697y', 'tos ylate', 'v559a', 'e490k', 'mandibulofacial', 'dysostosis', 'neurosecretory', 'reggio', 'emilia', 'padova', 'cremona', 'organisation', 'precepts', 'immuno stainer', '3diaminobenzidine', 'chromogene', 'dakopatts', 'sp19', '4a4', 'lk2 h10', 'nonimmune', 'uno', 'thermoblock', 'biometra', 'gottingen', 'centri', 'sep', 'runned', 'annals', 'lymphoepitelioma', 'lymphoepithelioma', 'immunor eacted', 'heterozygosis', 'immunoprofile', 'chemoradiotherapy', 'adenoid', 'tsuchida', 'petrini', 't1889', 'yoh', 'strobel', 'bisagni', 'disel', 'del577', 'buti', 'chuah', 'hamada', 'nessuno', 'giaccone', 'palmieri', 'girard', 'keratinization', '579del', 'somatostatin', 'antiangiogenetic', 'categorize', '164 2', 'interviews', '3z', 'm1628tf', '1708', 'rlu', 'v1804', 'monogenic', 'uc v', 'ucvs', 'interspecies', 'd1546n', 'e1644g', 'l1664p', 't1700a', 'v1713a', 'q1785h', 'g1788d', '24endtand', 'ux13', 'ralph', 'ecor1', 'pgtb9', '1t29', 't2k', 'aaa73985', 'aag43492', 'aat44835', 'aat44834', 'mullata', 'aat4483 3', 'aac48663', 'aal176094', 'aax92675', 'aad00168', 'aac36493', 'aak83825', 'aal13037', 'aar89523', 'renderbyattribute', 'findhbond', 'hispanics', 'brct domains', '1766s', 'l1546n', 'm933', 'h2116r', 'e1664g', 'nonpaternity', 'aus', 'somber', 'neutrals', 't1700', 'gly1613', 'ser1613', 'discipline', 'forefront', 'consultations', 'surpassing', 'confuse', 'y285c', 'd595v', 'd931y', 'k935i', 'n181s', 't244r', 'r306s', 'v348l', 'h618p', 'puroerbb4', 'acyt',

'neuregulin', 'nrg', '93k', 'v28', 'v58', 'f288', 'y285', 'incoming', 'd595', '212k', 'heterodimerizaton', '2ahx', '3u7u', 'phenylalanines', 'valines', 'r306', '1n8z', 'f273', 'y268', 'p270', 'v427', 'p294', 'c293', 'c304', 'c262', 'c289', 'f297', 'vacated', 'stabile', 'c593', 'd585', 'p594', 'l597', 'k606', 'k592', 'relocate', 'e591', 'q558', 'f583', 'r612', 'relocates', 'k592i', 'f603', '3bce', 'k955', 'd767', 'v763', 'a769', 'l770', 'i947', 'm951', 'v954', 'm958', 'i959', 'v954r', 'k751r', 'wo', 'd931', 'k935', '193k', 'erb b4jm', 'phophorylation', 'k714', 'e715', 't716', 'c814', 'l815', 's816', 'p817', 'cup', 'p790', 'l932', 'e934', 'e937', 'timepoints', '151k', 'mutants34', 'irrespectively', 'h115n', 'footprinting', 'rrrcatgyyy', 'bln22', 'wrn', 'act ivities30', 'olignucleotide', 'his178', 'his296', 'his297', 'h168n', 'h179n', 'h297n', 'h193n', 'h214n', 'h233n', 'purifying', 'oligonucleotides44', 'footp rinting40', 'discernibly', 'h1229', 'mainiodide', 'nucleases', 'iii54', 'endo nuclesae', 'hap1', 'h115', 'overproduced', 'relevantly', 'r143a71', 'r175p72', 'urticaria', 'pigmentosa', 'dysphagia', 'prominence', 'patchy', 'cecal', 'echocardiographic', 'knees', 'perineum', 'millimeters', 'palpated', 'gastrectomy', 'reconstruction', 'elective', 'wife', '1669', 'invitation', '0007', 'innocuous', 'digits', 'elbows', 'puberty', 'fade', 'speckled', 'incident', 'peristalsis', 'dysmotility', 'fasman', 'amgen', 'oaks', 'lipofectaminetm', 'macvector', 'brookhaven', 'mmnacl', 'theleft', 'cylinder', 'minima', 'block ers', 'hannover', 'immunohistopathologic', 'palm', 'microbeam', 'hay', 'primaries', 'conjunctiva', 'longley', 'hodi', 'lutzky', 'alexis', 'conjunctival', 'molecgist', 'bordeaux', 'paoli', 'calmettes', 'chu', 'strasbourg', '1772', '559del', 'anywhere', 'gsf', 'send', 'blay', 'germain', 'laye', 'physician', 'immunostains', 'lapp', 'swog', 'eortc', 'exclusions', '491', 'afip', '1767', 'tyr570', '558del', 'miettinen', 'microscopes', 'confusing', 'populationbase d', '1769', '422', 'rfs', 'forthcoming', '1684', 'races', 'benefiting', 'hosp italized', 'elastosis', 'bastian', 'hospitalization', '465', 'n566d', 'l576', 'abe', 'carvajal', 'kitexon11mutation', 'asubset', 'choroidal', 'lmost', 'kit copy', 'vamedical', 'georgia', 'mailcode', 'l471', '6834', '6787', 'corless c', '0575', '6821', 'chelex', '95jc', 'nikiforova', 'cactagggtttcattccatt g', 'tccttaatacagaatatggtaaaga', 'tgaardactgagtgacaaactgggtgggtfluorescen', 'lc640', 'agcaggtgggttggga', 'lc480', 'mixand', '60jc', '72jc', '45jc', 'f15', 'aaagatttgatggtagtgc', 'gaaactaaaaatccttgca', 'tgtggtaaaggaaac gtgagta', 'iablkfq', 'gapd', 'cactaggcgctcactgttct', 'gcgaactcacccgttg', 'txred', 'tgggaaaggtaaggctcga', 'iabrqsp', 'aggtagaaaactgaaaaagacatgc', 'gcctttatctcgtagatgaagt', 'aggcatgctatccacaggtgattg', 'idtdna', 'kitmutant', '6822', 'ddct', 'chebychev', 'inequality', 'tricolor', 'scfd2', 'lnx', '73jc', 'h obbs', 'rhnp800p', 'anorectum', 'evqwkv', 'f10', 'sixmucosal', '6823', 'g13r', 'immumunohistochemistry', 'willmorepayne', 'nonmucosal', 'cpreselected', '5591', 'v569g', '6824', 'vulvar', 'vaginal', 'cont', '6825', 'woenckhaus', 'chiriac', '11f', 'nonchoroidal', '915', '6826', 'spindled', 'y503', 'h10471', 't574', 'i563', 'zytomed', 'nucleospin', 'macherey', 'nagel', '11r', '12r', '18f', '18r', '15r', '20f', '20r', '7r', '10f', '10r', 'fluka', 'anneal', 'intraexonic', 'k642', 'whorls', 'brisk', 'mitotically', '8700', 'melanomas', 'established11', 'predisease', 'strenuous', 'ambulatory', 'sedentary', 'housework', 'federal', '722f1', '365h22', 'hinxtion', '799e21', '19f13', '461g20', 'motorized', '400mgorally', 'ctep', 'protocoldevelopment', 'ctcae3', '6w eeks', 'decreaseorlessthan20', '1week', 'design32', 'efigure', 'inadvertentl y', 'csd34', 'nokit', 'etable', 'wean', 'methadone', 'contraindicated', 'proto colduetocyp3a4interactions', 'optional', 'utations', 'melanoma8', 'gist35', 'gist37', 'v559c', 'v642e', 'pathogenetically', 'preselected', 'thorac', 'no ntumoral', 'hospices', 'civils', 'claude', 'bernard', 'declare', 'hopital', 'issn', '1556', 'g863d', 'g13v', 'chemorefractory', 'nexavar', 'branchburg', 'avastin', 'nonsquamous', 'aflibercept', 'motesanib', 'amg', 'dematteo', 'sak urai', 'mib1', 'spiral', 'oedema', 'hyalinised', 'shrunken', 'goldman', 'melo', 'gbm3', 'bannayanazonana', 'seiryo', 'machi', '3937', '3943', 'astec', 'ac

ggtagacacgcttccctggattgg', 'ctgtcagtgggaacaagaagtggaga', 'ttctgccatctctctcctc c', 'tttcatgggttttatccctc', 'shuzo', 'ctcgtcattgttctgttcc', 'cgggacaaagcaa ttgaag', 'his2', 'suprec', 'ttgtttagggcaggggagt', 'ctggcccctgtcatcttct', 'gcc cttcctcagcatcttat', 'ggaagagaatctccgcaaga', 'cacagctagaacttatcaaacc', 'tgcaca tattcattacaccagg', 'ggattatagaccagtggcac', 'agcatttgcagtatagacggt', 'sepagen e', 'sanko', 'junyaku', 'd10s579', 'd10s215', 'd10s541', 'tacgcggatccatgacagc catcatcaaagag', 'agccaaagcttcagactttgttaatttgtatgc', 'phk101', 'm199del', 'phk102', 'phk103', 'phk104', 'phk105', 'phk106', 'dthiogalactopyranoside', 'phenylmeth', '3938', 'histologya', 'gradeb', 'his2c', 'mutatione', '10q23 f', 'r249g', 'agg3ggg', 'e7x', 'gag3tag', 'm243l', 'atg3ctg', 'cgt3tgt', 'cat 3stat', 's240g', 'agt3ggt', 'c135g', 'tgc3ggc', 'p2221', 'ccg3ctg', 'cgc3tgc', 'p177r', 'ccc3cgc', 'd186g', 'gat3ggt', 'apa', 'ao', 'd259v', 'gac3gtc', 'ggc 3gac', 'y234c', 'tac3tgc', 'cgg3tgg', 'tgt3tat', 'h179r', 'cat3cgt', 'cgt3ca t', 'cga3gga', '742insc', 'cct3ccct', 'ttt3gtt', '514del', 'gtcagag3gg', 'tat 3tgt', 'cgc3cac', 'tcc3ttc', 'l348x', 'ttg3tag', 'ganglioglioma', 'meningiom a', 'neurocytoma', '3939', 'ylsulfonyl', 'nanocep', 'a2b1', 'gangliogliomas', 'suitably', 'ptena', '10q23b', '3940', 'r15s', 'r15i', 'r129e', 'tfl', 'tv', 'cd58', 'tnfrsf14', 'stat6', 'studies9', 'genes17', 'hlh', 'immunoreceptor', 'tyr196', 'dlbcl14', 'aupd', 'reaffirming', 'rel', 'lymphomas30', 'strelka', 'xt', 'dbsnp132', 'varsan2', 'dnacopy', 'redrawn', 'bowtie2', 'ccf', 'op9', '7500', 'described57', 'akta', 'schwannomas', 'meningomas', '146r', 'l141p', 'a211d', '1535p', 'halflives', '4982', 'knudsen', 'l64p', 'singlenucleotide', 'schwannomin', 'blastocyst', 'schreiber', 'keegan', 'mec1', 'tel1', 'morrow', 'descending', 'autopsied', 'intercrosses', 'pellucida', 'devd', 'woo', 'prein cubation', 'catastrophe', 'schlegel', 'pines', 'matsuoka', 'chaturvedi', '09 7', 'elledge', 'lengauer', 'abraham', 'shiloh', 'stalled', 'aphidicolin', 'cha', 'robertson', 'partnered', 'recombinational', 'illegitimate', 'sensors', 'methylating', 'papadopoulos', 'cahill', 'promoterless', 'au1', 'seckel', 'masramon', 'mlh', 'hsrs', 'bartek', 'exaggerated', 'opportunistic', 'integrants', 'transcarbamylase', 'stemmed', 'litters', 'mnng', 'atrip', 'wholly', 'exacerbated', 'rampant', 'recombinogenic', 'postreplication', 'interdependence', 'datto', 'junjie', '33258', 'ribonucleotide', 'satellite', 'rainbow', 'chrom a', '8300', 'sensys', 'ntrk1', 'fusionseq', '1835', 'co9', 'sept14', 'kif5a', '5411', 'abrupt', 'mprip', 'samsung', 'gsnap', 'dasper', 'nts', 'satisfy', 'unmatched', 'cosine', 'genepattern', 'bcs', 'pax2', 'noble', 'donis', 'mulligan', 'men2a', 'men2b', 'fmtd', 'raue', 'gdnf', 'artn', 'nrtn', 'pspn', 'gpi', 'erk5', 'mtc', 'men2', 'asai', 'borrello', 'santoro', '618', 'met918thr', 'carlson', 'hofstra', 'gimm', '321', '891', 'berndt', 'carlomagno', 'zd6474', 'axitinib', 'skinner', 'quantization', 'puc18', 'hts', 'pentagastrin', '153 0', 'erc', 'topflash', 'gupta', 'henderson', 'generalization', 'pyrazolopyrim idines', 'knowles', 'brandi', 'counters', 'alps', 'pbls', 'sinopulmonary', 'disc', 'pbl', 'erythematous', 'chimaeric', 'cleave', 'cd69', 'pokeweed', '110 k', 'zvad', 'mucocutaneous', 'staphylococcus', 'aureus', 'cowan', 'microbeads', 'onestep', 'microl', 'flice', 'immunological', 'described25', '395', 'd59 4', '1597s', 'tak', 'uninvolved', 'qual', '1790t', '089', 'd594n', '1780g', '1790ct', 'comprehensiveness', 'overlook', 'prohibitively', '3395c', 'p1132 l', '3103g', 'd1035n', 'c154t', 'r52x', 'monotherapies', 'rp2d', 'palmar', 'actinic', 'keratosis', 'arthralgia', 'cfDNA', 'historic', 'cobas', 'braf1597 r', 'uva', 'discontinue', 'retinopathy', 'occlusion', 'pruritis', '0973', 'suited', 'alopecia', 'corticosteroid', 'cobimetinib', 'val600glu', 'sir', 'comp assionate', 'commencement', 'm1b', '1119', 'aihw', 'flaherty', 'hauschild', 'larkin', 'maldonado', 'bowyer', 'dahlman', 'menzies', 'melbourne', 'spitzoi d', 'v600m', '1790delinstc', 'leu597ser', '1801a', 'lys601glu', '1798', '1799 insatacag', 'thr599', 'val600insaspthr', '1802delinsagat', 'lys601delinsglull e', 'c1794', '1796dup', 'thr599dup', 'hazel', 'keratoses', 'ssm', 'lmm', 'lentigo', 'maligna', 'krauthammer', 'whiteman', 'neoantigens', 'snyder', 'ctla', 'thinner', 'nct02296112', 'victorian', 'vms', '1788', '1823', 'hrm', 'multino

mial', 'melanoma3', 'predose', 'csr', 'enroll', 'whichever', 'existent', 'embolism', 'combi', 'deah', 'brodie', 'callebaut', 'mornon', 'heterodimerizes', 'welcsh', 'rutter', 'wce', 'preimmune', 'microcapillary', 'nanoelectrospray', '1249', 'coin', 'synaptonemal', 'transverse', 'comigration', '17q22', 'genemap', 'clonetech', '1043', 'asynchronously', 'tops', 'lowermost', 'costaining', 's35', 'cleaver', 'k52r', '6xhis', 'callahan', 'pro919ser', 'germlines', 'p47a', 'm299i', 'intimately', 'nbs', 'mrn', 'drapkin', 'scp1', 'friedberg', 'nonlethal', 'unquestionably', 'blr', 'pgex2tk', 'pgst', 'strips', 'lcq', 'quadrupe', '1300', 'microchemistry', 'howley', 'not1', 'sal1', 'fitzgerald', 'adr', '2774', 'prim', 'wavemaker', 'melt', 'ewen', 'livingston', 'rad4', 'rad16', 'reca', 'uvrd', '633', 'finch', 'hansen', 'nakayama', 'hydrolyzes', 'singlestranded', '2345', 'cour', 'singlestrand', 'g50', 'cdc', 'nicks', 'incisions', 'mgatp', 'sigal', '3266', 'sedimentation', 'uvrabc', 'displacing', 'matson', 'richardson', 'stahl', 'gough', 'dispensability', '2137', 'trpl', 'fluoroorotic', 'monoq', '460', 'centricon', '1964', 'unreacted', 'hci', 'kci', 'suction', 'polystyrene', 'sarstedt', 'y1', 'safeguard', 'xpg', 'ercc5', 'ercc6', 'mailing', '2300', 'acci', 'a9', 'pst1', '782', 'lowercase', 'designates', 'nsii', 'cj236', 'p35', 'gpt', 'fluence', 'stransferase', 'immunize', 'mobilites', 'kilodaltons', 'overloaded', '645', 'lv', 'striped', '1111', 'fk', '4131', 'qm', 'ul5', 'eif', 'bloom', 'rothmund', 'thomson', 'nappi', 'netn', 'gellert', 'm13mp18', 'wilsonville', 'microspin', 'duplexed', 'boric', 'e67', 'rabbits', 'thioredoxin', '1006', 'motors', 'prerequisites', 'supercoiled', 'oligoribonucleotide', 'oligodeoxynucleotide', 'caretakers', 'multimers', 'multimer', 'p10341', 'd593v', 'g595r', 'industries', '0063', 'tripeptide', 'villous', 'signalings', 'schopfer', 'garnett', 'rushworth', 'prevail', 'transphosphorylate', 'dankort', 'dhomen', 'ci1040', 'disappointingly', 'solit', 'sebolt', 'leopold', 'mm415', 'sw620', 'r1881braf', 'fractionations', 'proving', 'd594abraf', 'k483mbraf', 'posit', 'd594a', 'paws', 'braflox', 'cameron', 'frye', 'comprehend', 'pathognomonic', 'msin1', 'hifs', 'kirc', 'dgdp', 'solenoid', 'clamps', 'k2', 'fractional', 'oppose', 'synergism', 'integrator', 'negate', 'trd1', 'trd2', 'subtherapeutic', 'blueprint', 'lsrfortessa', 'nsg', 'centennial', 'fascinated', 'dexter', '1914', 'hunt', '1917', 'revolution', 'spyros', 'artavanis', 'tsakonas', 'notching', 'kidd', 'wharton', 'neuroscience', 'fortini', 'nineties', 'ellisen', 'gallahan', 'jhappan', 'deluge', 'cementing', 'notch3', 'haines', 'okajima', 'rebay', 'ntm', 'furin', 'blau Mueller', 'jagged', 'dll13', 'dll14', 'jag2', 'gets', 'adam10', 'metalloproteases', 'psen1', 'psen2', 'nicastrin', 'ncstn', 'pen2', 'pharynx', 'aph1', 'stropooper', 'schroeter', 'bray', 'metalloprotease', 'selkoe', 'cbf1', 'lag1', 'cls', 'mastermind', 'maml1', 'nam', 'kovall', 'joshi', 'palomero', 'hey', 'iso', 'jarriault', 'fryer', 'shuts', 'castel', 'ntziachristos', 'mizutani', 'sowilhol', 'sale', 'hes1gfp', 'fucosylation', 'lei', 'aifantis', 'boehmer', 'radtke', 'icn1', 'trb', 'capobianco', 'jme', 'sulis', 'alls', 'asnafi', 'lic', 'munster', 'breit', 'graall', 'clg', '58881', '58951', 'mrd', 'clappier', 'p2515rfs', 'fabbri', 'puente', 'wickremasinghe', 'oscier', 'weissmann', 'roulland', 'love', 'kiel', 'spen', 'dtx1', 'jundt', 'schwarzer', 'ebf', 'rangarajan', 'klinakis', 'attested', 'fx', 'fucose', 'fucosyltransferase', 'pofut1', 'aph1a', 'kannan', 'lobry', 'maml3', 'hes5', 'zweidler', 'mckay', 'kuang', 'nowell', 'ranganathan', 'robbins', 'colaluca', 'ayyanan', 'izrailit', 'meurrette', 'weijzen', 'harrison', 'cui', 'mazzone', 'eliasz', 'haruki', 'licciulli', 'hassan', 'agrawal', 'pickering', 'proweller', 'rothenberg', 'peifer', 'rudin', 'sriuranpong', 'kluk', 'salcido', 'fujimoto', 'guichard', 'alagille', 'mcDaniell', 'croquelois', 'viatour', 'dill', 'tschaharganeh', 'villanueva', 'jeliazkova', 'akhoondi', 'ortica', 'hepatocytic', 'clevers', 'babaei', 'jadi di', 'camps', 'miyaki', 'sancho', 'peignon', 'rodilla', 'jagged1', 'ridgway', 'sonoshita', 'pdac', 'miyamoto', 'mizuma', 'plentz', 'bailey', 'mazur', 'hanlon', 'asnaghi', 'bedogni', 'brn2', 'thurber', 'gsi', 'huynh', 'outstanding', 'pui', 'santagata', 'timmerman', 'sethi', 'brabletz', 'sahlgren', 'oskarsson'

n', 'mccaffrey', 'benedito', 'phng', 'gerhardt', 'noguera', 'troise', 'hoey', '2011c', 'niches', 'homing', 'blacklow', 'ferrando', 'goblet', 'ameliorate', 'asi', 'nrr', 'emanating', 'decoys', 'funahashi', 'hicks', 'moellerling', 'sa hm1', 'vlimas', 'cyclind', 'sawai', 'brd', 'bradner', 'thienodiazepine', 'fi lippakopoulos', 'kopan', 'stockhausen', 'invent', 'lobov', 'yatim', 'pu1', '0 3084014', 'omic', 'pdx', 'tnbc', 'gsis', 'site2', 'nicd1', 'hadju', 'cheney', 'hcc1599', 'stemelite', 'arizona', 'hpb', '3608', '2118', 'cerfe', 'xentech', 'hbcx', 'des', 'rinaires', 'minist', 'agriculture', 'che', 'ste123', 'legisla tion', 'fisheries', 'truong', 'maxf1162', 'oncotest', 'aa1077', '69504', 'pyr ophosphatase', 'tap', 'firstchoice', 'am1700', 'xenome', 'tophat2', 'hbcx14', 'prealigned', 'phs000178', 'v8', 'vendor', 'multiarray', 'summarization', 'co llapseddataset', 'nonaltered', 'blur', 'manufactured', 'cartridges', 'nsolve r', 'hes4', 'heyl', 'nrarp', 'phosphoribosyltransferase', 'peptidylprolyl', 'lollipop', 'unscaled', 'topnotch', 'nicd3', 'enst00000360016', '93rd', '97t h', 'tnbcs', '89e', 'nbpf8', '2462', '12b', 'aa0869', 'lighter', 'inappropriately', 'boost', 'leong', 'karsan', 'freeing', 'disintegrin', 'jk', 'kbf2', 'hamidi', 'ankryin', 'deltex1', 'malyukova', 'maser', 'chiang', 'reedijk', '251 5', 'suprabasal', 'demehri', 'dapt', 'stransky', 'zage', 'duality', 'commits', 'lics', 'moran', 'crusio', 'chromosomal', 'furinlike', 'noncovalently', 'serrate', 'lag2', 'dsl', 't6e', 'mastermindlike', 'propodium', 'draq5', 'dn d', 'lnr', 'ank', 'accentuates', 'misexpression', 'hox11', 'hox1112', 'tal1', 'lyl1', 'enl', 'calm', 'af10', '1574', '1622', '1575', '1594', '1601', '247 1', 'deltap', '2473', '2556', 'amyloidogenic', '2019', 'ese', 'eligibility', 'multigenerations', 'greece', '5331', 'engines', 'tcoffee', '3dcoffee', 'tunicate', 'sickkids', 'rulai', 'cshl', 'sgi', 'abcf', 'equivocal', 'tmlink', 'redesigned', 'tubule', 'lakhani', '5215', '8714', '5242', 'a1708v', '8237', 's 26701', 'tesoriero', '0010', 'incompatibility', 'asp1840', 'd1155h', 'r841w', 's1982t', '1225', 'del3', '1605', '2640', '1379', 's384f', '8795', 'maximizing', 's8691', 'g2274v', 'g1738', 'c1265s', '00001', 'wappenschmidt', '5331g', '5871', '8395g', '410', '4719', 'mircovic', '1738e', 'phosphospecificity', 'fruitful', 'eses', 'ask1', 'adipocytic', 'lipomatous', 'map3k', 'axioskop', 'lipomas', 'lp', 'map3ks', '1p36', 'stochastically', 'bohmann', 'angel', 'hattori', 'shaolian', 't10', 't11', 'ponceau', 'lpl', 't14', 't16', 'adipocytes', 'wisdom', 'tls', 'habener', 'poli', 'ccna2', '1974', 'homemade', 'preadipocyte', 'd6', 'nateri', 'toft', 'tominaga', 'grandson', 'v804m', 'c634r', 'aggressively', 'a883f', 'disclose', 'scraper', '852', 'emptive', 'distinctively', 'empowered', 'crlf2', 'il7r', 'sh2b3', 'charities', 'hyperdiploidy', 'hypodiploidy', 'tcf3', 'pbx1', 'ikaros', '1589', 'egas00001000654', 'estimator', 'p eto', 'p2ry8', 'lnk', '1241', 'cells30', 'ik6', 'rcsd1', 'zmiz1', 'ssbp2', 's 17a', 'ima', 'tinib', 's20', 's18', 'bcp', 'bfm', 'dcog', 'coall', 'finnzymes', 'goethe', 'vienna', 'pervanadate', '2025', 'supermix', 'plus2', 'miami', '6217', 'moderated', 'k67', 'op', 'jh6', 'hcx', '6x', 'boxplots', 'autophosphorylating', 'sediment', 'meat', 'qqtof', 'ida', 'acetaminophen', 'ibuprofen', 'salicylic', 'rsd', 'muc6', 'portrait', 'demethylated', 'pold3', 'helicobacter', 'genes12', 'oncosnp', 'cancer15', 'til', 'knot', 'proclivity', 'frizzled', 'pnt', 'ctnnal1', 'cancer19', 'acvr2a', 'lines23', 'cancers25', 'cdx2', '12c', '27632', 'evading', 'prevailed', 'gistic2', 'mucosae', 'bnip3', 'tp7 3', 'as1', 'ascl2', 'interconnecting', '19a', 'epigenome', '22b', '069', '23 a', 'intramucosal', 'gly17val', 'angioimmunoblastic', 'analyzable', 'waived', 'subjecting', 'eber', 'studies1', 'cleaving', 'purchase', '3165', 'quad', 'lrr', 'aroma', '984', 'pathway1', 'map2k', 'strategic', 'methodspatient', 'data clinical', 'lad', 'ingram', 'vicc', 'waivers', 'profilingtumor', 'k57n', 'studiesfunctional', '14746', 'p90rsk', 'thr359', 'ser363', 'rsk1', 'analysisfisher', 'resultsmutation', 'screeningamong', 'errb2', 'mutationsamong', 'm146i', 'g301x', 's331r', 'g13c', 'f531', 'e102', 'i103del', 'casescharacteristics', 'mutationsthe', 'mutationscomparisons', 'subsetsclinical', 'analysisearly', 'cancersmedian', 'mutantsto', 'ei102del', 'e203k', 'discussionsomatic', 'fis

chmann', 'f53', 'k57', 'd67', 'y130c', 'm146', '325901', 'nonrecurrent', 'botleneck', 'furthering', 'fat4', 'dsc1', 'fatalities', 'melanomas4', 'sequence d5', 'lau50', 'translesional', 'polymerases7', 'lau149', 'lau618', 'dipyrimidine', 'cells5', 'melanomas9', 'survival9', 'lau63', 'lau165', 'gly593ser', 'leu597arg', 'gln61arg', '610g', '622g', 'glu207lys', 'tumor12', 'glu203gln', 'pathway13', 'pro124ser', 'phe53leu', 'asn382his', 'ser154phe', 'pro124gln', 'pro124leu', 'phe57', 'syndrome16', 'ser150', 'tyr155cys', 'database4', 'lau108', 'reduced9', 'downregulator', 'melanoma19', 'lys22arg', 'tumor20', 'arg24cys', 'tumors21', 'map2k1e203k', 'map2k1p124s', 'proliferation22', 'map2k2e207k', 'vary17', 'kinase17', 'arg49', 'lys203', 'e207k', 'pathway26', 'hyperploid', 'analyzed9', 'cys275tyr', 'polyphen27', 'tp53bp1', 'lys803thr', 'rad23a', 'silenced9', 'integrity28', 'hdac3', 'expression9', 'stability30', 'dchs1', 'planar', 'cancer32', 'adenocarcinomas36', 'ldl', 'pathway37', 'lrp10', 'lrp1', 'lrp6', 'desmocollin', 'progression39', 'desmoglein', 'melanoma41', 'study6', 'dsc2', 'dsc3', 'respectively6', 'syndrome11', 'inhibitors42', 'sizeable', 'dormant', 'grin', 'plcb4', 'akt8', 'r38h', 'q546k', 'q546p', 'h710p', 't1025s', 'm1043i', 'm1043v', 'h1047y', 'rcas', 'transformers', 'eots', 'nononcogenic', 'nontargeted', 'e52k', 'e116k', 'p217k', 'g912r', 'd1016g', 'e542', 'q546', 'p539', 'd1045k', 't1025', 'q1033k', 'd1017g', 'e576k', 'd1046k', '165360', 'untypical', 'martinelli', 'niemeyer', 'caligiuri', 'sargin', 'makishima', 'muramatsu', 'dikic', 'postendocytic', 'piper', 'lechner', 'highway', 'epsin', 'sorkin', '005188', 'cblwt', 'cblc381a', 'cblk382e', 'cbl390y', 'cblr420q', 'mirko', 'sarah', '021009', '17608', 'lbe', 'turbofect', 'taufkirchen', 'desoxycholate', 'schwalbach', 'mapk3', 'nonclathrin', 'nc', 'cme', 'sigismund', 'madshus', 'stang', 'haglund', 'dmire2', 'egf488', 'deubiquitylating', 'rsb', 'anovas', 'pror', '688', 'd390y', 'tegf', 'pcpt', 'immuncytochemical', 'immunocytochemical', 'swaminathan', 'tsygankov', 'p60', 'p240', 'hyperstimulation', 'dhillon', 'rasopathy', 'pathomechanistic', 'derange', 'q367p', 'sanada', 'commit', 'tsg101', 'bache', 'allanson', 'nadeau', 'slavotinek', 'biecke', 'genin', 'phenocritical', 'unclarified', 'pharmacogenomic', 'zfs', 'insulators', 'multivalent', 'insulation', 'plk', '16q22', 'insulator', 'pediatrics', 'bergoni', 'gchgcaattgtataactgactcaactatagcgaggaggttttaccatggaaaggtaatgcagtcgaagccattgt', 'spacer', 'reticulocyte', 'k344e', 'h345r', 'r339w', 'r448q', 'c2hc', 'zf3', 'zf7', 'pcite', '11zf', 'recreatin', 'zns04', 'deoxyinosinic', 'deoxycytidylic', 'ctss', 'rw', 'banner', 'arfproto', 'pcat2b', 'phiv', 'af145468', 'af145477', 'tfiiia', 'nullizygous', 'interrogates', 'globe', 'intravascular', 'coagulation', 'foundationone', 'supracilavicular', 'miliary', 'subcarinal', 'hip', 'headache', 'slurred', 'drag', 'egfr1858r', 'nr6', 'egfrwt', '1068', 'geometrically', 'assaysfor', 'z1', 'sextuplicate', 'fusionthe', 'filament', 'portability', 'accountability', 'hipaa', 'replications', 'normality', 'cropped', 'illustrator', 'chaperonin', 'radford', 'dobson', 'crowding', 'vectorial', 'frydman', 'cylindrical', 'hartl', 'hayer', 'catalysts', 'chaperonins', 'membered', 'bukau', 'horwich', 'groel', 'nonnative', 'archaeabacteria', 'leroux', 'thulasiraman', 'gavin', 'pappenberger', 'quasi', 'llorca', 'willison', 'rommelaere', 'melville', 'ohh', 'clifford', 'tbd', '1213', '315kb', 'l158p', '163', 'e204', '143a', '147a', 'w117', 'l118', 'f148', 'i151', 'l153', 'v155', 'chaotropic', '289kb', 'y112h', 'g114r', 'l118p', 'a149p', 'n150k', 'i151s', 'w117c', 'l118r', '484kb', '630kb', 'labile', 'gimc', 'elutes', 'vivohek293', '404kb', 'denaturant', 'duan', 'unassembled', 'conformers', '439kb', 'translations', 'projecting', 'farr', 'vitroin', '191kb', 'cdc20', 'camasses', 'plaxco', 'swapping', 'resonates', 'large', 'rstested', 'peltoma', 'misincorporations', 'yamane', 'modrich', 'niels', 'wind', '9600', 'lumc', 'hays', 'mismatching', 'nicked', 'lnfam', 'jiricny', '371c', 'pasching', 'pcite4a', 'beeston', 'cite', 'optiquant', 'bsrbi', 'peakscann', 'lajolla', 'mappmmr', 'pjhg3', 'pcrbased', 'chromium', 'liacs', 'lovd2', 'pathologyb', 'mmrd', 's44f', 'g67r', 's93g', 'i219v', '655a4g', 'r265c', 's406n', '1852', 'webpage', 'o0', 'blueankh', 'shimodaira', 'spina', 'trojan',

'twostep', 'tournier', 'inflict', 'fanci', 'heptapeptide', 'disabling', 'pc130', '9010', 'wh0051755m2', 'h9010', '12ca5', 'tetramethylrhodamine', 'teva', 'veliparib', '38101x', 'cttacgcugaguacuucga', 'gugggguguuggacagugua', 'gaaagaa gacaaacagaaa', 'gagacuagacaauugagaaa', 'gaagaacagtattgacata', 'gtagatacaagtaga gaat', 'scudierio', 'ascei', 'bc140854', '28102', 'psfb', 'appends', 'bc129121', 'grays', 'radsource', 'cdk12s', 'k756r', 'd877n', 'k975e', 'autorad', 'g909r', 'melphalan', 'p12661', 'biallelically', '20101', 'hgs', 'ovca', 'months4', 'aberrations8', '489', 'capture9', 'commentsthis', 'trials11', 'tiers', 'certification', 'producedfull', 'tablefigures', 'indexnext', 'tablemutation', 'analysisabstract', 'commentswe', 'totalling', 'gigabases', 'results13', 'fat3', 'gabra6', 'arg8821leu', 'tyr901cys', 'lys975glu', 'leu996phe', 'ovcaf ull', 'tableprevious', 'indexwe', 'man18', 'asn581ser', 'glu545lys', 'his1047 arg', 'gly12asp', 'swissprot', 'chasm19', 'commentssomatic', 'aberrations21', 'previously22', 'tumours47', 'gistic21', 'zmynd8', 'irf2bp2', 'id4', 'drugbank', 'huex', 'genes24', 'factorization', 'observations25', 'cxcl10', 'cxcr3', 'muc1', 'mcm2', 'slpi', 'myofibroblasts', 'angpt12', 'angpt11', 'controls26', 'amt', 'ccl21', 'sparcl1', 'rab25', 'cancer27', 'rand', 'sets25', 'signature31', 'diseaseabstract', 'commentsseveral', 'subnetworks', 'network32', 'hotnet33', 'samples34', 'hotnet', 'foxm1', 'spoke', 'indexpublished', 'inhibitors35', 'emsy', 'c11orf30', 'reported40', 'paradigm41', 'database42', 'aurb', 'aurkb', 'ccnb1', 'damage43', 'tissue44', 'ovca22', 'discussionabstract', 'glioblastoma47', 'mutations48', 'mutations51', 'aetiological', 'chasm20', 'artefacts', 'approaches47', 'cyck', 'rnapii', 'rbp1', 'isomerizations', 'pausing', 'tefb', 'cyc', 'ctk1', 'lsk1', 'ctk', 'incapacitated', 'insult', 'prev', '44343', '37600', 'cdk13', '72835', 'rcrargrarurgrarcrcrururgrarargrcrurud tdt', 'figure5c5c', '37007', 'cagcagaacagacgacactcgaggcttcaagcacaccag', 'ctgg tggcttgaaggcctcgagtgtcgctgttctg', 'wces', 'nbp1', '06519', 'novus', 'glut atione', 'bioer', 'ab5095', 'psliib', 'coa', 'ddh2o', '32643', 'rnazol', '480 ii', '23921', 'chek1', '8408', '366363', '32233', 'watt', 'magnet', 'ab5731 1', 'ab5408', '60th', '100th', 'mx3005p', 'acuri', 'statistica', 'welsch', '4nst', 'l122fs4', 'q602', 'w719', 'e928fs27', 'r8821', 'y901c', '1996f', 't1014', 'q1016', 'k975', 'r882', 'y901', '1996', 'reorients', 'isolations', 'prm', 's325n', 's352f', 'e1024', 'figure1d1d', 'l122fs', 'eager', 'void', 'prohibiting', 'spacious', 'demolishes', 'pt893', 'organizing', 'figure2d', 'figures3', 'hiv1', 'sliib', 'tethers', 'mms221', 'nek9', 'orc31', 'terf2', 'figure4d', 'figure4e4e', 'ccnk', 'figure5b5b', 's9d', 'tsss', 'figure5b', 'figure5c', 'figure6b', 'incapacitate', 'cki', 'cycc', 'cyct', 'karnitz', 'transcribing', 'icls', 'nucleoprotein', 'revelations', 'equips', 'stall', 'malfunctions', 'g464e', 'n486', 'p490del', 'q201h', 'dysmorphia', '1391', 'i220i', 'rs10250', 'g634g', 'rs9648696', 'd151d', 'rs17851657', 'q113q', 'dysmorphisms', 'q257r', 'q210h', 'artificially', 'pathogenetics', 'tov', '112d', '21g', 'a1847', 'peo1', 'drescher', 'beatrice', '030662', 'gtaaaaacgacggccagt', 'caggaaa cagctatgacc', 'abi3730xl', 'jsnp', 'sprot', 'histiocytoses', 'therapeutisc', 'erdheim', 'jxg', 'dorfman', 'rdd', 'histiocytes', 'ddpcr', 'rnf11', 'clip2', 'lmna', 'cd207', 'cdcs', 'enrichments', 'irf7', 'runx3', 'gpr82', 'ccr7', 'baalc', 'id1', 'ill1', 'map2k1k57n', 'anakinra', '2mg', 'spermatic', 'map2k1q56p', 'infiltrations', 'orbits', 'cavernous', 'sinuses', 'arafs214a', 'clofarabine', 'arafs214c', 'retina', 'fundoscopy', 'taper', 'phagocyte', 'granulomatous', 'lesional', 'unamenable', 'ntrk', 'overhaul', 'mandating', 'piti', 'salp', '250ng', 'duty', 'hifi', 'qubit', 'fluorimetric', '300ng', '06465684001', 's214a', 's186r', 'a225v', 'p539h', 'r603q', 'q58del', 'q58', 'e62del', 'f681', 'hex', 'qx200', '30ng', 'cartridge', 'fpkm', 'fpkms', 'cuffdiff2', 'histiocyte', 'nebnext', '10ml', 'streck', '640t', 'infinite', 'ramp', 't100', '5ng', 'quantasoft', 'bmpr1a', 'cdc73', 'ptch1', 'sdha', 'sdhaf2', 'sufu', 'tmem127', 'gnb1', 'bronze', 'asu', 'consulted', 'lectinib', 'migii', 'msigdbv4', 'srp065600', 'gse74442', 'pvhls', '1186', '1786', 'pcc', 'luminol', 'pet28a', 'tag1', '28a', 'tween20', 'gsh', 'odd', 'vcb', 'ph7', 'r167q', 'l18

8v', 'n78s', 'andb', 'k171', '1158', 'r167', 'popular', 'proteasomes', 'melanomaassociated', 'unhelpful', 'hugoapproved', '600160', 'alcorta', 'haferkamp', 'mcconnell', '301c', '421c', 'artefactual', 'helen', 'westmead', '2145', 'wmi', 'usyd', 'coffee', 'igs', 'mrs', 'z0', 'greenblatt', 'pq', 'scoreo60', 'scorez60', 'zc25', 'wmm1175', 'r24p', 'nm39', '1105', 'pvp16', 'spectrin', '401c', '41c', 'completetm', 'prb807', '9308', 'bf683', 's8dup', '243del19', 'a73a', 'r144c', 'o85', 'm53i', 's56i', '4273', '4973', 'g67s', 'n71k', 'v126d', 'mckenzie', 'postinduction', 'b62pw', 'scoreb', 'gvgdc', 'scoree', 'scoref', 'l16p', 'g23d', 'l32p', 'g35a', 'a36p', 'i49s', 'a60r', 'a681', 'a68tg', 'e69g', 'n71i', 'n71s', 'p81t', '197r', 'r99p', 'r112g', 'r144cg', 'a148tg', 'anumber', 'bb62pw', 'dprediction', 'eblosum62', 'fgrantham', 'o60', 'gombart', 'hussussian', 'cespedes', 'a68t', 'piccinin', '695', '3374', 'localizationc', 'indexd', 'bindinge', '0f', 'speckling', 'weaver', 'feldhaus', 'freedberg', '2001c', 'pegfpspectrin', 'cellspercentage', 'mean7standard', 'sphase', '697', 'mean7sd', 'newton', 'isoc2', 'mislocalization', 'malhotra', 'debnia', 'confidently', 'alas2', 'slc25a37', 'prpf8', 'hnrrnpd', 'bclaf1', 'splicesomal', 'branchpoint', 'tf1', 'skm1', 'gillingham', 'stealth', 'pmaxgfp', 'hemin', 'hbg1', 'cd36', 'cd235a', 'retroscript', 'ftmt', 'klf1', 'lewes', 'ivt', 'fastdb', 'stk6', 'hitchin', 'handle', 'edger27', 'dexseq28', 'supplementary', 'gse63569', 'g2m', 'loc100505759', 'mkrn1', 'wrnip1', 'zc3h7a', 'crebzf', 'sgk494', 'wsb1', '1419', 'srsf11', '0027', 'abcc', 'ncifcrf', 'eder', '1177', 'glrx5', 'alad', 'fech', 'urod', 'hbq1', 'hba2', 'hbb', 'polo', 'tnfr2', 'analysisand', '3506', '1924', 'ezh1', 'epb42', 'rbm25', 'prpf40a', '3097', 'ndufaf6', 'ppox', 'hmbs', 'uqcc1', 'biogenesis31', 'numa1', 'chuk', 'pias1', '0498', 'bcl2a1', 'gypb', 'hbbp1', 'importer', 'exporter', 'sideroblast', 'map3k8', 'clec5a', 'cbl1', 'missspliced', 'misslicing', 'exacerbation', 'u2af145', 'shedding', 'ofmds', 'states2', 'level5', 'graduate', 'hongo', '8655', 'shirokanedai', '8639', 'manheim', 'theodor', 'kutzer', 'ufer', '68167', 'nakuchi', 'tennodai', 'ibaraki', '8571', '10division', 'ohtsuka', 'toshima', '11division', 'fujigaoka', 'kanagawa', '8501', '12munich', 'lebsche', 'platz', '81377', '13hematology', 'cedars', 'sinai', '90048', '14national', '117456', '15division', 'gung', 'tung', 'hwa', 'u2af35', 'sc35', 'u11', 'u2af65', '39ss', 'urp', 'u1snrnp', 'yncuray', 'ryyyry', 'uhm', 'sf3b', 'u2snrp', '59ss', 's34y', 'q157r', 'q157p', 'xp22', 'i202n', 'n261y', 'c302r', 'h330r', 'n382k', 'i53t', 's40x', 'r68sp', 'a96fs', 'e118fs', 'n126x', 'l237fs', 'k257sp', 'w291x', 'g323fs', 'n327fs', 'f239v', 'e148x', 'e362x', 'e133g', 'c326r', '12q13', 'surf', 'ubql', 'a57s', 'i141m', 'y772c', 'm117i', 'k166t', 'e373d', 't374p', 'm667v', 'y347x', 'a26v', 'p383l', 'p15h', 'm58i', 'p2121', 'd442n', 'p540s', 'e622d', 'k182e', 'g347v', 'r18w', 'm144i', 'l187v', 'g372v', 't454m', 't474a', 'y476c', 'a508g', 'mutationontaster', 'anaemia23', 'instability24', 's34', 'q157', 'h662', 'e622', 'welldocumented', 'build131', '39sss', 'appreciate', 'colours', 'themutant', 'u2af35was', 'cooverexpression', 'thefunctions', 'translation27', 'mutantu2af35', 'withwild', '15a', 'smg5', 'dhx34', 'upf1', 'smg1', 'upf3b', 'btz', 'upf2', 'smg7', 'smg6', 'smg9', 'magoh', 'smg8', 'y14', 'eif4a3', 'pym', 'upf3a', 'hnag', 'antiu2af35', 'antiflag', 'mockwt', 'cnormalized', 'intergene', 'usedfor', 'themann', 'cd342ckit1scaillin2', 'cd342ksl', 'marrow28', 'chimaerism', 'egfppositive', 'cellswasassessed6', 'byflow', 'typeu2af35', 'unmasked', 'unquestionable', 'homostasis30', 'abnormalities31', 'neoplasms33', 'mutations35', 'unlimited', 'proliferation1', '5q2', 'myeloproliferation36', 'annv', 'absorbance6', 'annv1', 'ksl', 'grubbs', 'methodsvii', '9211', 'fgf8', 'antihormone', 'ablating', 'proliferates', 'estrogenic', 'cryosections', 'endojournals', 'redmond', '5s', 'antirat', 'rocker', 'stereomicroscope', 'microimaging', 'thornwood', 'grid', 'fronts', '1c1c', '3a3a', '3b3b', '4a4a', '4b4b', 'nonadherent', '5a5a', '5b5b', '5c5c', 'dampen', 'researches', '7th', 'pathvysion', 'wiesbaden', 'bench', 'slips', 'pylorus', 'polypoid', 't4a', 'adata', 'tyr641', 'monomethylating', 'prc2s', 'enzymology', 'epigenetics', 'phf19', 'pcl3', 'lymphogenesi

s', 'velichutina', 'prc1', 'bmi1', 'hodgkins', 'adenosylmethionine', 'oligonucleosomes', 'exosites', 'nucelosome', 'exosite', 'methylations', 'monomethyl', 'trimethyl', 'lsmt', 'y245a', 'y245f', 'dimethylate', '1067', 'hydroxyl', 'attendant', 'simplifying', 'ly19', 'karpas', 'wsu', 'dlc12', 'foregoing', '004456', '153207', '003797', '005610', '51004', 'm250s', '21st', 'flashplate', 'bicine', 'bsg', 'smp410', 'propogation', 'filterplate', 'oligonucleosome', 'msfbn6b', 'nucelosomes', '2289', 'dh12', '2959', '1836153', 'bursts', '9733', '9755', '39377', '3638', '800cw', 'licor', '32212', '21076', 'ezh2y641', 'ezh2y641f', 'ezh2wt', 'nontyrosine', 'zest', 'hmtases', 'g9a9', 'ab1791', 'h3k9me3', 'ab8898', 'h4k20me3', 'ab9053', 'ab4469', 'ab6002', 'ab12073', '612667', '118p', 'antibiotin', 'orfeome', 'ezh2y641n', 'pcdna6', 'ezh2h689a', 'reinberg', 'psf9114', 'psf91', 'e86', 'gagagaacgcgttaaccatgtaccatacgacgtcccgactacgctatgggccagactgggaag', 'gagagacaattgaaccatgtaccatacgacgtcccagactacgctatgggccagactgggaag', 'gagagagtcgactcaaggattccatttctcttt', 'me3', 'blastidine', 'totalh3', '44n', 'dohh2', 'colloidal', 'destaining', 'sciox', 'via1', 'scintisafe', 'econo1', '1900tr', 'ehmt2', '208j', 'threaded', 'molsoft', 'llc17', 'dlcl', 'mutf', 'monospecificity', 'dlcbcls', 'md903', 'dul', 'toledo', 'deazaneplanocin', 'activity20', 'methytransferases', 'monomethyltransferase', 'setd7', 'data9', 'y1067f', 'trimethyltransferases', 'pdbid', 'me0', 'jmjd', 'allelism', 'ezh130', 'allele31', 'interactor', 'multicomponent', 'smurf2', 'h3k27me', 'ectoderm', '452', '180k', 'sudhl', '1428', '160k', 'stat5b', 'ppase', 'sh2b', 'phosphodegron', 'h4k20', 'h3k18ac', 'hypomethylated', 'h3k9ac', 'swine', 'streptavidine', 'lsab', 'unevenly', 'h3k4me', 'pt1a', '792', 'h3k18', 'methylases', 'jmjd3', 'jumonji', 'lengthened', 'hypoacetylation', 'kamangar', 'dores', 'wf', 'crossref', 'eac', 'engl', '2241', 'quarters', 'hj', 'surg', 'tomoda', 'aerodigestive', 'tomizawa', 'brennan', '15q25', 'neumann', 'schumacher', 'prasad', 'ahuja', 'hmlh1', 'esteller', 'adv', 'deacetylate', 'hyperacetylation', 'feng', 'yq', 'groudine', 'hmts', 'peterson', 'curr', 'euchromatin', '5616', 'ikeda', 'xiangya', 'changsha', 'jamieson', 'slideoriginal', '00kb', 'strip', '042', 'hdac5', 'suv39h1', 'lsd1', 'hmt', 'epigenomics', 'methylomes', '1262', 'ja', '4358', '4364', 'licht', '1872', 'bk', 'lia', 'o', 'jl', 'dev', '985', 'js', 'norman', 'shilatifard', 'klose', 'buttner', 'luscher', 'ash2', 'lan', '941', 'hypoacetylated', 'wh', 'hyperacetylated', 'ju', 'ng', 'jong', 'eun', 'sy', 'yoo', 'chiu', 'wilton', 'gerald', 'deacetylases', 'transl', 'fy', 'zz', 'dnmt', 'firestein', 'wc', 'plos', 'albany', 'biophys', '1804', '1689', 'lett', 'matsubara', 'matsushita', '65e', 'euploidy', 'codeletion', 'skmg3', '8mgb', 'dac', 'metaanalysis', 'q1481x', 'ptp', 'r', 'q987k', 'n1128i', 'and33c', 'ih', 'ihas', 'r1088c', 'p1311t', 'pyxxq', 'dbss', 'globocan', 'amifostine', 'ellipticine', 'prima', 't123p', 's240n', 't123a', 'rby377', '1cuas53', 'rby41', 'rby171', 'prb16', 'pflmi', 'prs413', 'cassettes', 'baby', 'dualluciferase', 'pp53', 'dbs', 'shortcomings', 'falsepositives', 'revertants', 'reisolated', 'f1131', 't118m', 'v122i', 'v122a', 'e285g', 's240r', 'gml', 'y205c', 'v173l', 'undervalue', 'staphylococcal', 'divert', 'superstable', 'n235k', 'chemists', 'spell', 'r172h', 'k132r', 'm246v', 'f270c', 'a276p', 'm237r', 't256a', 'iggo', 'isrec', 'epalinges', 'h193r', 'c238f', 'h1791', 'r249w', 'p1771', 'r283h', 'pflm', 'sgra', 'stu', 'ddh', 'mercapto', 'superatant', 'spo13', 'foa', 'v157f', 'c141y', 'p2781', 'lavender', 'specifies', '8298', 'shortcoming', 'recapitulation', 'weaknesses', 'aids', 'shiraishi', '3583', '1794', 'myh11', '1761', 'amlcg', 'qiаcube', 'bioruptor', 'diagenode', '50mb', '004119', 'humphries', 'terry', 'pmapk', 'cd135', '1762', 'genecards', 'multichip', 'irizarry', 'routines', 'biostatistics', '1763', 'catalase', '1764', 'ccna1', 'dqa1', 'lactam', 'proportionate', 'outcompeted', 'interobserver', 'u138mg', 'scd', 's23', 's29', 'pta', 'cytoscan', 'processes1', 'previously11', 'previously12', 'antigenically', 'm13f', 'genscript', 'aseptic', 'ewsr1', 'individualization', 'cd99', 'fev', 'mesoblastic', 'a673', '1094', 'debatable', 'collaborations', 'qualifying', 'ampliseq', 'pgm', 'torrent', 'defuse', 'alternated', 'nexus', 'biodiscovery', 'tryptase', 'gdan

sk', 'itching', 'cd2', 'anaphylaxis', 'abi3130', 'pmscvpuro', 'thefollowing', 'rps6k', 'scion', 'mcdonnell', 'leiomyomatosis', 'k509i', 'verstovsek', 'ustun', 'cd401', 'lmp1', 'nfkbie', 'to0', 'cyld', 'ptpn1', 'ontologies', 'muc4', 'cases6', 'pmbcl', 'dave', 'bjorkman', 'peaper', 'cresswell', 'townsend', 'zinkernagel', 'ctls', 'garrido', 'hicklin', 'jordanova', 'glycosylphosphatidyl inositol', 'dustin', 'springer', 'bolhuis', 'kanner', 'cytolysis', 'altomonte', 'gwin', 'billaud', 'daudi', '2126', 'hc', 'stam', 'reintroduce', 'colon a', 'takai', 'w6', 'parham', 'tyr119', 'ser98', 'tonsillar', 'hbl1', 'ts2', 'koopman', 'iceberg', 'riemersma', 'boon', 'bruggen', 'karre', 'steidl', 'ppp6c', 'h114y', 'micronucleation', 'kops', 'bischoff', 'karthigeyan', 'lentin i', 'dependant', 'stefansson', 'brautigan', '3429', 'hinted', 'saps', 'ankrd', 'ppp1r1', 'haften', 'manganese', 'saps2', 'ankrd28', 'prickett', 'nitrop henylphosphate', 'destabilises', '3430', 'openms', 'kohlbacher', 'ppp', 'saps1', 'asn113', 'his163', 'his237', 'asp53', 'his55', 'asp81', 'his114', 'nitrophenol', 'n53', 's167f', 'r264c', 's2701', 'k132n', 'p186s', 'p259s', 'g52k', 'h55y', '1110f', 'g112e', 'q189r', 'l305f', 'n52', '3431', 'sipp6c', 'sisaps', 'ppp6r', 'scx', 'mascot', 'ms2', 't288', '3432', 'z51152', '5768', 'mps1', 't481', 'molinari', 'tpx2', 'sipp6c08', 'numa', 'kettenbach', 'oa', 'pt288', 'pt481', 'dou', 'si08', 'sipp6r', '3433', 'ppp6', 'lb2518', 'upshifted', 'auroraa', 'mln8537', 'si08resppp6cwt', 'ppp6c08', 'sipp6c07', 'micronucleated', 'si08resppp6ch114y', 'multilobed', 'si08resppp6cpd', 'macville', 'rp e1', 'nonmodal', 'catalyst', 'inh', '3435', 'crasta', 'rogakou', 'ppp6ch114y', 'ch2ax', 'mln8357', 'si07', 'outset', '3436', 'congression', 'boveri', 'mcgranahan', 'micronucleus', '0166', '3437', 'undirected', 'raud', 'hoffelder', 'micronuclear', 'pinpoints', 'sporn', 'comediator', 'smad6', 'ssxs', 'kretzschmar', 'macias', 'heterooligomer', 'rsks', 'kuriyan', 'cowburn', 'heterooligomers', 'ser465', 'ser467', 'intein', 'chitin', 'thioester', 'spectroscopic', 'unligated', 'as3', 'speeds', 'ultracentrifuge', 'homoand', 'heterotrimerization', 'calonge', 'schutte', 'raxis', 'chess', 'rworking', 'h5', 's465', 's467', '1dev', 'klaulis', 'barried', '2600', 'homotrimeric', 'dpc4', 'cssms', 'nucleating', 'lys375', 'lys420', 'tyr426', 'arg428', 'heterooligomerization', 'met466', 'ser464', 'arg427', 'cys463', 'met327', 'thr328', 'his331', 'val419', 'arg310', 'asp450', 'asp300', 'tyr268', 'thr303', 'asp304', 'gly301', 'phe302', 'ser306', 'ser269', 'gln447', 'val319', 'tyr280', 'val431', 'thr432', 'trp437', 'overlaying', 'invariably', 'sma2', 'chromatographs', 'k375a', 'hydrodynamic', 'lys375ala', 'k428a', 'k420a', 'k507a', 'y426a', 'y513a', 'r428a', 'r515a', 'd300h', 'r310h', 'd351h', 'r361h', 'd450h', 'd537h', 'v319d', 'v370d', 'w368h', 'r420h', 'n387p', 'r441p', 'y406h', 'd493h', 'heterocomplexes', 'tendencies', 'stoichiometries', 'pseudophosphorylated', 'asp351', 'arg361', 'lys507gln', 'arg515gly', 'lys507', 'arg515', '145', 'huse', 'wieser', 'willis', 'tyr426ala', 'sara', 'tsukazaki', 'nucleate', 'sicheri', 'nonideal', 'ptxb1', 'biolab', 'gyra', 'mesna', 'mercaptoethanesulfonic', 'tne', 'spacegroup', 'i213', 'amore', 'navaza', 'rotational', 'optima', 'an60', 'ep on', 'centerpieces', 'sednterp', 'hid', 'storrs', 'h2347', 'h82', 'palbociclib', 'table22', 'figure3a3a', 'mention', '4562', '3247', 'hcc1195', 'sgx', 'hcc15', 'pen', 'hextuplicate', 'undruggable', 'unsolved', 'defeating', 'rasheed', 'fascinating', 'eckhart', 'fortuitous', 'thwarted', 'backup', 'holderfield', 'pathwayras', 'spred', 'gysin', 'energized', 'biopharmaceutical', 'rascental', 'wigler', 'taparowsky', 'painfully', 'sight', 'villalonga', 'potenza', 'balmain', 'lampson', 'buhrman', 'viewpoint', 'roock', 'a59t', 'shirouzu', 'anecdotes', 'remind', 'viciiana', 'inculpates', 'digit', 'marte', 'brems', 'ira1', 'sec14', 'stowe', 'lowy', 'feramisco', 'ying', 'shokat', 'ostrem', '4luc', 'dichlorophenoxy', 'piperidin', 'sulfanylbutanamide', 'f78', 'i100', 'recharging', 'hypervariable', 'alcal', 'isoformsthe', 'palmitoyl', 'picomolar', 'attacked', 'abandoned', 'maurer', 'trahey', 'mattos', 'cetus', 'infusion', 'phosphoroamidate', 'trick', 'rethinking', 'carboxymethylation', 'luck', 'geranylgeranylated', 'court', 'wahlstrom', 'depalmitoylation', 'thioesterase',

'apt1', 'palmostatin', 'delocalizes', 'zimmermann', 'bivona', 'lys147', 'lys104', 'sirt2', 'hdac6', 'nitrosylation', 'enos', '1990s', 'moodie', 'aelst', 'warne', 'thing', 'preincubate', 'foothold', 'coinfection', 'macdonald', 'hindsight', 'mirzoeva', 'turke', 'giroux', 'hatano', 'saba', 'leil', 'blasco', 'abou', 'fares', 'sobczak', 'gonz', 'lez', 'kashatus', 'tbk1', 'neel', 'tiam1', 'puyol', 'stk33', 'weiwer', 'nobody', 'crispr', 'payloads', 'deploying', 'ralgef', 'phosphoerk', 'mek', 'littleton', 'tif', 'myod1', 'myogenin', 'c2c12', 'oncoscan', '10e', '51mb', 'sbs', 'ribogreen', 'clipped', 'fastx', '50bp', 'clipping', 'chr22', 'mapq', 'bams', 'indelrealigner', 'recalibrated', 'tablerecalibration', 'unified', 'python', 'flagged', 'vcf', 'reordering', 'mpi_leup', 'intersected', 'jb', 'sms', 'ctr', 'fred', 'foreskin', 'dendograms', 'devising', 'clm', 'histo', '721', '1079', '4366', 'blm', 'flt4', 'disseminating', 'gil', 'retrieve', 'q472h', 'buster', 'po4', 'decorated', '55b11', 'patterson', '1942', 'lingen', 'punched', 'jc70a', 'envisiontm', 'acis', 'juan', 'digitally', 'cfx96', 'sbe', 'biosystems', 'calculator', 'pike', 'mafs', 'bioinformatics', 'ikzf2', 'confounder', 'amppnp', 'p123', 'pim2', 'pim3', 'u2s', 'nfatc1', 'untranscribed', 'qian', '1xzq', 'bilobular', 'capk', '1atp', 'e89', 'swayed', '1xr1', 'd186', 'n172', 'e121', 'o2g', 'od2', 'bulge', 'phkg', 'camki', 'ttn', 'sfrsk2', 'pkaca', 'ck1d', 'p123m', 'rendition', 'm123', 'e124', 'h68y', 'p81s', 'n82k', 'l193f', 'e135k', 'h68', 'y53', 'i66', 'v117', 'e35', 'l193', 'e135', 'n82', 'nd2', 'flips', 'f187', 'glu121', 'exocyclic', 'electronegative', 'k031', 'k063', 'oxindole', 'oxindoles', 'fgfr135', 'e562', 'e81', 'l83', 'campaign', 'bioactive', 'halogens', 'encloses', 'dihydroxy', 'quinolin', '144', 'v52', 'a65', 'l174', 'i185', 'untapped', 'tactic', '002648', '7169', 'pet29a', 'cobalt', 'argonne', 'mosflm', 'scala', 'program_s40', 'elves', 'resolve42', 'cnx', 'cns44', 'accelerys', 'refmac5', '715', 'redundancyb', 'completenessb', 'factorc', 'rfreed', 'discouraging', 'bronchoalveolar', 'puse', 'gliomas15', 't751inss', 'dsk5', 'tyrosine1068', '3afigure', 'tripled', 'ninth', 'gefitinib22', '4figure', 'erythroblastosis', 'recommendations21', 'undocumented', 'y1092', 'rc20', 'reexposure', '2233', 't751insi', '2229', 'reencoded', 's752insq', '2326', '2504', 'quit', 'and3b', 'cos1', 'hplid', '15of', 's0361', 's0388', 's0389', 'm95712', 'autoregulation', 'elrea', 'ir3t', 'ing', 'perception', 'arteriovenous', 'appearances', 'trails', 'cardiomyopathies', 'hypertensive', 'embryogeneis', 'brafq257r', 'brafg596v', 'gastrula', 'dorsalventral', 'mek1wt', 'mek2wt', 'hgti', 'hatching', 'varient', 'brafs467a', 's467a', 'g596v', '5402', 'and7a', 'natally', 'propel', 'fam131b', 'neurospheres', 'p61braf', 'akap9', 'agtrap', 'hrasv12', 'cytoselect', 'ltx', 'p221', '9b11', 'cignal', 'chromophore', 'k483m', 'illuminated', 'and4b', 'opto', 'unsaturated', '1100delc', 'i157t', 'froma', 'phylogenetically', 'plk3', 'machineries', 'nevanlinna', 'biomedicum', 'fin', '00029', 'falck', 'somaticmutations', 'allinen', 'vahteristo', 'lfl', 'siddiqui', '1q23', 'highrisk', 'multiplicative', 'cybulski', 'caligo', 'offit', 'bernstein', 'osorio', '860', 'oldenburg', '1035', 'broeks', 'kilpivaara', 's428', 'rad53', '1673', 'ancestries', 'walsh', 'slovak', 'goode', 'baysal', 'lipton', 'hofmann', 'sullivan', 'huzarski', 'nonbrca1', 'bartkova', 'nonproliferating', 'gorgoulis', 'cds1', '1848', 'epidemiologically', 'complexities', '2488', 'jerusalem', '948', 'tctt', 'disposable', 'r01', 't32', 'uncles', '6174delt', '10gy', 'talking', '608', 'r15', 'senesced', 'r14', 'arthritis', 'ab1', 'perp', 'ionized', 'figure11b', 'solo', 'atf3', 'interleukins', 'trp53', 'pqcxip', 'cartilaginous', 'amar', 'g97', 'g97d', 'bleeker', 'paugh', 'carbons', 'amines', 'morton', '236k', 'r100a', 'pusch', 'y139d', 'y139', '1t01', '132k', 'y179', 'dinucleotides', 'youssoufian', 'stereospecificity', 'kranendijk', 'neomorphs', 'redisssolved', 'butanol', 'v71i', 'v178i', 'hemerly', 'marucci', 'murugan', 'rakheja', 'f394', 'ailt', 'f394i', 'f394v', 'g70d', 'a134d', 'lemons', 'a134', '973', 'noushmehr', 'chondrosarcomas', 'urinalysis', 'sensitively', 'ser36', 'eda', 'ikbkg', 'met37', 'noninfectious', 'abscess', 'calmette', 'bcg', 'postvaccination', 'ivig', 'humoral', 'lymphedema', 'thrive', 'drb1', 'f

ludarabine', 'thymocyte', 'tacrolimus', 'gvhd', 'prednisolone', 'hemolytic', 'discontinuing', 'immunol', 'tetanus', 'toxoid', 'lymphotoxin', 'sk1', 'cd45ra', 'cd45ro', 'cd54', 'cd86', 'pbmc', 'cd14', 'tcell', 'cd27', 'dosedependent', 'tlr', '2246', '1699', '1171', '1290', '1238', '9800', '2457', 'ige', '1169', 'hypermorphic', 'teeth', 'petersen', 'fleming', 'hohenstein', 'lognormal', 'discriminatory', 'barker', 'overwhelmingly', 'deffenbaugh', 'classifiable', 'checks', 'aberrancies', 'cyclosome', 'smurf1', 'oscillating', 'shscr', 'nitrilotriacetic', 'noninfected', 'cdc27', 's2i', 's2j', 'hbl', 's3k', 'processive', 'inaccessibility', 'ferritin', 'aldolase', 'ovalbumin', 'refractive', 'harness', 's4i', 's4j', '4l', 's4l', 's4m', 'brafr509h', 's4q', 's4r', 's4s', 'brafe586k', 's5g', 's5h', 's5i', 's5j', 's5k', 'nonphosphorylatable', 's5l', 's5n', 'lee011', 'pd0332991', 's6f', 'codepletion', 'hydroxytamoxifen', 'sel', 's4o', 'bh2', 'bh1', 'smac', 'diablo', 'oligomerizes', 'tsp53', 'steven', 'jefferson', '9661', 'a1978', 'hsp60', '1722', 'grp75', 'lindsten', 'transcriptions', '1tsr', 'apbs', 'c277f', 'k120e', 'lp53', 'prescission', 'cysteinyl', 'bmh', 'delineating', 'oligomerizing', 'undertake', 'tomita', 'bh4', 'gateways', 'omm', 'apoptosis2', 'sequester', 'domain9', 'mdm210', 'protein11', 'binding12', 'proteins12', 'tad1', 'tad2', 'formation10', 'x118', 'endeavored', 'surface12', 'arrest21', 'senescence22', 'apoptosis23', 'proteins24', 'constructs26', 'c27', 'dbd28', 'reports18', 'trosy', 'previously14', 'others12', 'endothermic', 'calorimeter', 'exothermic', 't30', 'nitroxide', 'mtsl', 'cys2', 'cys122', 'ilv', 'protocol29', '1r2d', '2ac0', 'interaction14', 'reported12', 'dna33', 'glu153', 'asp158', 'glu180', 'lys20', 'asp186', 'leu137', 'met237', 'repositioned', 'x134', 'predicted35', 'reported37', 'function40', 'cancers35', 'h178d', 'lys120', 'e180k', 'i114a', 'k20a', 'e158k', 'reports41', 'mutation42', 'c277w', 'lys12043', 'cys277', 'interference', 'assay40', 'partners24', 'd186a', 'interactions43', 'unilamellar', 'luvs', 'groove29', 'luv', 'x148', '73749', 'his113', 'ile114', 'phe105', 'glu158', 'bax48', 'p5350', 'cerulean', 'domains12', 'reported18', 'dbds', '1821', '1076', 'p5324', 'p5313', 'domains33', 'overrode', 'exploits', '53bp143', '53bp244', 'sv4045', 'activity21', 'myst', 'hmof', 'tip60', 'debilitates', 'k120', 'k120r', 'hmdm2', 'moz', 'morf', 'kusch', 'legube', 'tyteca', 'berns', 'kitayner', 'acetylatable', 'hgcn5', '9k', 'ard', 'occlude', 'chymotryptic', 'peaked', 'deissler', 'p53er', 'chipuk', 'minimizes', 'zupnick', 'k120m', 'debilitated', 'villunger', 'doyon', 'acetylation', 'rea', 'akhtar', 'squatrito', 'aspp', 'aspp1', 'aspp2', 'friedler', 'bulavin', 'jack', 'knights', 'allis', 'tumors1', 'thr670', 'kinase23', 'syngenic', 't338i', 'patients13', 'cancer18', 'proteins20', 'invoked', 'pop', 'peyk3', 'peg4000', 'cryoprotected', 'soaks', 'beamlines', 'lb2', 'm14', '944', 'secrete', 'hnsccs', 'mm200', 'sequence', 'dideoxynucleotides', 'j02958', '2264', 'masa', 'y1253d', 'm1268t', 'labware', 'spheroid', 'musculoskeletal', 'y1248c', '1253', '1130', '1118', '1112', 'h1112r', 'v1110', 'h1112', 'cholangiocarcinomas', 'cytogenomics', 'bgj398', 'hf', 'k517r', 'ccl', 'milliliters', 'pt185', 'py187', 'calnexin', 'gm130', 'unreduced', 'endohf', 'shot', '80i', 'nis', 'ig3', 'endoglycosidase', 'oligosaccharides', 'zoomed', 'punctae', 'g271e', 'w290c', 'lipoatrophy', 'propnosis', 'underdeveloped', 'facies', 'hutchinson', 'guilford', '216400', '278700', '610651', '278720', '278740', '278760', '278780', '278750', 'atherosclerosis', 'progeroid', 'sclerotic', '0193', 'senile', 'corneal', 'sensorineural', 'hyperextensible', 'fractures', 'occipital', 'ofc', '45th', 'bitemporal', 'toes', 'palms', 'soles', 'bowed', 'thinning', 'phalanges', 'arachnoid', 'calcifications', 'flexion', 'tooth', '95th', 'mandibular', 'overlying', 'radii', 'pakistani', 'eczema', 'canthal', 'recessed', '12th', 'disproportionate', 'intellect', 'fluent', 'novocraft', 'gen', 'consistencies', 'd3a7', 'd13', 'sgc', 'hmlh3', 'porto', 'uf', 'paresis', 'compressed', '100b', 'gz', 'homepage', 'ema', 'hemiparesis', 'striatin', 'strn', 'ftc', 'pdtc', 'extrathyroidal', 'prescreened', 'paracentric', 'pptfig', 'lys1150', 'papillae', 'aminopyrine', 'diaminopyrimidine', 'sectionmaterials', 'amakye', 'patched1', 'derepr

essing', 'bccs', 'lbp', 'vismodegib', 'sekulic', 'gorlin', 'axelson', 'oro', 'gajjar', 'atwood', 'bestow', 'dijkgraaf', 'd473h', 'yauch', 'buonamici', 'm etcalfe', '114x', 'd473g', 'w535l', 'microfluidic', '2365x', 'ly2940680', 'd473', 'h231r', 'w281c', 'q477e', 'v386', 'v386a', 'q635e', 'w535', 'ayers', 'rond', 'riobo', 'v321', 'f460', 'l412f', 'katritch', 'f460l', 'h233fs', 's533n', 'dey', 'asz001', 'ato', 'arsenic', 'trioxide', 'decorate', 'sarin', '1412', 'prescreening', 'extraterminal', '18325', 'tailing', 'glm', 'dbsnip138', 'esp6500', 'ambry', 'seattleseq138', 'nucleofected', 'minikit', 'ch310t1', 'pgl3b', '6xglics', 'pcdhgfp', 'pcdhmcherry', 'sbi', 'cherry', 'serpentine', 'pastorino', 'pt12', 'prerx', 'pt10', 'pt09', 'mki67', 'pt06', 'pt08', 'hpi s', '0449', 'brinkhuizen', 'pricl', 'kandoth', 'rez', 'bigas', 'sweeney', 't241m', 'i408v', 'a459v', 'c469y', 'brastianos', 'pt03', 'pt04', 'pt11', 'stjernqvist', 'w281', 'i408', 'c469', 'dbp', 'h470', 'v404', 'c281', 'pt02', 'm321', 'v408', 'overestimates', 'c3h10t1', 'sasai', 'finite', 'ptch1loxp', 'rosa26lsl', 'tdtomato', 'tomato', 'r400', 'e518', 'n521', 'sonidegib', 'lde225', 'n219', 'd384', 'n219d', 'd384n', 's387n', 's387', 'gether', 's616g', '376', 'e380', 'jq1', 'lackner', 'inukai', 'negrini', 'nedelcu', 'pesz', 'lessenin g', 'basaloid', 'pt05', 'das', 'thakur', 'lorusso', 'irbs', 'nct00607724', 'nct00833417', 'nct01367665', 'sint', 'augustinus', 'hum00069052', 'hum00050085', 'proteogenex', 'pcas', 'herv', 'klk2', 'gsp', 'pcr4', 'as2', 'w3', 'alias', 'g270', 'g268', '4r', 'multicolor', 'dec', 'consult', 'mmlv', 'chori', 'm incing', 'diagnostica', 'transactivators', 't25', 'thiazolyl', 'retrotransposon', 'hervk', 'unigene', 'alox15b', 'alox12b', 'territories', 'factories', 'urges', 'reconsider', 'tads', 'ggaa', 'predestined', 'juxtapose', 'pia', 'waiting', 'turp', '8p21', 'invasively', 'liposomal', 'expansions', 'fusions1', '1q32', '198k', 'gus', 'hprd', 'plg', 'mpin', 'hyperchromasia', 'dorsolateral', 'rpn1', 'gene25', 'guarantees', 'titrate', 'mittnacht', 'peeper', 'albanese', 'winston', 'aktas', 'kerkhoff', 'diehl', 'ciechanover', 'renny', 'glotzer', 'lanker', 'disappears', 'baldin', 'redirects', 'sapk', 'p54', 'k35m', 't286a', 'fig2c', 'coprecipitating', 'adventitious', 'immunodepletion', 'prephosphorylated', 'repurification', 'boudewijn', 'kauffmann', 'zeh', 'vanhaesebroeck', 'dudek', 'myrakt', 'coreceptor', 'chased', 'corollaries', 'kauffman', 'beals', 'deficent', 'wtgsk', 'coselection', 'boyle', 'tequation', 'cancels', 'dissipate', 'unmapped', 'extinguishing', 'periodicity', 'labaer', 'nishi', 'fornerod', 'reits', 'rubinfeld', 'yost', 'korinek', 'clock', 'hatakeyama', 'calbo', 'farkas', 'muraoka', 'collecchi', 'lovec', '1994a', '1994b', 'd1b', 'cumulatively', 'benzeno', 'gladden', 'adelaide', 'bani', 'hani', 'beerm', 't286r', 'delta266', 'delta289', 'moreno', '134k', 'te3', 'te7', 'te12', 'p287', 'doble', 'woodgett', 'p287a', 'kyse520', 'precipitating', 'd289', 'd292', 'dajani', 'd292a', 'd289a', 'gamma32', 'axin', 'gid', 'gid380', '139k', 'unprimed', 'hedgepeth', 'fraser', '292a', '113k', '172k', 'k114e', 'coqueret', 'lys114', 'ser780', 'kitagawa', '197k', '72k', 'approx4', 'canonically', 'haar', 'fiol', 'sears', 'hagen', 'ascribes', 'zwijsen', 'petre', 'sicinski', 'lands', 'bodrug', 'rimerman', 'relocalizes', 'stambolic', 'crm', 'npc', 'nigg', 'resnitzky', 'metabolize', 'p9', 't23', 'sevenfold', 'importins', 'exportins', 'neufeld', 'alternately', 'implicit', 'lees', 'rustgi', 'sterner', 'reste', 'biowhitaker', 'summers', '1393', 'hexahistidine', 'antipain', 'chelate', 'rangtp', 'ey', 'tumorigenicity', 'bard', 'cmv2', 'ab4', 'redistributed', 'reassembles', 'isopeptide', 'breakages', '1470', 'c393', 'c243', 'agtccgaaattctcgccagatagttctacc', 'agtccgaaattcactgatactgctgggtta', 'agtccgaaattcgtagtggtgtggga', 'pyfp', 'tgatgcgtcgacactgatactgctgggtta', 'tgatgcgtcgacgttagtggtgtggggaa', 'cesium137', 'gammacell', '591', 'baer', 'bl578', '33285', 'bx40', 'thaw', 'refrozen', 'rethawed', 'bromphenol', '1637', 'subfragments', '1315', 'es corted', 'n304', 'brcts', 'ab5', 'endog', 'saturable', 'nibrin', 'attract', 'mpes', 'receptorlung', 'cancermutationpleural', 'invades', 'lymphatics', 'fibrinogen', 'transudates', 'mpe', 'nonmalignancy', 'retrived', 'heparinised', 'agttgtggagccttacacc', 'taaaattgattccaatgccatcc', 'nm005228', 'lobectomy'

s', 'p771', 'h773insynp', 'h773y', 'k860i', 's720f', 'e746v', 'anorexia', 'pn eumonitis', 'dyspnoea', 'soh', 'matsumoto', 'cough', 'hilar', 'plain', 'clump s', 'scanty', 'napsin', 'pn2', 'axillar', 'correspondent', 'esx', 'involutio n', 'nitric', 'nos2', 'crsp3', 'drip130', 'sur2', 'nonsequence', 'ets1', 'pos its', 'f9', 'hcollagenase', 'n233', 'n270', 'elf3l142p', 'uesugi', 'n173', 't tgaggatccgcaggccaggtccctactac', 'cgagctctagcattttaggtgacac', 'tgdd', 'tgtg', 'c235', 'tgddr', 'c354', 'eppley', 'composer', 'sybyl', 'tripos', 'probabl e', 'gabp', 'pdef', 'elf5', '1awc', '1bc8', '1k79', '1yo5', '1fli', 'chemiflu orescence', 'ecf', 'odn', 'mutets', 'tggcgaggagttcctgtttcccttcggcgc', 'tggc gaggagttgaattcacccctctcgccgc', 'tagc', 'radiolabeling', 'preblocked', 'decros s', 'cepheid', 'superarray', 'tactagcggtttacggcgc', 'tcgaacaggaggagcagagacg a', 'immunoabsorbed', '8334', 'leu275', 'leu283', 'ser308', 'lys318', 'tyr32 6', 'tyr337', 'elf3i279p', 'elf3a312p', 'elf3k320e', 'elf3r331p', 'k320e', 's 330a', 'n270i279p', 'n270a312p', 'n270r331p', 'n270k320e', 'i279p', 'nonconse nsus', 'ser136', 'glu144', 'd133a', 'l142p', 'w137p', 's136a', 'unstably', 'a nisomycin', 'dorset', 'chalfont', 'backcross', 'coitum', 'synchronization', 'na4p2o7', 'restimulation', 'ani', 'harnessed', 'kendall', 'weir', 'kwei', 'returns', 'clarifying', 'titf1', 'nmf', 'cna', 'adler', 'midst', 'ji', 'riely', 'cordero', '8p12', 'h522', 'michel', 'ohgaki', 'dok2', 'niki', 'exemplifi es', 'prevails', 'postgerminal', 'guardian', 'centroblasts', 'tmd8', 'ly1', 'ly7', 'ly4', '00025', 'actn4', 'rasali1', 'fbxo32', 'pag1', 'phospo', 'count erbalance', 'cal27', 'sacramento', 'jsi', 'epitect', 'methprimer', 'mdanderso n', 'hemacytometer', 'and3d', 'devise', 'ip3', 'plgf', 'psoriasis', 'detach', 'carmeliet', 'eriksson', 'alitalo', 'neuropilin', 'migdal', 'gerber', 'plcgam ma', 'cunningham', 'extraembryonic', 'chemotactic', 'raftk', 'angelo', 'shibu ya', 'sinusoidal', 'hydrolyses', 'sn', 'dag', 'brock', 'cgmp', 'pericellula r', 'endothelia', 'kukk', 'chorioallantoic', 'erythroleukemia', 'hypoplasti c', 'g857r', 'karkkainen', 'yancopoulos', 'cediranib', 'ingelheim', 'fluosta r', 'bmg', 'labtech', 'biovision', 'emd1214063', 'ced', 'calu1', 'diametrical ly', 'hyperphosphatemia', 'neoangiogenesis', 'igi', 'igii', 'igiii', 'igiii b', 'igiiic', 'hpsgs', 'klotho', 'rsk2', 'spry', 'sef', 'xenogafts', 'figure1 1', 'portrays', 'pam50', 'k659e', '4p16', 'fgf18', 'evasive', 'fgfr2iiib', 'c learest', 'r251q', 'e219k', 'd530n', 'i642v', 'a648t', 'ridaforolimus', 'vali dates', 'toxicology', 'keratopathy', 'sulfatinib', 'nct01225887', 'nct0088817 3', 'levantinib', 'nct01441297', 'nct01529112', 'ly2874455', 'jnj', '4275649 3', 'diuretics', 'pyelonephritis', 'mucositis', 'cep8', 'nct01202591', 'qualify', 'nct01244438', 'mfgr1877s', 'cen8', 'empiric', 'cghs', 'impressively', 'table44', 'hyperphoshatemia', 'joe', 'faber', 'schwab', 'grail', 'ccs', 'su peroxide', 'dismutase', 'rhod', 'zcchc7', 'epicenter', 'tetsu', 'sansom', 'po dsypanina', 'huh7', 'hle', 'fgf15', 'leptin', 'figure3c', 'figure3d', 'figure 6f', 'figure7', 'preoperatively', 'y53h', 'e124q', 'e142d', 'cradle', 'r122', 'q127', 'e142', 'arrhenius', 'photomultiplier', 'pmtv', 'carboline', 'imidazo pyridazole', 'k00487', 'k018444a', 'k00207a', 'midpoints', 'gtot', 'fi1', 'ty r53', 'oe2', 'h11', 'gh2o', '1wild', 'thiogalactoside', 'energic', 'de52', 'nit ritiltriacetate', 'precasted', 'etch', 'concentrators', 'fplc', 'absorptivit y', '48930', '1at', 'spectrofluorimeter', 'cuvettes', 'thermostated', 'nonlin early', 'phosphoenolpyruvate', 'afrrewspgkeakk', 'thermocontroller', 'kmvalue s', 'coating', 'lifescience', 'programmable', 'decomposition', 'natick', 'de composed', 'gunfolding', 'quantifies', 'kunfolding', 'yn', 'dependences', 'y 2', 'fi2', 'tyr721', 'ontariochem', 'mscf', 'femto', 'nitrosourea', 'molecula rmd', 'peg400', 'monobasic', 'quanta', '502ay', 'imparted', 'liabilities', 'n 655', 'n655k', '4u0i', 'disengaged', 'v654', 'episode', 'infarction', 'compar ators', 'nct01874665', 'approx60', 'cases1', 'ser259', 'hcm', 'syndrome8', 'e ngendering', 'approx10', 'to1', 't599i', 's2571', 'p261s', 'asp486', 'thr49 1', 'nonsyndromic', 'improbable', 'l613v', 't491', 's259a', 's621a', 'd594v', 'described6', 'openbiosystems', 'karlsruhe', 'h2n', 'pgex4t', 'propionate', 'propane', 'villigen', 'isotherm', 'bifc', 'bimolecular', 'kerppola', 'waldb

ronn', 'excite', '4695', 'nbt', 'rafwt', 'jetpei', 'ps338', 'spr', 'cm5', 'bindings', 'v263a', 'setups', 's259f', 's259', 's621', 'p261a', '6a6a', 'millimolar', 'exiting', 'druggability', 'hdac4', 'discriminator', 'to3', 'profit', 'carm1', 'arg17', 'nitrogens', 'kix', 'ere', 'gelcode', 'destained', 'nh4hco3', 'savant', 'trifluoroacetic', 'bioworks', 'ser10', 'adomet', 'ps2', 'prmt1', 'prmt5', 'broadens', 'reused', 'proinflammatory', 'ikk2', 'antiinflammatory', 'krasg12dpdx1', 'pdacs', 'panins', 'ikk1', 'ikk3', 'nemo', 'g12dpdx1', 'g12dikk2f1', 'flpdx1', 'krasg12dikk2', 'g12dtnfafl', 'krasg12dtnfa', 'tnfa', 'jci45797ds1', 'figure1f', 'figure1g', 'creas', 'igfr1', 'rtnf', 'figure3b3b', 'figure3d3d', 'figure3e3e', '1685458', 'il1b', 'figure4c4c', 'figure4d4d', 'figure6d6d', 'bay11', '7082', 'figure8', 'figure8d', 'deubiquitinase', 'fueled', 'pancreatitis', 'thiazolidinedione', 'metformin', 'tnfafl', 'ikk2f1', 'svjaa', 'toby', 'luminy', 'diterran', 'transnetyx', 'flmx1', 'difluorophenacetyl', 'phenylglycine', 'dl1', 'trichrome', 'immunolabeled', 'bm8', 'sangram', 'sisodia', 'marc', 'montminy', 'salk', 'bonferonni', 'vgp', '4182', 'wnt5a', 'grm', 'jagged2', 'delta1', 'delta3', 'delta4', 'nic', 'maml', 'delte', 'x', 'competence', 'rgp', 'wm35', 'wm278', 'wm3248', '1205lu', 'wm164', 'w489', 'mcdb153', 'l15', 'overlain', 'py397', 'she78', '3900', 'ab6556', 'nonimmunogen', 'pseudotyped', 'nalon', 'tampa', 'multiplicities', 'inefficacy', 'photos', 'transductant', 'municipalize', 'pfak', 'fom104', 'guaranteed', 'pi3ka', 'kt', 'womb', 'tomb', 'dii4', 'articulate', 'duplicitous', 'variously', 'multi modular', 'iterated', 'lin12', 'supposes', 'k2555', 'r23', 'i1446', 'shd', 'e1447', 'q1734', 'bsu36i', 'jeff', 'zuniga', 'pfl', 'cker', 'aprotonin', 'r1609s', 'enclosed', 'l1575p', 'l1601p', 'l1594p', 'dl1', 'i1617t', 'v1677d', 'l1679p', 'shds', 'f1593s', 'v1577e', 'l1597h', 'r1599p', 'i1617n', 'i1681n', 'a1702p', 'i1719t', 'reposition', 'tempered', 'circumvents', 'dyscohesive', 'expendable', 'extrathymic', 'evolves', 'abovementioned', 'longterm', 'laparotomy', 'intraabdominal', 'feline', 'proximally', 'klf5', 'emphasise', 'recognise', 'skp', 'hypothesised', 'spruck', 'barrel', 'apex', 'absorptive', 'r479q', 'shall', 'apcmin', 'ser62', '610154', 'tec', 'ec3', 'homoeostasis', 'sw1463', 'c32', 'shown31', 'submaximal', 'guanylate', 'girdle', 'palisading', 'itraconazole', 'proactively', 'orchestrating', 'cdkis', 'nevins', 'trxg', 'hanson', 'paro', 'schuettengruber', 'invertebrates', 'bmi', 'intimate', 'trx', 'rowley', 'ayton', 'rabbits', 'indispensability', '2003b', 'af4', 'commonality', 'diversified', 'elutriation', 'cyclohexamide', 'prometaphase', 'fyrc', '2722', 'synchronize', 'thp', 'nomo', 'noncleaved', 'insightful', 'denominator', 'enu', 'dire', 'lohuizen', 'plans', 'hoxa7', 'hoxa11', 'trcp1', 'oligolectamine', 'mll1', 'labtek', 'alexa488', 'alexa568', 'octamer', 'phosphokinase', 'yano', 'forster', 'malformed', 't5', 'hemisphere', '128k', 'prophase', 'expedita', 'antagonistically', 'instant', 'krause', 'hirschsprung', 'hscr', 'm980t', 's765p', 'e921k', 'p9731', 'dopaminergic', 'chromaffin', 'parafollicular', 'neurturin', 'persephin', 'artemin', 'glycosyl', 'gfra1', 'neuroma', 'marfanoid', 'habitus', 'aganglionosis', '13q22', 'phosphatebuffered', '0016', '5085', 'gast', 'ptc2', 'nissui', 'kogyo', '2149', 'kbq', 'iwashita', '1114', 'faintly', '1015', 'r897', '27a', 'recklinghausen', 'neuromas', 'inguinal', 'gutmann', 'ballester', 'verlag', 'heath', 'poulet', 'anneals', '1212', '1155', 'n308', '1361', '4071', '1383', '4255', 'k1423e', 'hypermutable', '1362', 'taqi', '4267a', 'morcos', '1434', 'neuropathy', 'deformity', '78k', 'sknas', 'disease3', 'imd', 'laminin', 'carcinoma26', 'previously5', 'cb17', 'mu6pro', 'bbsi', 'adeasy', 'dr4', 'fasl', 'cd95l', 'r4', 'universitaire', 'vaudois', 'lausanne', 'chuv', 'nmvc', 'fluorogenic', 'imr', 'neuroectoderm', 'ganglioneuromas', 'ncam', 'kaufmann', 'fulda', 'apoptogenic', 'apoptosome', 'baylin', 'vita', 'vh', 'xs', 'shep', 'imr32', 'u251mg', 'd283med', 'h9', 'reh', 'intergroup', 'deisenhofen', 'benzyloxycarbonyl', 'bachem', 'trevigen', 'ulm', 'hyperfilm', 'fernandes', 'autoclave', 'e2f3', 'dek', 'cdh11', '29a', '29b', '34c', 'rassf1a', 'dapk', 'cdh13', 'o6', 'methylguanine', 'tff1', 'dkk1', 'tens', 'chromothriptic', 'nations', 'nephrotoxicity', 'cellulitis', 'beam',

'mef2', 'intersecting', 'r661w', 'lxcxe', 'p57', 'mad2', 'condensin', 'lh', 'degenerating', 'bipolar', 'mdmx', 'ophthalmic', 'congestion', 'onwards', 'c losing', 'confounders', 'cerevesiae', 'rad9', 'prepublication', 'montiero', 'hanafusa', 'pvt', 'u100', 'dyedideoxy', '373a', 'ochre', 'vista', 'bioplate s', 'peptone', 'YPD', 'decanted', 'lioac', 'hourly', 'hemagglutination', 'iod inated', '22c', 'q1756c', 'stop1829', '22d', 'luce', '33c', '2575delc', 'penu ltimate', 'substratum', 'roots', 'opacities', 'nicknamed', 'hypothesizing', 'cterm', 'liposome', 'keyhole', 'limpet', 'hemocyanin', 'silverscan', 'rxn', 'sorvall', 'figures33', 'submembranous', 'filopodia', 'microspikes', 'rufflin g', 'granules', '6figure', 'spikes', 'nonionic', 'phosphoimager', 'nur', 'ros enbaum', 'keratohyalin', 'relaxed', 'hayashi', 'takeshima', 'maeda', 'bretsch er', 'nhe', 'rf', 'lamina', 'zucman', 'p53c', 'm133', 'swung', 'y126', 's24 1', 'crevice', 'depriving', '6319', 'olive', 'mad1', 'asparagines', 'ptre', 'fluidics', '25mer', 'cgccggcagaccagcatgac', 'tggagactctcagggtcgaaa', 'ggcgt ttggagtggtagaaatc', 'ctgtccccctgccgtccca', 'cctatggaaactacttcctgaaaacaa', 'a cagcatcaaatcatccattgc', 'ccgggatgaccggagcacctg', 'tgagatgtacgaggtgcacaca', 'c gcacgctctgctgaatg', 'cccgagaggctttccgagtggc', 'ccgcccgtggacacagact', 'ttgcc gtcagaaaacatgtca', 'a427', 'f113c', 'scaber', 'r1101', 'e271k', 'v274f', 'k32 0n', 'h1931', 'scc25', 'r290', 'pc13', 'h211', 'lx1', 'lu135', 'g244c', 'ov 3', 'n87', 'mkn1', '118mg', 'r213q', 'katoiii', 'h69', 'sc263', 'silentfect', 'disulfophenyl', 'monosodium', 'kishida', 'dehydrogenases', 'redness', 'green ness', 'blc', 'map4', 'transrepressional', 'umuc3', 'pupupucwwgpypypy', 'tran sfac', 'msin3a', 'transducting', 'reorganizing', 'p27kip', 'hsmad1', 'crl4dca f1', 'promitogenic', 'mesotheliomas', 'okada', 's518', 'kissil', 'lallemand', 'cul4b', 'baits', 'meso', 'nonnuclear', 'neddylated', 'ddb2', 'csa', 'connel l', 'harper', 'crl4', 'dyrk2', 's518a', 'mesothelial', '1417', 'hexagon', 'de construction', 'epistasis', 'stoichiometrically', 'saturates', '1566', 'conse nsual', 'itga6', 'ilk', 'limk1', 'p130cas', 'lamellipodia', 'asynchronized', 'shuttle', 'curto', 'breadth', 'hijacks', 'scchn', 'tendon', 'aktser473', 'd ihydrochloride', 'denver', 'topographical', 'a702s', '748', 'tyrphostin', 'ef gr', 'notions', 'expecting', 'ttp', '0011', 'unvalidated', 'nowadays', '1799g t', '1800tg', 'conventionally', 'scalp', 'useless', 'glut4', 'tirf', 'tir', 'movie', 'swapped', 'rabgap', 'nonpermeabilized', 'hyperinsulinemia', 'speci alist', 'cranium', 'feeds', 'hnf4a', 'hnf1a', 'ketone', 'apposition', 'mutati ons9', 'bruton', 'cancers19', 'phds', '86k', 'crosstide', 'cells27', '125k', '0041', 'inconspicuous', 'gliosarcoma', '3350', 'hydrolytic', 'domain13', 'm ecp2', 'glycosidic', 'alkali', 'symmetrically', 'hemimethylated', 'jj', 'tp g', 'glycosylases', 'bandshift', 'preferring', 'repair24', 'muth', 'piperidin e', '17q12', '8q11', 'smgs', '1424v', 's478n', 'rpl22', 'lys15', 'subcohort s', 'reclassify', 'asns', 'infinium', 'mc1', 'mc3', 'supercluster', 'membersh ips', 'bcl7', 'types24', 'memo25', 'gsk3beta', 'a96g', 's403i', 'specialty', 'nt770a', 'gln257', 'nt158t', 'f53s', 'nt389a', 'nt170t', 'f57c', 'f57', 'v6 00eb', 'som', 'f53smek1', 'y130cmek1', 'wtmek1', 'k97mmek1', 's218d', 's222dm ek1', 'f57cmek2', 'wtmek2', 'k101mmek2', 's226dmek2', 'cataract', 'overlays', '7362', 'acvr1', 'capitalize', 'conceptually', 'noisy', 'stitching', 'api', 'humsavar', '20140807', 'pathogeneity', 'curate', 'github', 'siglist', 'fo lder', 'snri', 'convolution', 'oni', 'kbin', 'proportionality', 'tli', 'tni', 'nonmatch', 'bernoulli', 'thresholding', 'boyangzhao', 'targetid', 'nonmatche d', 'biotab', 'exalt', 'readme', 'fibrodysplasia', 'ossificans', 'progressiv a', 'fop', '100000', 'spe1', 'ryr2', 'melamed', 'comorbidity', 'comorbid', 'c aller', 'r552k', 'g434r', 'c441f', 'm239k', 'n233y', 'm269r', 'e846k', 'r552 s', 'col3a1', 'ehlers', 'danlos', 'figs44', 'agnostic', 'rupturing', '4t1', 'r241', 'k22m', 'smethod', 'doubles', 'bmp2', 'decidualization', 'inlung', 'inmelanoma', 'discovers', 'rigor', 'hypermorphs', 'activation7', 'dh', 'e13 9d', 'm504v', 'ile36', 'traverses', 'mutants3', 'malignancies19', 'palpebra l', 'fissures', 'arched', 'posteriorly', 'cryptorchidism', 'previously4', 'de formities', 'diathesis', 'dysmorphology', 'rras', 'eyebrows', 'accuprime', 'p

olyplus', 'ex6', 'occipitofrontal', 'cesarean', 'walked', 'unassisted', 'strabismus', 'marmorata', 'cafe', 'naevi', 'creases', 'joints', 'denied', 'pph2', 'chr2', 'al1', 'underexpressed', 'underrepresented', 'methylanthranioloyl', 'assisting', 'timescales', 'decisive', 'counteraction', 'krev', 'rap1', 'd32y', 'k39n', 'p42t', 'c77f', '7pik3cawt', 'rat1a', 'personalizing', 'artifactual', 'ccnd2', 'irak4', 'mountains', 'gse16619', 'toyama', 'lentivector', 'p200', 'scratched', 'scratching', '8q22', 'paragraphs', 'ncor', 'smrt', 'sagi', 'al a59', '1049', 'robocycler', 'polyhistidine', 'branson', 'mmedta', '1044', '927', 'phe958', 'coloredwhite', 'c3g', 'arf1', 'glu97', 'd28', 'hersey', 'hayward', 'vx', '11e', '137f5', 'akts473', 'aktt308', 'mcs', 'pro124', 'unexposed', '652', '6q27', 'cep110', '1for', 'erythrophagocytosis', '7t', 'fim', 'zmf 198', 'u937', 'kg1', 'hsb', 'molt', 'hl60', 'mia', 'paca', 'minnesota', 'marathon', 'montigny', 'bretonneux', 'gt11', 'nome', 'grenoble', 'dbest', '7760', 'meylan', '2628', 'pvu', 'morpholino', 'antibacterial', 'absorbency', 'oakville', 'eco47iii', 'lipofectin', 'interrupts', 'autoantigen', '3892', 'cenp', '1658', 'immunolocalization', 'pericentrin', 'midbody', 'ninein', 'cep250', 'guasch', '787', 'basophilia', 'bilineal', 'ems', 'retrospect', 'manthorpe', 'kjeldsberg', 'catovsky', 'posner', 'abruzzo', 'jotterand', 'bellomo', 'inhorn', 'awkward', 'scientifically', 'leukemoid', 'starry', 'paler', 'surround', 'clot', 'filing', 'perinodal', 'hilum', 'venules', 'subtotal', 'dusty', 'cd68', 'fgfr1op1', 'endogen', 'p11p22', 'fgfr1op2', 'tifi1', 'trim24', 'q37', 'lr rfp1', 'flightless', 'myo18a', 'cpsf6', 'mym', 'groth', 'lardelli', 'phosphotyrosines', 'fop1', 'fgfr1op', 'tifi', 'oligodimerization', '13q11', 'fostere d', 'kasyapa', 'popovici', 'proteomics', 'flj14235', 'trim4', 'roumiantsev', 'appellation', 'agerstam', 'richebourg', 'overemphasize', 'lbl', 'cytosis', 'paratrabecular', 'succumbing', '0167', 'climent', 'bcr1', 'kunapuli', 'fgf1', 'overemphasized', 'baselga', 'swain', 'burgess', 'neuregulins', 'leahy', 'dawson', 'ogiso', 'untethered', 'garrett', 'heterodimerization', 'jura', 'hobro', 'tzahar', 'yamauchi', 'gotoh', 'alvarez', 'koutras', '484', 'nonprotein', '1m6b', '3lmg', 'v104', 'a232', 'p262', 'g284', 'd297', 'g325', 't355', 's846i', 'e928g', 'cystine', 'alvarado', 'q271', 'e928', 'imce', 'abaco', 'perbb3', 'v104m', 'p262h', 'g284r', 'cw', 'dv', 'soule', 'nrg1', 'brummer', 'bundy', 'ocum', 'hoeflich', 'luciferace', 'a232v', 't389k', 'v714m', 'q809r', 'karunagaran', 'hegde', 'derberg', 'horn', 'jaiswal', 'junttila', 'unmarked', 'schaefer', 'mehd7945a', 'goodin', 'tmab', 'anit', 'ragweed', 'nonliganded', '1971', 'unbridled', 'leone', 'ink4c', 'ink4d', 'kip2', 'tgfbeta', 'ekholm', 'hypothetically', 'jamal', 'preudhomme', 'beck', 'hangaiishi', 'beato', 'lecouter', 'drexler', 'chim', 'uchida', 'yamada', 'kees', 'ramakers', 'dicciani', 'ohnishi', 'tien', 'kamijo', 'krimpenfort', 'bv173', 'granta', 'jadale', 'nb4', 'masuda', 'boussard', 'fenaux', 'lai', 'wattel', 'inactivations', 'skorski', 'hsiao', 'kawamura', 'sakashita', 'cesarman', 'nishimura', 'kaneko', 'bueso', 'kawamata', 'faderl', 'seliger', 'oculocutaneous', 'gatti', 'pl1', 'vorechovsky', 'stoppa', 'stankovic', 'tavor', 'shikama', 'giles', 'blobel', 'borrow', 'panagopoulos', 'preleukemic', 'adjei', 'pepb2', 'aml2', 'cbfbeta', 'eap', 'nucifora', 'romana', 'shurtleff', 'mclean', 'cave', 'castilla', 'horwitz', 'osato', 'imai', 'rana', 'pabst', 'madden', 'harrington', 'drummond', 'hewitt', 'englert', 'kts', 'nephroblastomas', 'wagr', 'aniridia', 'baird', 'denys', 'drash', 'nephropathy', 'carapeti', 'hosoya', 'cd38', 'simmons', 'schmid', 'enormously', 'methylthioadenosine', 'mtap', 'nobori', 'ikappabalapha', 'inokuchi', 'nakahara', 'straus', 'uninhibited', 'bub1', 'e88k', 'd9s171', '9p22', 'jetquick', 'genomed', 'pter', 'ninefold', 'cdkn2', 'snx31', 'stk19', 'contextualized', 'cdnk2a', 'nikolaev', 'appreciation', 'pleasance', 'cytidine', 'lohr', 'blistering', 'sunburns', 'garibyan', 'zaidi', 'derijard', 'infers', 'nonsilent', 'anticorrelation', 'nonexonic', 'aura', 'r264', 'd89', 'd89n', 'p901', 'nexin', 'band4', 'ghai', 'cully', 'conte', 'jaffe', 'eff ecting', '85c', 'rac2', 'p291', 'rhot1', 'p301', 'p29', 'rac3', 'e31', 'refer encing', 'vigil', 'g13i', 'e478k', 'emuss', 'dogma', 'cites', '182a', '1676

t', '626a', '181ac', '182ca', '38gg', 'invex', 'leverages', 'shendure', 'fullness', 'disorders2', 'bulbs', 'heads', 'd47a', 'e49a', 'bolstering', 'damperd', 'nanoclusters', 'testify', 'nuremberg', 'superiore', 'ircs', 'pcm1', '123k', '4056', 'strn3', 'cells12', 'katholieke', 'universiteit', 'jan', 'alk1', 'lmp', 'perforin', 'harvests', 'fosmid', 'abvd', 'roosbroeck', 'cy7', 'cd3e', 'cd8a', 'preliminarily', 'wi2', '4059', '300000', '200000', '50000', 'm12', 'n10', 'anymore', 'm10', 'm11', 'endoplasmatic', 'monotonous', 'doughnut', 'bony', 'nup214', 'lymphoma40', 'lls', 'bcl11b', 'ctcl', 'gata3', 'lmo2', 'nf2', 'sil', 'ccrf', 'loucy', '8402', 'zary', 'commonalities', 'ptcl', 'piperidinyl', 'suberoylanilide', 'hydroxamic', 'ptcls', 'indexing', 'overhangs', 'biospring', 'parse', 'cabig', 'caarray', 'gskdata', 'rearranges', 'mycosis', 'fungoides', '2024', 'adjoining', 'cd133', 'rejection', 'immunoregulatory', 'signified', 'cav1', 'il10', 'tfbs', 'phosphosulfate', 'th2', 'th17', 'nfkb', 'tcf1', 'pneumonia', 'multiorgan', 'leek', 'embnet', 'excesses', 'cd7', 'b41', 'totipotent', 'polycythaemia', 'thrombocythaemia', 'acml', 'eos', 'busquet', 'heiss', 'murati', 'torsten', 'andreas', 'alerted', 'peeters', 'lacrionique', 'najfeld', 'griesinger', 'cirmena', 'trilineage', 'dominguez', 'nebral', 'patnaik', '2q13', 'bilineage', 'coh', 'epoxide', 'striated', '1244', '1473', '1784', 'oligodeoxyribonucleotides', 'eu', 'mdm1', '12q15', '4b5', 'smp14', 'wax', 'diamino', 'axiocam', 'extrapolating', 'c33a', 'quintuplicate', 'refed', 'antivinculin', 'pab1801', '4b2', '2a10', 'maggiore', 'novara', 'vmercate', 'riboprobes', '1064p9', 'immortalizes', 'hosting', 'ugent', 'mcb', 'conceal', 'oca', 'luna', 'hypomorphism', 'ladanyi', 'leach', 'supernumerary', 'ancillary', 'explant', '1440', 'zack', 'marchetti', 'job', 'paraganglioma', 'electromobility', '986', 'lind', 'archer', 'subdividing', 'hemangioblastomas', 'spindly', 'pvhl30', 'pvhl19', 'attaches', 'hemangioblastoma', 'rcc10', 'miguel', 'esteban', 'imperial', 'supreme', 'psuperretro', 'gchgcccggatccgcaccatgccccggaggggcgagaactgggagc', 'ccgggctcgagtcaatctccatccgttgatgtgcaatgc', 'gchgcccggatccgcaccatggaggccggcggccgcggccgtgc', 'gatcccctggatcactccaggtagttcaagagactacacctggagtgatcaccattttgaaa', 'agctttccaaaatggatcactccaggtagtcttgaactacacctggagtgatcaccagg', 'na2vo3', 'bestatin', 'microcentrifuged', '11e12', 'gelmount', 'qe', 'viewing', 'diagrammatically', 'rc161', '2qw', '1188q', 'y98h', 'y98n', 'r167w', 'v84l', 'upheld', 'figure4b4b', 'and4e', 'y112n', 's65w', 'figures1b', 'figure9', 'figure10', 'cobblestone', 'egln3', 'culling', 'harten', 'taron', 'pt3n0m0', 'blend', 'genegel', 'genephor', 'tyr845', 'diluting', 'nichirei', 'bronchiolar', '3383', 'nonpapillary', '3972', 'ex18', 'balancing', 'obviating', 'rantes', 'atopic', 'dermatitis', 'cxc', 'ponent', 'milliplex', 'jmp', 'hyperlipidemia', 'appetite', 'brinkman', '0044', 'pulopustules', 'periungual', 'onycholysis', 'corneum', 'xerosis', 'th1', 'adiponectin', 'cachexia', 'cogent', 'rps', 'collectives', 'pts', 'dyenamic', 'pharmadx', 'conjunctivitis', 'synopsis', 'ctx', 'rtx', 'dcr', 'polyploidy', 'explanatory', 'microphotographs', 'del19', 'prescription', '2000s', 'fees', 'shipped', 'microdissector', 'lmd2000', 'qiaampl', 'egfr21', 'biologist', 'shipping', 'v16', 'rareness', 'ndm', 'uke', 'folliculitis', 'decalcification', 'requesting', 'falsely', 'scorpion', 'reassess', 'prescriptions', 'scrupulously', 'inequalities', 'biomarqueur', 'ifct', 'nonrandomized', 'accrue', 'consort', 'itt', 'turnaround', 'signing', 'transitioned', 'centimeters', 'benefitted', 'interruptions', 'mandate', 'a859t', 'targetedagent', 'boramae', 'cbc', 'tccaaatgagctggcaagt', 'tcccaaactcagtgaaacaaa', 'atgtggcaccatctcacaattgcc', 'cacacagcaaagcagaaactcac', 'gctcagagcctggcatgaa', 'catcctccctgcatgtgt', 'tgaa gcaaattgccaagac', 'tgacatctcaggatgc', 'microwaving', '2237a', 's752insv', 'awbf', 's752insa', '2251', '2239t', 't751insp', 't751insa', 'cea', 'pbac', '2156g', '2125g', 'e709k', '2575g', 'carcinoembryogenic', '2709c', 't903', 'aggravation', '093', '052', 'reaped', '2155g', 'a859', 'extractable', 'nonmutants', 'adjunctive', 'arteaga', 'sordella', '82k', 'laci', 'pless', '1621', 'bitransgenic', 'disengaging', 'medeiros', 'hemangioma', 'chesi', 'richelida', 'cappellen', 'fgfr2k', 'fgfr3k', 'tdi', 'wilkin', 'calvert', 'vikkula',

'lymphoedema', 'sutures', 'ravel', 'mcgillivray', 'zankl', 'k526e', 'n549t', 'e565a', 'e565g', 'k641r', 'g663e', 'r678g', 'deutz', 'terlouw', 'grigelionine', 'mortier', 'n546k', 'fgfr2ks', 'cocrystallization', 'y769', 'lobate', 'phosphorylated', 'favelyukis', 'tetrapeptide', '767eey1770', 'uninterpretable', 'igf1rk', 'py657', 'r649', 't661', 'k658', 'v679', 'y680', 'e565', 'k641', 'relaxes', 'k526', 'd530', 'e526', 'r664', 'y657', 'k569n', 'n659', 'schubert', 'mctigue', 'shewchuk', 'strucuture', 'unphosphoyrlated', '815', '2i1m', 'n546', 'k517', 'e534', 'reconfiguration', 'tightened', 'unphosphorylate d', 'y656', 'nh4so3', 'stoichiometrical', 'spec', 'hkl2000', 'fgfr1k', 'lsqkab', 'ddmp', 'roselab', 'spectrophotometric', 'kkeeeymmmm', 'casnellie', '5p12', 'cholesterol', '022970', '000141', 'bek', 'url', 'spermatogonial', 'cwr22', 'sfani', 'cv1', 'dehydroepiandrosterone', 'hydroxyflutamide', 'stanolone', 'cwr', 'tsu', 'pr1', 'wainstein', 'gregory', 'bracket', 'lubahn', '4048', 'gcatc', 'gtatc', 'zapii', 'pcmvhar', 'androstenedione', 'dhea', 'circulates', 'sulfatase', 'nonhyperplastic', 'v715m', '9m', 'ici', 'v730m', 'kemppaine n', 't877s', 'q919r', 'c619y', 'heptad', 'mohler', 'gre', 'ar4', 'antiandrogens', 'adrenally', 'flut', 'sv4o', 'albert', 'brinkmann', 'gin', 'q902r', '5cr', 'a72it', '15m', 'v71sm', 'cancer5', 'v715', 'mmtvpa3luc', 'phhluc', 'pestell', 'lavalette', 'jil', 'charcoalstripped', 'estradiob', 'stcraloid', 'shering', 'kenilworth', 'nilutamide', 'roussell', 'uclef', 'romainville', 'jcj', 'casodex', '10', 'rlus', 'cotransfccted', 'cxpressing', 'lucifcrase', 'nsi', 'nst', 'androstenedionc', 'a721t', 't8775', 'cstradiol', 'gabactosidase', 'ses', 'andro', 'stenedione', 'v7i5m', 'hydroxyflutamidc', 'certam', 'bicabutamide', 'viva', 'cxperiments', 'sponses', 'hydroxyfiutamide', 'cently', 'intravesical', 'differentiation10', 'ubc12', 'dsh1', 'cubi', 'cubiii', 'uc14', 'bc3c', 'sh4', 'sh6', 'cbt', 's653c', 'd277', 'l15f', 'r143q', 'd277h', 'y239', '75um', '0279', 'ecd2', 'ecd1', 'ecd3', 'voluminous', 'adaption', 'weinberg31', 'lfing', 'gunes', 'mchugh', 'ser46', 'r8581', 'ht1376', 'tccsup', 'uc3', '647v', 'bftc', 'jmsu', 'sw1710', '253j', 'jer', 'tsong', 'cubii', 'planted', 'cbot', 'gctcgaggcc', 'gcagtgagca', 'ccatggagct', 'ggcggccttg', 'gtctagata', 'cactggcacg', 'tccagaccc', '23888', 'transfectability', 'lifters', 'peroxidases', 'flowing', 'uncentered', 'hsp90ab1', 'microtiazolium', 'r156h', 'r267q', 'dulic', 'farmer', 'jf', 'trkb', 'neurotrophins', 'bdnf', 'h1466', 'h820', 'h128', 'c14', 'kyushu', 'rasband', 'y516', 'victor', 'sfks', 'midgett', 'casali', 'chk', 'vidal', 'ebs', 'fig2a', 'intimal', 'fig4a', 'upd', '5m', 'puckered', 's3s', 'sauer', 'redundantly', 's5m', 'tian', 'ning', 'tice', 'courtneidge', 'simulating', 'maa', 'extrahepatic', 'uences', 'intuitive', 'kaposi', 'bene', 'neuroblasts', 'bromberg', 'turkson', 'catlett', 'falcone', 'desperately', 'mta', 'citations', 'cyquant', 'b27', 'poultry', 'lectin', '5mg', 'y877', 'regain', 'protumorigenic', 'ppi', 'behalf', 'creative', 'commons', 'noncommercial', 'blackwell', 'layout', 'proteinscence', 'stearoyl', 'bmax', 'simplifies', 'leakage', '1816', 'optics', 'fatal2', 'gene5', 'hap', 'cancer13', 'kcnr', 'azd3463', 'childhood1', 'patients8', 'proliferation9', 'dis ease26', 'physiochemical', 'orthotropic', 'cells34', 'autonomic', 'family9', 'tumors16', 'heritability', 'activity29', 'm1149t', 'cases5', 'catecholamine', 'impetus', 'y35a', 'asp65', 'tyr35', 'enthalpic', 'autoinhibits', 'carboxamide', 'rapgap', 'y35', 'mazhab', 'jafari', 'gmppnp', 'shielded', 'pulls', 'frech', 'ser68', 'noncatalytic', 'r15g', 'carboxylates', 'q61e', 'arg15', 'timescale', 'sot', 'krengel', 'y35f', 'allin', 'kj', 'entropically', 'penalty', 'overabundance', 'tyr34', 'srp', 'rangap', 'kanamycin', 'htrf', '2103', 's221', 'pi103', 'as605240', 'perifosine', 'i391m', 'g466', 'txg', 'navelbine', 'vasudevan', 'moelling', 'brachmann', 'torc2', 'suspects', 'aorta', 'gonad', 'mesonephros', 'frs', 'gab', 'icsbp', 'irf8', 'nrptps', '9126', 'm5g', '117', '1030', 'shps', 'csw', 'y542', 'backside', 'y63c', 'e69q', 'f711', 'a72s', 'e76d', 'q79r', 't42a', 'y62d', 'i282v', 's502t', 'g503r', 't73i', 'd61y', 'd61n', 'e69k', 'a72v', 'e76g', 'l560f', 'subverted', 'kinking', 'araki', 'misdiagnoses', 'macrocytic', 'bn', 'leukemiaassociated', 'spry1', 'jarvis',

'ena', 'vasp', 'nfg', '5r', 'mandates', 'y177f', 'y177', 'mohi', 'etten', 'erythroleukemias', 'keen', 'bentires', 'alj', 'virulent', 'miniature', 'dos', 'overridden', 'pyk2', 'ubiquity', 'imposing', 'skull', 'jaw', 'prof', 'rinse', 'unknowns', 'pone', 'e013', 'e014', 'e015', 'gprotein', 'hydroxylamine', 'pvuui', 'quantitating', 'zip', 'lacal', 'heteroduplexed', 'undissected', 'celery', 'reannealing', 'useable', 'dfcilu', 'raging', 'phi29', 'jumpstart', 'nlaiii', 'bstui', 'dnasep', 'triethylammonium', 'intercalating', 'e749del', 'homodplexes', 'rerun', 'q787q', 'r836r', 'g779f', 't783', 't783a', 't783i', 'tt55', '1792', 'tt53', 'tt90', 'tt60', 'r838r', '1792del', 'reevaluate', 'dnmt3b', 'unpredictable', 'dnmt3l', 'banked', 'calgb', 'r882h', 'r882c', 'yamas hita', 'icf', 'nonpromoter', 'simpson', 'flt3itd', 'pwpp', 'homotetramer', 'obs', 'g543', 'cautious', 'imprint', 'platypus', 'gestalt', 'sotos', 'opitz', 'standardised', '1958', 'cls', '176876', '182530', '190070', '164760', '176872', 'keilhack', 'digilio', 'sarkozy', '164757', '601263', 'narumi', 'schulz', '190020', 'kerr', 'zampino', '609291', 'nystrom', 'dysmorphologists', 'anthropometric', 'otorhinolaryngologic', 'voron', 'nfns', 'stumpf', '3500ht', 'ampdester', 'p173', 'gain3', 'mapper', 'coffalyzer', 'thr241pro', 'glu275lys', 'trp531cys', 'leu597val', 'thr599arg', 'lys601gln', 'pefp', 'cambre', '10cm', 'thr241arg', 'shchelochkov', 'makita', 'leu245phe', 'phe595leu', 'asp638glu', '1406g', 'gly469glu', '1785t', '1789c', '1796c', 'doliocephaly', 'polydramnios', 'webbed', 'hypomineralization', 'lids', 'epicanthal', 'pilaris', 'sucking', 'hyperhidrosis', 'n07', 'n23', 'hemangiomas', 'macrosomia', '1501g', 'glu501lys', '1495a', 'lys499glu', 'respects', 'glu501', 'asp594', 'polyhydramnios', 'armour', 'gln709arg', 'nosologically', '5fu', 'colo320', 'widr', 'kyse180', 'kyse', 'nanodrop2000', 'sty', 'snp6', 'partek', 'dice', 'clea', 'equalize', 'mau', 'tai', 'warfarin', 'hsc43', 'recommenced', '2355', 'merelbeke', 'cov', '5min', 'biom', 'eia', 'coinjection', 'scapulae', 'forthe', 'mobilize', 'hrg', 'nusse', 'varmus', 'wnts', 'christian', 'mailstop', 'childrens', 'rowe', '2987', 'graz', '12q24', 'xq28', 'shortage', '3425', '2193', '40k', '00028', '91k', 'w802', 'whites', 'lz', 'biding', 'talin', 'mcgiffert', 'woodlawn', '60637', 'paltermax', 'gctggcgctaaagaataaccacccatatatcttagac', 'ctaccagatacctaccagtatctccgtactatcttgtc', 'ctttacccgactcttgagccctggctcgc', 'cagtcctccttggctgattgtctggatggatgc', 'geneloc', 'weizmann', 'rehovot', 'nist', '8444', 'dnaanalysis', 'primertoolspage', 'tk6', 'transilluminator', 'd11s929', 'qloh', 'nutley', 'elutions', 'sidak', 'greenhouse', 'geisser', 'v10', '1620f', 'rs2227988', '72515', '72517', 'r830k', 'a848t', 'h398y', 'c384r', '1380p', 'nh3t3', 'behoove', 'degradation', 'hermann', 'erbs', 'downplays', 'oncogenome', 'chelix', 'exibility', 'l485', 'p490', 'ly3009120', 'sheng', 'lilly', 'eli', 'corporate', '46285', '4549', '1414', '0896', 'g596', '2016cancer', 'e586', 'g469', 'brafmutant', 'benefi', 'rmation', 'bxpc', 'v487', 'p492', 'h2405', 'ling', 't488', 'q493', 'q490', 'q494', 'p940', 'phosphomek', 'pacaca', 'rafb', 'softagar', 'pla', 'colonies40', 'cellomics', 'arrayscan', 'vti', 'hcs', 'e586k', 'p61v600e', 'arg509', 'crafef478k', '5c9c', 'lys483', 'disfavors', 'helixout', 'arg506', 'asp449', 'thr508', '087', 'turquoise', 'lled', 'uorine', 'a489', 'toggle', '1486', 'r506', 'r509', 'd449', 't508', 'thr488', 'ala489', 'val487', 'pro490', 'asn486', 'leu485', 'antiproliferation', 'vem', 'tra', 'cparp', 'hij', 'bxr', '12181', 'rxl', '11831', '3027', '4845', 'brafengaged', 'nonv600', 'shortens', 't799', 'investigatedthe', 'transformingactivityofthe', 'oncogenewith', 'medullary', 'fmtcs', 'totalof', '18mutantgenes', 'revealedthatcodon634', 'mutationshavethe', 'mutationwas', 'comparedwith', 'lions', 'mightimpairtransportof', 'plasmamembrane', 'observedfor', 'hirschspnmg', 'potentialto', 'sdisease', 'medullaiy', 'extracel', 'lular', 'frequenfly', 'cova', 'lent', 'cranio', 'caudal', 'sprung', 'strated', 'cdnaclonecontainingtheentirerecodingsequence', 'aptag', 'video', 'sequencedto', 'mutationswere', 'preparedfrom', 'transfectantsas', 'membraneswere', 'reactedwith', 'retantibodyor', 'chemi', 'retic', 'ulum', 'c61', '8s', 'c620r', 'c609y', 'c609w', 'pleu', 'ochromocytoma', 'c618r', 'c618s', 'haemopoietic', 'v65

81', 'f938s', 'p960s', 'k1026e', 'y1035c', 'val658', 'homlogy', 'hjak1', 'pri mestar', '180a', 'y81a', 'tyr701', '9171', '9138', '3344', 'tyr1022', 'tyr102 3', '422g', '611964', 'tyr1034', 'tyr1035', '3eyh', 'f958s', 'lys1026', 'kne w', 'val617', 'starving', 'leu80', 'tyr81', 's646p', 's646f', 'and5a', 'subt1 ety', 'l624', '629w', 'ile919', 'leu922', 'pro960', 'arg879', 'pro933', 'p933 r', 'and3b3b', 'dusa', 'behaviours', 'val658f', 'xiang', 'needless', 'nfkb2', 'synergizing', 'tumorgraft', 'swerdlow', 'grewal', 'chiarle', 'grandis', 'kos kela', 'gse6338', 'y640f', 'd661y', 'jerez', 'piva', 'dy', 'lymphokine', 'sal vaged', 'nil', 'traf1', 'grr', 'lyt', 'overtime', 'pststat1', 'sanda', 'concurr ence', 'lymphokines', '2014b', 'triage', 'voena', 'multiparametric', 'szj', 'implanting', 'bioethical', 'y1230', '1299f', 's323g', 's1058p', 'h1838', 's w900', 'sw1573', 'chlorophenyl', 'methylpiperazin', 'pyrrol', 'oxo', 'phospho epitopes', 'py1230', 'py1021', '1252', 'prestarved', 't421', 's424', 'u95av 2', 'u95', 'unmasking', 'mach3tm', 'walnut', 'creek', 'vulcan', 'nonimmuine', 'immunoscoring', 'phosphoepitope', 'bhattacharjee', 'u95a', 'assure', '7926', 't127', 't74', '1229f', 't63', 't117', 'c1124a', 't123', 't103', 'n103', 'n6 3', 'n80', 'h249', 'tig', 'pennacchietti', 'antivascular', 'py31', 'ribozym e', 'y1234', 'y1235', 'multisubstrate', '5d5', 'm13r', 'genecodes', 'glucuron idase', 'tggttggagagctcatttgga', 'gcactctcgccgtgactgtt', 'tttgccgatttcatgac t', 'mgbnfq', 'cattaaaggagacaccatagcta', 'cctgatcgagaaaccacaacct', 'catg aagcgaccctctgtatccc', 'bhq', 'karen', 'ervin', 'pflag5a', 'prk5tkneo', 'ex1 4', 'autocloned', 'polytron', 'kinematica', 'preclearance', 'rhuhgf', '1365', 'schwall', '3d6', '587f11', 'h596', 'dinucleotidic', '1964', 'y1365', 's985 a', 'of10', 'h838', 'h2882', 'rerf', 'calu6', 'mkn45', 'h19934', 'mkn455', 't agaagagccagccagtgt', 'cgaatgcaatggatgatctg', 'aaagccgctcaactacatgg', 'tgcttt gaatgcgtcccagag', 'tcaaattgaacagcgagctaaat', 'ttgaaatgcacaatcaggctac', '00000 7', '99615', '99728', 'gattgctgttgtctcaatatc', 'tgtcaaataacttacttggcagagg', 'acagtctacaagggcatctgga', 'aactccacacatcactctggtg', 't1095', 't1323', '3075', '3215', 't1021', '2732', 't1148', '2029', 't1165', '1611', 't1181', 't1241', 't1307', 't1384', 't1065', 'and14', 'vgs', 'exclusionary', '26met', 'nongene tic', 'pha665752', 's780', 's795', 'imunoblotting', 'velcade', 'timed', 'cpo 6', 'barcoding', 'plummeting', 'modernized', 'hydroxyglutaric', 'l370', 'hitmob', 'candra', '006218', '006209', 'hmller', 'mani', 'megm', 'fingerprinted', 'attb2r', 'attr2', 'attr1', 'attr4', 'stbl3', '300g', 'refreshing', 'opaque', '8x105', 'pik3cawt', 'pik3cae545k', '1x105', 'caattaaccctcaactaaagg', 'ccgccac tggctggata', 'xpress', '17', 't1025t', 'engineer', 'e545e', 'h1047h', 'e542 e', 'e542v', 'q546e', 'e545q', 'e453k', 'indendent', 'a1020v', 'vetted', 'c30 75t', 'a3140g', 'a3140t', 'g1633a', 'pik3cat1025t', 'pik3cag1049r', 'pik3cam1 043i', 'pik3cak111n', 'n345i', 'hmechs', 'precept', 'n345l', 'metabolomics', 'cyclinb1', 's1981', 's345', 't68', 'claudin', 'pik3cae542v', 'pik3cae453k', 'scalability', 'deconvoluting', 'coined', '1960', '1965', '1969', 'spranger', 'lumbar', 'brachydactyly', 'gly380arg', 'lys650glu', 'lys650met', 'tavormin a', 'spondylolisthesis', '30th', 'percentiles', '35th', 'thighs', 'depresse d', 'orthopedic', 'bspmi', 'fibulae', 'radiograph', 'raffioni', 'naski', 'fat hers', 'nhuc', 'dailey', 'eswarakumar', 'legeai', 'mallet', 'sahni', 'sible y', 'avis', 'wallerand', 'karoui', 'oers', 'mghu3', 'tcc97', 'pierrot', 'torr ecuadrada', 'subconfluence', 'y762f', 'ronchetti', 'santamaria', 'espana', 'korkolopoulou', 'agazie', 'hart', 'bonaventure', 'lievens', 'liboi', 'y762', '191170', 'brosh', 'petitjean', 'roussel', 'kuo', 'igg2b', 'ez1', 'roy', 'tb x2', 'tip3', 'dhar', 'topbp1', 'hamartomata', 'antagonising', 'antagonises', 'characterising', 'r611w', 'a614d', 'f615s', 'c696y', 'v769e', 'n525s', 'k59 9m', 'r367q', 'a607t', '609inss', 'p1202h', '1240', 'nrui', 'pqe', 'nobokun i', 'miescher', 'rap1gap', 'onda', 'additives', 'permeabilised', 'leicadm', 'rxa', 'laemelli', 'muci', 'alpha32p', '1828insgca', '616', 'destabilised', 'technology14', '81k', 'ubiquitination15', 'phosphorylation16', '1240x', '10 99x', '607x', '1532', 'immunofluorescent', 'rhebgap', '69k', 'matrix25', '1292 a', '1814a', 'dtu', 'webface', 'jobid', 'netgene2', 'mutation26', 'hamartomat

ous', '1631', '1713', '1134', '1535', 'r611', 'a614', 'f615', 'c696', 'surprisingly', 'aicher', 'areessential', 'immunoprecipitable', '1753', 'y1571h', 'prosite', 'stabilises', '1712', '3413insg', '1558', '1087', 'freud', 'fluoroesc ein', 'hernia', 'uncomplicated', 'irritability', 'sepsis', 'puncture', 's33 a', 'polakis', 'dishevelled', 'dsh', 'conductin', 'lustig', 'behrens', 's45', 't41', 'aberle', 'orford', 'guerin', 's45a', 't41a', 'dsg', 'nontranscription al', 'phosphorvlation', 'cateninin', 'accumula', 'threoninc', 'carcinogencsi s', 'coun', 'tries', 'occasioned', 'mvc', 'accepled', 'ol', 'indicale', '1thi s', 'kducation', 'lor', 'kulure', 'mh02', '5449', '5433', 'yusuke', 'eadheri n', 'cudhcrin', 'adenomalous', 'fodrin', 'polarize', 'cateninresult', 'stabil iza', 'cateninand', 'cateninmay', 'highmobility', 'cateninhas', 'cateningen e', 'cateninactivation', 'ikcda', 'diag', 'nosed', 'noncanccrous', 'manufactu r', 'reversely', 'nologies', 'subse', 'ccataacaactgtttgaaaatcc', 'tgagctcgagt cattgcatac', 'cagtccttcactcaagaacaag', 'atagttaatcaagttacaactgc', 'gtttggctg aaccatcacag', 'gttatgcaaggcccagcgg', 'atcatgcgttctcctcagatg', 'caccagcta', '2524', 'tcttacacctggactctggaat', 'tcttacactgaactctggaat', 'actctgg', 'uvaggcaa tcctg', 'atgtg', 'ggaatcc', 'attctggt', 'aatcctga', 'ggaaga', 'ggaatcctggaag a', 'ctggactctggaatcctgg', 'cateninfound', 'aatcc', 'aacgcactgcc', 'gccattaca actctccacaac', 'gatagcaccttcagcactctg', 'tcataactggctagtggtaggac', 'ccaaggagacc ttccatccc', 'agatgctcaactaaacagg', 'gtgaaggcgaactgcattctg', 'ggtgttatagaggt ctgtg', 'ctgaactagtcgtggaatggc', 'atcatgcaccccttcgtgag', 'ttcaatgggagaataaag cagc', 'atgttccacaaccgaattgttata', 'gtcagctcaactgaaagccg', 'tgcgttccgaatgtctgagg ac', 'gttcagaacatacagctaaagg', 'cfor', 'igof', 'fmbio', 'ccagcgtggacaatggcta c', '2525', 'ctileninmutatinns', 'amang', 'tumorpoint', 'linn', 'tilt50', 'ta 42t70t30t3it65', 't67t9', 'ta99intenititial', 'deletions8affected', '32323334 414545deletion', 'umino', 'acidsmutationongac', '51', 'ion', 'dala', 'x87838', 'elhidium', 'stratagenc', 'repealed', 'ditluoridc', 'c19220', 'frag', 'subst i', 'tution', 'cateninmutations', 'cateninat', 'catet1', 'bloi', 'caleninis', 'siles', 'yldsgihsgattapslsg', '4intron', 'txon', '2526', 'cateninor', 'cate ninaccumulation', 'gatekeeping', 'hdlg', 'zw3', 'htcf', 'apc331', 'apc1309', '2075', 'apc1941', 'apc2644', '2644', '2843', 'txv', 'noncolorectal', 'fe9', 'ser45', 'ser33', 'kentucky', 'generality', 's45f', 't44a', 's33y', 'lef1', 'rubinfeld', 'molenaar', 'travis', 'wetering', 'dld', 'cmt', 'sw948', '158 k', 's37a', 'ective', 'giese', '48c', '1178', 'cuto', 'arry', '1000nm', 'ptrk a', 'g466v', 'g596r', 'houben', 'tzivion', 'y341', 's494', 'r469', 'd470', 'hekman', 'g466eb', 'g466vb', 'g596rb', 'k483', '195kb', 'favata', '810kb', 'g4 66ebraf', 'g596rbraf', 'y341ac', 's338ac', 't491a', 's494ac', 's621ac', '429k b', 's338a', 'y341a', '239kb', 'edc', 'y340', 'karasarides', 'neill', 'chiloe ches', 'phosphos338', 'genes15', 'unsorted', 'xpo1', 'klh16', 'spiked', 'cheking', 'brafd594v', 'nonsupervised', 'tgfb2', 'fst', 'sfrp4', 'nondegradabl e', 'hackensack', 'nytran', 'explants', 'toluidine', 'pixcell', 'paint', 'precipitant', 'precautions', 'dntcf4', 'ptopflash', 'concatemerized', 'pfopflas h', 'replete', 'convenience', 'triphasic', 'histogenesis', 'itpr1', 'fra', 'd ichotomy', 'dickkopf', 'adhering', 'hca', 'contraceptives', 'hcav', 'contrace ptive', 'rooks', 'mccune', 'albright', 'maturity', 'calderaro', 'bacq', 'bluteau', 'jeannot', 'bhca', 'ihca', 'bioulac', 'uhca', 'bihca', 'poussin', 'laumontier', 'ihcas', 'bhcav', 'bihcas', 'uhcas', 'slco1b3', 'kpna4', 'ddx11', 'ki aa1109', 'q211p', 't1055r', 'rak', 'yim', 'v378', 'f379', 'vf', 'k380', 'deli nse', 'vk', 'e346g', 'e346', 'y350c', 'y350', 'delinskln', 'vdelinskln', 'del insl', 'alphascreen', 'crp', 'talpaz', 'allografted', 'a723d', 'v658', 'r72 4', 'staerk', '1023', 'a418', 'f421', 's703', 'a723', 's45p', 'k335i', 'k33 5', 'n387k', 'n387', 'cotransfections', 'glul', 'k335t', 'lohs', 'noninflammatory', 'partition', 'stefanska', 'acevedo', 'letouz', 'govindan', 'peg3', 'a1 26', 'monoamine', 'maob', 'killela', 'grivennikov', 'musteanu', 'fasan', 'pton toglio', 'melki', 'cadoret', 'omics', 'verifications', 'humancnv370', 'duo', 'humanomniexpressbeachips', 'tumult', 'calvo', 'ptc1', 'ptc3', 'ncoa4', 'el e1', 'y1062', '22q34', 'ay803272', '147171', 'figure2e', 'figure44b', 'belaru

ssian', 'figure55b', 'ex8a', 'ex10b', 'akap', 'accidental', 'figure77', 'reco mbine', 'randomness', 'papss1', 'pt4bn3', 'sulfurylase', 'phosphoadenosine', 'tripartite', '34q34', 'tax1bp1', 'd594e', 'adenylylsulfate', 'fchsd1', 'lcm n', 'unresected', 'rktr', 'sievert', 'fgf6', 'thumbs', 'hexameric', 'mc3t3', '6046', 'galacto', 'tropix', 'radioligand', 'calvaria', '33d', 'suture', 'sa gittal', 'osseous', 'tendons', 'condensations', 'aer', 'p60src', 'gap1', 'sam uel', 'lunenfeld', '8262', 'pawson', '1840', '77a', 'sacs', 'lossof', 'port', 'matsuda', 'strained', 'kazlauskas', 'margolis', 'sumimoto', '10mers', '110 2', 'concentra', 'concen', 'trations', 'koyama', 'tional', 'fluorography', 'nakata', 'myristylation', 'oocyte', 'y527', 'milligram', 'rick', 'fps', 'avm', 'cms', 'klippeL', 'cloves', '1900', '601728', 'halo', 'fistula', 'donahue', 'ipsilateral', 'penzberg', 'cnag', '5y', '2y', '1384', '4y', '1494', 'gln54 6', 'flaccid', 'paraplegia', '2081', 'tetraplegia', 'aneurysm', 'headaches', '2344', 'discomfort', 'paralysis', 'stroke', 'naevus', 'epistaxis', 'incline d', 'angiographic', 'vasomotor', '5q14', 'mef2c', 'cardoso', 'glomuvenous', 'gvm', 'saxton', 'qu', '1968', 'cytochemical', 'preprint', '26b', 'rome', 'c 459g', 'fragale', 'hof', 'deepview', 'peitsch', 'f513', 'nonbonded', 'molmo l', 'summation', 'y63', 'f71', 'e76', 'i282', 'a42', 'l43', 'e110', '1560', 'h53', 't42', 'meunier', 'fukuda', 'jongmans', 'repulsive', 'drifted', 'intr adomain', 'd308', 'miscarried', 'dab2', 'dab2ip', 'phoshO', 'sgk', 'intend', 'figure4e', 'phoshorylation', 'koepp', 'strohmaier', 'yada', 'oberg', 'pung a', 'inuzuka', 'wertz', 'galan', 'pashkova', 'liou', 'wulf', 'peptidyl', 'ppi ase', 'isomerizes', 'tun', 'kyi', 'ayala', 'bao', 'ryo', 'suizu', 'rustighi', 'pani', 'girardini', 'foxos', 'brenkman', 'stanya', 'fb13a', 't205', 'thr20 5', 'w34a', 'k63a', 't205a', 'stukenberg', 'kirschner', 'zita', 'barbash', 'd irupting', 'abolishment', 'glomulin', 'tron', 'w34', '5h5h', 'gefi', 'thelanc et', 'thailand', 'arrhythmia', 'angina', 'congestive', 'randomisation', 'guan gzhou', 'string', 'eff', 'centres', 'voice', 'therascreen', 'clopper', 'progr essionfree', 'posthoc', 'lw', 'nov', 'gure', 'estimable', 'described9', 'rece ptive', 'aff', 'subanalyses', 'nej002', 'doublemutant', 'iliac', 'aro', 'cyp1 7a1', 'leuprolide', 'c19', '878', 'histogene', 'hotstart', 'inpp4b', 'ddc', 'inventoried', '1944', 'lhrh', 'ignore', 'eosinstained', 'pericardiac', 'kat ou', 'pref', '1349', '7006', '00782', 'ggatcgccctc', 'ttcatgc', 'aac ttggcga ctatctgc', 'microsampling', 'lmd', 'wetzler', 'dr6g', 'dtamra', 'dr110', 'dro x', 'fluorescence', 'dtamara', 'ggaggaccgtcgcttggt', 'tcttga', 'catgctcggtgt t', 'caccgcgacctggcagc', 'contiguously', 'nonsensitive', 'fluorescence', 'js t003692', 'jst004556', 'gefinitib', 'ddg', 'dda', 'dele750', 'insip', 'dela85 9', '1883', 'bai1', 'deglycosylation', 'polyt', 'kinetochore', 'lucc2', 'and4 c', 'pfb', 'microhomology', 'fog', 'deformation', 'p5371', 'pugpuca', 'tgpycp y', 'pentamer', 'aggca', 'p53bp1', 'tgccct', 'prset', 'coinfecting', 'm13k07', 'doublestranded', 'py10', 'py28', 'plgk669z', 'kpnixhoi', 'cm1', 'n9', '9132 0', '5196', '2773q', 'c277q', 'a1g', 'microcentron', 'pep4', 'trp1d', 'a1 c', 're2', 're3', 're7', 're8', 're9', 'r248k', 'r283q', 'k120n', 'c277r', 'a 1t', 'g3t', 'g3c', '5198', 'thukral', 'activityb', 'g2c', 'g2t', 'c4g', 'c4 t', 'a5g', 'a5c', 'a5t', 'r273k', '5199', 'pugpucaztgpycp', 'ggcaagtct', 'p entamers', 'acceptors', 'ggca', '5200', '37a', 'grave', 'tumors8', 'tumors1 3', 'sequencing13', 'others17', 'kdm5a', 'others34', 'arid1', 'tumors42', 'genes26', 'described46', 'far1', 'bins', 'querying', 'discordantly', '1kb', '5u m', 'coffin', 'siris', 'agenesis', 'dysgenesis', 'sim', 'hypertrichosis', 'cc ds', 'osa', 'hindbrain', 'ribosomes', 'neuroepithelial', 'repositions', 'stud y9', 'qiashredder', 'tcf7', 'ultrasonic', 'caco', 'glutamin', 'fopflash', 'pf a2', 'dehner', '2188', '1777', 'axioplan2', 'morphometric', 'deconvolution', 'neuro2a', 'duolink', 'starter', 'd1h2', 'pttg1', 'ppard', 'wnt3a', '60325 4', 'tcf712', 'heuvel', 'suppressor1', 'carcinomas2', 'cancers2', 'nucleus5', 'interactions8', 'activity14', 'ligand9', 'protein16', 'construct14', 'analys istable', 'molscript26', 'raster3d27', '38k', 'members5', 'proteins1', 'droso phila18', 'elegans19', 'mutations1', 'screens18', 'arg441pro', 'leu440arg',

'pro445h', '55k', 'asp351his2', 'arg361cys17', 'val370asp17', 'asp493his1', 'asp450glu12', 'deprive', 'arg420his', 'pro445his', 'phenotypes18', 'overexpressed15', 'colouring', 'wedging', 'grasp29', 'r361c', 'v730d', 'd537e', 'g508s', 'burying', 'asp351his', 'arg361cys', 'ovarian2', 'cancers17', 'asp450glu', 'asp493his', 'sma4', 'imporant', 'gly508ser', 'approx180k', '181k', 'trimerization', 'approx3', 'mrof', 'oligomerization15', 'mrconsistent', 'mr180k', '24k', 'heterohexamer', 'undulations', 'disks', 'kinase30', '62k', 'ultrafiltration', 'vapour', 'monomethylether', 'peg5000', 'mpeg5000', 'macroseedling20', 'f4132', 'axisiic', 'rigaku', '200hb', 'generator', 'mpeg', 'thimerosal', '3pboocch3', 'uranyl', 'shelxs', 'mlphare22', 'flattening', 'squash23', 'plor25', 'beta1', 'mrstandards', '440k', '25k', 'bps', 'smad5', 'smad8', 'miyazono', 'sbes', 'agac', 'moustakas', 'kardassis', 'r497h', 'k507q', 'r515g', 'p1021', 'p130s', 'mb468', 'nagarajan', 'meanplusminuss', '18k', '143v', 'r100t', 'moren', '153k', 'r497', '70k', '102k', 'aliex', 'hoeschst33342', '20k', 'smad10', 'andin', 'oswego', 'bisacrylamide', 'prerunning', '665', 'anin', 'fromtop', 'topto', 'preimmobilized', 'mconcentration', 'thermodenaturation', '1440r', 'p445h', 'i527r', 'f346v', 'w398v', 'purple residues', 'andblue', 'andorange', 'inred', 'seme', 'xenopusis', 'psuedosymmetric', 'wellconserved', 'prolinerich', 'rsmads', 'mixer', 'gtct', 'sbr', 'd537y', 'smad4s', 'are3', 'hsmad4', 'xsmad2', 'haemagglutinin', 'xfoxh1a', 'xfast', 'caga12', 'pointmutated', 'efha', 'iscoves', 'galactostar', 'flowcytometric', 'westernblotting', 'hasmad4', 'hasmad4d351h', 'efhasmad4d537y', 'eflacz', 'hasmad4d537y', 'dnabound', 'supershifts', 'cagac', 'endogenousactivated', 'jps', '414', 'cac02', 'colo201', 'colo741', 'gp5d', 'hca7', 'hcc2998', 'hdc135', 'hdc143', 'hdc57', 'hdcc114', 'ht55', 'lim1215', 'lim1863', 'lim1899', 'lim2099', 'lim2405', 'lim2550', 'lim2551', 'ls174t', 'rw2982', 'skco1', 'sw1116', 'sw1222', 'sw403', 'sw837', 't84', 'analyzers', 'biomek', 'robots', 'pixsys', 'nanolite', 'cartesian', 'human610', 'mgc34440', 'mgc60396', 'mgc8602', 's2761', 'd300n', 'p3051', 't413n', 'd450e', 's647p', 'd258n', 'r268c', 'p336s', 'r373q', 'd408y', 's425c', 't204d', 'pcaga12', 'lumistar', 'galaxy', 'd43b4', 'blink', 'maxentscan', 'pro356', 'd351del', 'p356delinsqk', 'spliceoform', 'ala406', 'lys428', 'reoccurring', 'asp537', 'leu540', 'a406t', 'k428t', 'r515t', 'a118v', 'arg100', 'r100g', 'k45n', 'g65r', '198f', 'l172m', 't197i', 's4641', 's467p', 's464x', 's422', 's423insrn', 's423vfsx13', 'a354t', 'e228k', 's2661', 's2661', 'sno', 'a22v', 'k43n', 'r93n', '769a', 'pai1', 'serpine1', 'd351', 'r361', 'gly65', 'a118', 'd258', 'd408', 't184a', 'herbst', 'kris', '87k', 'submitting', 'breathlessness', 'ck20', 'keam', 'heo', 'daehak', 'jogn', 'kimdw', 's10147', 'adc', 't725t', 'p741p', 'g779g', '17921', 'p794p', '18141', 'v819v', 'd837d', '1692f', 'p694s', 'i706t', 't710a', 'e711', 't712', 'k714n', 'a755d', 'k737t', 'v8341', 'r836c', 'g873q', 'generalizing', 'e711t712', 'g873e', 'dupn771', 'dupa767', 'menchester', 'ness', 'sunaga', 'initia', 'erlo', 'deoxy', 'prelimi', 'nary', 'tients', 'morial', 'cgmh', 'cmrpg', '350031', '0615b', '0086c', 'tology', 'recov', 'ered', 'hav', 'atinine', 'postopera', 'stud', 'ceive', 'craniotomy', 'agnosis', 'lyzer', 'amplifica', 'tox', 'icity', 'radiologi', 'sponse', 'egfrtki', 'nificant', '3999', '1months', '2months', '9months', 'ficacy', 'orrs', '27months', 'summa', 'rized', 'paronychia', 'guid', 'patents', 'regimen1', 'status2', 'erlotinib3', 'delk745', 'v774dup', 'cision', 'keratitis', 'hepatotoxicity', 'propriate', 'study3', 'patients4', '686', 'liberation', 'literatures', 'oncologist', 'lifestyle', 'genes1', 'dates', 'imputed', 'birthday', 'ec1', 'complied', 'salpingo', 'breasts', 'awaiting', 'activity34', 'd281g', 'l22q', 'w23s', 'r267p', 'h1048', 'acetyalse', 'doublings', 'anova', 'm01', '5274', '12the', '5nm', 'egfrmutant', 'hmgal1', 'krt6a', 'krt14', 'krt5', 'h1573', 'word', '15q15', 'manuscripts', 'sp3', '435s', 'gsk1363089', 'diversify', 'cautionary', 'carboxylase', 'cabozantinib', 'foretinib', 'epiphenomenon', 'mcrc', 'tpc1', 'figure1c1c', '1553', 'specifics', '500ng', 'cl1', 'moabs', 'argyrophilic', 'q15q33', 'decompensation', 'ffler', 'endocarditis', 'leriche', 'conus', 'cauda', 'embolisms', 'gaagg

tgtgtttgttgcggt', 'nest', 'ttgacggccactttcatcgt', 'tggcttcttctgccaaagca', 'float', 'gpiap1', 'acagcacccacaaccaccagaa', 'catcgatctcgtaacgtggct', 'git2', 'ggcagacagaagttagtcggt', 'gttcaacgcccattagttgcc', 'agggagatgtggtagcacc', 'prkg2', 'prkg2f', 'catggcactagatcgagaggt', 'prkg2r', 'aagcatcttgacggcca', 'cttca', 'ggacagccaaagctagagatga', 'atcgtggctgagaatgctca', 'ldi', 'ighj', 'pkpni', 'cctcagccacagccgcagcagtgagaagc', 'pbr1', 'tcagagagtcttcccacccaactt', 'gagtcc', 'ninpdcfrb', '057169', '005898', '006259', 'racepcr', 'cdk5rap2', 'utmost', 'pointer', 'trip11', 'cev14', 'pde4dip', 'specci', 'hcmogt', '14q24', 'kiaa1509', 'nde1', 'supplies', 'caprin', 'm11s1', 'grk', 'cat2', 'git', 'cgk', 'cgkii', 'pkrg1', 'pkrg2', 'renin', 'eosinophilia', 'pound', 'erythema', 'dermographism', '128kb', 'drenching', 'hydrochlorothiazide', 'amitriptyline', 'simvastatin', '423n19', '13i6', '347kb', 'vwr', 'marrowmax', 'specc1', 'ccdc77c', 'caprin1', 'particulate', 'guanylyl', 'nonerythrocytic', 'nonerythroid', 'counterbalanced', 'p101', 'compartmentation', 'broth', '4a6', 'prenylation', 'vacuoles', 'k13e', 'cbr3', 'antitumoral', 'tcrb', 'nordic', 'arg465', 'arg479', 'strategene', 'sic1', 'multiubiquitin', 'ishikawa', 'scfhcdc4', 'sip1', 'd408h', 'w406a', 'il11', 'tnfaip6', 'fermt1', 'olfm2', 'wnt11', 'y226a', 'y297a', 'w326a', 'k341a', 'e267a', 'inducibility', 'schiro', 'stauber', 'e25021', '0025021', 'dhyan', 'credited', 'mda69637', 'jdrf', 'r21', 'ca135483', 'ca009135', 'p30', 'ca014520', 'funders', 'interests', 'wisc', 'proteinbinding', 'immunosuppression', 'tumorsuppression', 'smadindependent', 'cofilin', 'plosone', 'mcf10a1', 'smadsmad', 'rsmad', 'helix2', 'corridor', 'binduced', 'jeg', 'binducibility', 'v356r', 'q322a', 'y324a', 'toroid', 'lumier', 'jeffery', 'smadbinding', 'flagtagged', 'f303a', 'r295a', 'r240a', '180u', 'g001', 'aa665', 'aa252', 'aa2', 'ppgy', 'qpsmt', 'stsvl', 'nkn', 't001', 'y297', 'r292', 'r268', 'r243', 'e337', 'y238', 'xfoxh1', 'w274e', 'skiw274e', 'p335a', 'p336a', 'c338a', '1340a', 's254a', 'y298a', 'v224a', 'r268a', 'y238a', 'r243a', 'r292a', 'y384a', 't002', 't003', 'choriocarcinoma', 'y384', 'y324', 'k341', '1340', 'w326', 'activation', 'y226', 'v356', 'w406', 'v224', 'keratinocytes', 'sbe12', 'g002', 'transcriptional', 'ski274e', 'jeg3', 'smad3w406a', 'sample1', 'sample2', '005045', '01261', 'trxsaki', 'trxsara', 'trxsip1', 'trxsurf2', '0464', '4233', '003008', '008461', '003333', '9814', '2267', '1123', '0085', 'mstat', 't004', 'adamts4', 'cst6', 'myl1', 'sgcg', 'mycsmad3', 'delppgy', 'g003', 'e267', 'gfpt2', 'npy', 'btreatment', 'g004', 'smad3s', 'g005', 'responsive', 'inducibility', 'cystatin', '31610', 't005', 'sarco glycan', 'duchenne', 'kindlin', 'hyaladherin', 'hyaluronan', 'fibroblastmyofibroblast', 'olfactomedin', 'olfactomedins', 'olfm1', 'fostering', 'seborrheic', 'uncultured', 'p716r', 'v77m', 'e574k', 's688f', 'p708s', 'g227e', 'v248d', 'fgfr1c', '1fq9', 'intentionally', 'g226', 'v247', 'g227', 'v248', '1245', 'meshes', 'r250', 'r251', 'g271', 'g305', 'c341', 'e159', 't212', 'h213', 'fgfr2s', 'e219', 'k35', 'e160a', 'h213y', 'canyon', 'm640', 'i642', 'innermost', 'a648', 'm537', 'e636', 's563', 'disposition', 'c606', 'm614', 'r759', 'e731', 'd756', 'e475', 't555', '1468', '1528', 'v532', 'm535', 'nearing', 'g701', 'k668', 'g701s', 'e636k', 's568', 'e475k', 'r759q', '1770v', 'refold', 'severalfold', 'c278f', 'pngase', 'carbohydrates', 'endoh', 'rfgf2', 'ao4', 'd22', 'uacc2534', 'd771n', 'lacrimo', 'ariculo', 'dento', 'ladd', '149730', 'a628t', 'ptyf', 'systems', 'vialight', 'cytotoxicity', 'mcdb', 'leibovitz', 'starve', '1nmol', 'j9', 'kyte', 'doolittle', 's473a', 'prkca', 'v60', 'b7', 'valent', 'kiszewski', 'amitai', 'hypovolemic', 'broudy', 'caruana', 'k839e', 'yanagihori', 'verzijl', 'd816i', 'm099', 'rna1', 'dna2', 'e001', 'm5411', 'e033', 'e047', 'e009', 'e012', 'e024', 'e025', 'e027', 'e031', 'e034', 'e035', 'e036', 'e049', 'dcm', 'm081', 'm084', 'e018', 'dcm3', 'e010', 'insff419', 'e050', 'e022', 'e038', 'e030', 'e026', 'e061', '419insy', 'e017', 'c443y', 'e032', 's476i', 'e054', 'itd501', 'itd502', 'e020', 'e041', 'itd505', 'e040', 'e044', 'e053', 'd572a', 'e037', 'e042', 'e043', 'e060', 'e048', 'm216', 'plei', 'bsma1', '2121', '1934', '2413', '2991', 'd419', '418d419y', 'a502y503', 'position2', 'agagtcatc', '2247', 'agatacatc', '2446', 'agaatcatc',

'acttacttcttcgac', 'phephe', 'gaaatcctg', 'tacaggctc', 'thr417tyr418asp419', 'acttac', 'aggctc', 'asp419', 'tttatcca', '1328', 'cys443', 'cagatttct', '1427', 'ser476', 'acttctgcctgcctatTTT', 'ser501ala', 'acttctgcctgcctatTTT', '1502', 'tctgcctatgcctatTTAac', 'ala502', 'tatttaacttgcatTTactttgcatttaaa gag', 'asn505phealaphe', 'tttataagag', 'lys509', 'atagccccca', 'asp572', 'gtct gatt', 'met541', 'x06182', 'tyr719', 'py719', 'fileoptions', '69795', '5507', '66707', '7841', '66526', '5973', '67032', '8706', '70751', '8826', '69836', '5905', '64939', '4957', '69350', '5509', '66869', '4732', '72160', '5464', '69483', '8044', '68531', '5819', '68904', '10487', '65330', '5106', '60394', '5212', '67922', '5577', '66245', '8212', '28780', '10664', '1688', '8040', '13733', '12833', '14727', '1089', '12829', '13266', '1803', '12213', '67019', '57624', '7192', '14429', '11259', '1416', '59224', '7741', '21743', '2692', '12045', '12434', '62184', '4197', '67474', '66178', '5882', '73569', '6956', '13226', '13469', '1512', '11147', '1639', '34023', '3324', '32606', '2807', '65419', '54359', 'e0411', '44131', '30290', '63898', '7419', '50404', '4751', '23766', '66224', '65933', '6508', '66258', '7454', '24654', '2719', '68563', '10774', '68893', '7190', '69413', '4040', '50468', '5672', '57655', '9215', '62026', '8657', 'calculable', 'goemans', 'tyr503', 'yuzawa', 's476', 'lennartsson', 'socs6', 'bayle', 'a533d', 'rosbotham', 'darier', 'papules', 'tharp', 'extracutaneous', 'stipulated', 'stratascript', 'geneclean', 'graffenstaden', 'u2f', '11f', 'ple1', 'theou', 'cctagtgtccattctgacg', '1176', '1195', 'cgattacgaaaccaatcagca', 'ils', 'compounding', 'ldk378', 'brigatini b', 'asp3026', 'i1171s', 'f15', 'table4', 'figure2e2e', 'overdose', 'lorlatinib', 'c1182', 'h1247', 'f1174c', 'f1174v', 'unsurprising', 'and2d', '0169', 'kodama', 'resists', 'g1269', 'fontana', '2xb7', 'desperate', 'regimes', 'triploid', 'irregularities', 'fam150a', 'fam150b', 'fam150', 'clb', 'pc12s', 'w1118', 'balbc', 'figure4f', 'penetrate', 'panerk', 'haoyuan', 'chemexpres s', 'mirrus', 'lite', 'x106', 'sparingly', 'pbt', '3633bc', 'flour', 'a488', 'flouromount', '50mm3', 'pyrrolidinone', 'danver', 'k379e', 'csh2', 'ks459de ln', 'dkrmns560del', 'a66', '604850', 'zk', 'pik90', 'bez', 'partnering', 'netsurfp', 'disinhibiting', 'pbsfi', 'sfii', 'g376r', 'gattatactttcacactaaggaaa gggagaaaataacaaattaatcaaataatttc', 'gaaatattttgattaatttttctcccttcattgtg taagagtataatc', 'caaataccaacaggatcaagttgtcaaagataatattgaagctgttaggg', 'ccctaca gcttcaatattatcttgcacaacttgatcctgttggatttt', 'gaatataacactcagttcaagaaaatcgag aatatgatagattatatg', 'catataatctatcatattctcgattttcttggaaacttgagtgttatattc', 'ggcagctgagtgatcgagaaatcattaaaccagacccattccagctgag', 'ctcagctggataaggctggttaatg atttctcgatactcagctgcc', 'gaagcaggcagctgagtgatcgagaaatttacaaacgtatgaacacgattaa cc', 'ggttaatgctgttacgttggattttctcgatactcagctgcctgcttc', 'gtatcgagaaatt gacaaacgtatgaagagcattaaaccagacccattccagctg', 'cagctggataaggctggttaatgcttc atacgtttgtcaatttctcgatac', 'gcattaaaccagacccattccagctgaaagacgagagaccaataactt g', 'caagtattggctctcgctttcagctggataaggctggttaatgc', 'ccagacccattccagctgag aaagagagaccaataacttgatgtggtt', 'caaccacatcaagtattggctctttctcgactggataaggctgg', 'gaaagacgagagaccaataacttgatgtggtt', 'ccgaacacccctttgagtc aacatcaagtattggctctcgctttc', 'ggaaagggggaaataacgaattaatcaaataatttcatc', 'gatgaaatatttgattaattcgattttcccttcc', 'sfi', 'leghorn', 'mp6', 'as604850', 'p4e', '3011s', '4292', '2967', '9275s', '9451s', '7176', 'i255f', 'ku70', 'pttg1ip', 'macrofollicular', 'd0', 'nonbiotinylated', 'caesium', 'ibl', 'vo', 'stopper', 'microsystems', 'signification', 'envisage', 'csnk1a1', 'miapaca', 'molt4', 'pmdlg', 'shfbxw7', 'exeter', 'garvan', 'jsf', 'creator', 'collapse', 'c7', 'nucleofection', '127k', 'polyubiquitylated', '84k', 'macaque', 'dex', 'hugene1', 'txnip', 'constituent', 'all2', 'sensitizer', 'bbc3', 's214c', '2501', 'hospice', 'jci72763ds1', '9x', '7x', 's257', 's257w', 'tracheobronchial', 's214', 'd429a', 's257p', '4720', 'essentiality', '3611', 'mill', 's43', 't452', 'cancers15', 'termini1', 'rbds', 'proteins21', 'ras23', 'vitr025', 'gtp30', 'activation44', 'domain50', 'pak2', 'paks', 'contaminant', 'bag1', 'cells56', 'title', 'remember', 'g465r', 'g465e', 'coastline', 'lentiginous', 'alm', 'tg1796', 'nrasq61r', 'gorden', 'overgrow', 'metastasic', 'edge

d', 'sdh', '18fdg', 'paragangliomas', 'pgls', 'carotid', 'stratakis', '397c', 'arg133', 'jcem', 'pericardial', 'jugular', 'glomus', 'foramen', 'constrictio n', 'painful', 'incontinence', 'tension', 'tonsillectomy', 'subclavian', 'art eries', 'catecholamines', 'mainz', 'sacral', 'pyrograms', 'parasympathetic', 'impending', 'hoxa3', 'hoxb2', 'acknowledgment', 'dalton', '3061', 'relentle ss', 'pdgfs', '1department', '2department', 'catholic', '4department', '5depa rtment', '6department', '8department', 'intraluminal', 'embolus', 'centrome r', '231c18', 'mito', 'extender', 'lowgrade', '10q22', '12q12', 'apaf1', 'ccn e2', 'mos', 'ddit3', 's133', 's392', 's46', 'y699', 'hsp27', 's78', 'kitk642 e', 'attracted', 'bernrried', '2q35', 'permanox', 'denatured', 'missouri', 'gadd153', 'ifng', 'flu', 'l12', 'namalwa', 'electro', 'rgg', 'bamhl', 'term inals', 'e2a', 'pbx', 'liberated', 'pa317', 'pg13', 'stempro', 'glycophorin', 'cd33', 'cd13', 'plucked', 'laddering', 'myelocytic', 'reprograms', 'vrk1', 'kadoch', 'delattre', 'riggi', 'stamenkovic', 'mscs', 'adipogenic', 'torchi a', 'bilke', 'gangwal', 'guillon', 'sknmc', 'margueron', 'visel', 'ramakrishn an', 'thurman', 'atac', 'kaikkonen', 'wdr5', 'elf1', 'gabpa', 'hollenhorst', 'otx2', 'mafb', 'vicente', 'duenas', 'riveiro', 'falkenbach', 'soengas', 'apis', 'duncan', 'cables1', 'arnason', 'ier3', 'sebens', 'muerkoster', 'tgfb1', 'cobaleda', 'trps1', 'cd73', 'chamberlain', 'prkcb', 'surdez', 'valbuena', 'm olitor', 'traktman', 'chp100', 'skes1', 'ew7', '1f6', '8007', 'photographic', 'trcn0000005322', 'trcn0000197134', 'trcn0000002133', 'fugenehd', 'sw41ti', 'sacrified', 'mikkelsen', 'aiden', 'sonifier', '22810', 'a302', '429a', 'm91 8t', 'l790f', 'y791f', 'tumoural', 'machens', 'elisei', 'romei', '0088', '011 7', 'mathe', '1389', 'pisa', 'euroclone', 'frogs', 'cattle', 'pigs', 'gifte d', 'ceccherini', 'genova', 'mutagenised', 'resuspending', 'lowmelting', 'cc h', 'ffus', 'inconsistency', 'focusforming', 'm918', 'anchorageindependent', 'teams', '20mb', 'overproliferation', 'cml4', 'kinases9', 'deutsches', 'ress ourcenzentrum', 'genomforschung', 'rzpd', 'cnat', 'spreadsheets', 'neots', 'e nsg00000110395', 'dropping', 'i7f', 'tgtggttcaacttaaacccctgga', 'i8r', 'gccag gccaccctgtatc', 'i8f', 'ggcctggctttgggttagg', 'i9r', 'cacaatggatttgccagt tcc', 'cble7f', 'gaatcctgatctgactggcttatg', 'cble10r', 'gctgcggcagaaggtaag t', 'p065', 'wtcbl', 'line17', 'wtfl3', '2111c3a', 'datapoints', 'upn', 'w51 51', 'isodisomy', 'ephb6', '1387', 'p417a', '1139t', '1249c', 'cblb', 'cblc', 'crisis22', 'n454d', 'gran', 'bucc', 'molm', 'subcytogenetic', '32ds', 'autozygosity', 'consanguinity', 'ubiquitinylates', 'cas', 'established26', 'mds28', 'c381', 'h398', 'c379a', 'c379', 'hyperresponsive', 'hyperresponsiveness', 'kinopathies', 'egfr790m', 'dr1', '0064', 'geftinib', 'depolarization', 'preexistence', 'prsk', 'rp23', 'ccsp', 'rtta', 'isoflurane', 'planes', 'affiliatio n', 'frederic', '94080', '5841', 'slicing', 'autoactivating', 'x10', 'hrk', 'microsomes', 'halflife', 'sg274', 'd477g', 'cilium', 'c490', 'c507', 'c31 4', 'l515', 'sag', 'kaad', 'tubulins', 'veh', 'ligandbinding', 'ciliogenesis', 'kabuki', 'readthrough', 'paving', 'kleefstra', 'schinzel', 'giedion', 'berdasco', '602113', '003482', 'alr', '300128', '021140', 'bogershausen', 'wol lnik', '147920', 'columella', 'dermatoglyphic', 'kuroki', 'niikawa', 'hypodontia', 'hannibal', 'micale', 'paulussen', 'banka', 'makrythanasis', 'kokitsu', 'gigantic', 'dillon', 'ansari', 'mandal', 'eissenberg', 'lederer', 'bellais', 'aminoglycosides', 'nonaminoglycosides', 'hemophilia', 'rnasy', 'sseldorf', '17b', 'priolo', 'p389', 'rozen', 'skaletsky', 'eef1a1', 'ferrero', 'pcrfl', 'rousset', 'hospira', 'bidou', 'firelfly', 'sermet', 'gaudelus', '319d', 'pcr fl319', 'wa1', '4sr', 'dyex', 'genetyc', 'brunak', 'reese', 'caratozzolo', 'ratbi', 'kb43', '15785', '16168del2425insttgatctcaa', '514c', 'arg172x', '1849delactc', 'thr616tyrfsx8', '2939a', 'asp980val', '3284', '6delaagt', 'asn1070', 'lys1094del', 'ser59argfsx86', 'gly134glufsx74', '13999', 'asn4614il efsx5', '14252', '5insgaaa', 'val4751', 'glufsx22', '14643', 'gln4882profsx3 6', '14644', 'glu4882profsx36', '4693', 'val1561argfsx11', 'brichta', 'kb18 6', 'kb153', 'kb3', 'kb48', 'kb83', 'f1a', 'hoxc6', 's100a', 's100a2', 's100a 5', 's100a6', 'figurefigure1g', 'eres', 'ere1', 'ere2', 'schematized', 'floqu

et', 'kb41', 'kb45', 'dynamism', 'misdiagnosed', 'rbbp5', 'c1430r', 'c1471y', 'phd4', 'deroo', 'korach', 'branchial', 'arches', 'otoliths', 'clefting', 'mekel', 'cartilages', 'hemming', '12844c', 'timine', 'aminoglycoside', 'amikacin', 'biocompatibility', 'ptc124', 'ntc13', 'rtc14', 'nb30', 'bordeira', 'carrico', 'ncoa6', 'cointegrator', 'asc', '5537', 'decide', 'raav', 'contradictive', 'trr', 'combinatory', 'pathways1', 'f13', 'set1', 'fyrn', 'compass', 'set1a', 'set1b', 'trimethylating', 'klh', '9g', 'ccnf', 'cases3', 'vavp', 'glomerulonephritis', 'syndrome24', 'expression45', 'inhibitors54', 'mutation5', 'sudh14', 'sudh16', 'sudh10', 'ly18', 'u2932', 'gilmore', 'studies5', 'pрутurbo', 'sacii', '200mm', 'ab8580', 'ezview', 'mbcl2', 'immunizations', 'biose arch', 'freund', 'sp7', 'coalescent', 'paracortical', 'imgt', '1070', 'ques t', 'p68', 'ultrasonicator', 'minielute', 'rmdup', 'stitched', '2kb', 'cabin 1', 'pmb1', 'regimen3', 'malignancy4', 'escape4', 'prdm110', 'gene4', 'mef2 a', 'mads', 'mcm1', 'agamous', 'deficiens', 'domain17', 'activity17', 'pathwa ys17', 'genes18', 'context20', 'mef2d', 'thymocytes21', 'gc23', 'analysis4', 'sets5', 'k4e', 'y69h', 'e77k', 'n81y', 'd83v', 'r171x', 'y201x', 'g242fs', 'p256fs', '1269fs', 'tonsils25', 'lymphomagenesis26', 'hbci', 'zmynd11', 'plot29', 'aracne', 'algorithm30', 'vivo31', 'madsmeff2', 'p3hr1', 'genes31', 'tuberorfp', 'trfp', 'proxy', 'tool33', 'cabin120', '154p', 's78r', 'l38i', '12172a', 'mef2b20', 'fsk', 'cyclase36', 'h89', 'inhibitor37', 'pka34', 'k72h', 'p dsm', 'sumo', 'kxe', 'kxexxsp', 'ubc940', 'sumoylated', 'k319', 'sumo1', 'mon osumoylation', 'rxs', 's324', 'predictably', 'proteins41', 'k319r', 's324a', 'cases4', 'bcl226', 'instructing', 'y69', 'g2e', 'k5e', 'i8v', 'k23v', 'partners43', 'aicda44', 'f146', 'mice48', 'translocations49', 'sequence16', 'ep3008', 'fbxo1150', 'tgfb1', 'fn1', 'domain1', 'domains2', 'bzlf1', 'cases13', 'lymphoma17', 'binding18', 'analysis21', 'cells2', 'activity18', '26a', '26c', 'y223', 'epha7', 'cxcl12', 'expression35', 'cells37', 'mice43', 'mice44', 'understudied', 'f71i', 'g503v', 'kure', '8574', 'aokiy', '7178139', '7178142', 'kurosawa', 'kondoh', 'ishikiriyama', 'kawame', 'nagano', 'kamasaki', 'sa pporo', 'yamanaka', 'okazaki', 'takada', 'sagamihara', 'nishio', 'tamai', 'im aizumi', '0239', 'mendez', 'servidei', 'kosaki', 'musante', 'cubitus', 'vulgars', 'americanbritish', '009775', '57c', 'fostercity', 'mdykddddk', 'ksii', 'niwa', 'e76a', 'mosby', '30c', 'ptstepqpyqpgenl', 'clone12ca5', 'matozaki', 'immunoprecipitants', 'psreluc', 'prltkluc', 'avacus', '1508g', 'pda', 'ductus', 'arteriosus', 'vsd', '8years', 'slanting', 'parpebral', 'valgus', 'hydrocele', 'nbtam', '181g', '188a', '211t', '218t', '236a', '844a', '1504t', '1510a', '1510', 'serumresponsive', 'hakak', 'ahmad', 'bentiresalj', 't507k', 'y62c', 'zumkeller', 'hecht', 'd61a', 'ekr', 'egfstimulated', 'age1', 'therapy7', 'reported13', '1687', 'hsa', 'hyaline', 'teratoid', 'tumors38', 'a08', 'described29', '4967', 'baf250b', 'desmoplastic', 'anatomically', 'genetica', 'hirohashi', 'conacci', 'sorrell', 'puig', 'and3c', 'fujita', 'thermoscript', 'prizm', 'photodocumentation', 'd84v', 'd84g', 'r87p', 'p48l', 'd74n', 'r87l', 'r801', 'h83y', 'aterials', 'ethods', 'unc7', 'deoxyribonucleoside', 'yue', 'u12818', 'u12819', 'u12829', 'lkb', 'mncl', 'capacitance', 'pcmvcd20', 'esults', 'iscussion', 'e26d', 'r124h', 'carets', 'm1250t', 'e768d', 's891a', 'keeper', 'v804', 'v8041', 'charmm22', 'counterions', 'v804g', 'cushion', 'rotatable', 'torsional', 'computations', 'electrostatics', 'biomolecules', 'corrections', 'grained', 'simulate', 's922', 'canceled', 'encounters', 'y806c', 'tale n', 'cas9', 'pi3ka', '2930', '16600', 'mel202', 'karpas422', 'talens', '2030', '2031', 'a11012', '7007', '1950', '1951', '1952', 'slc3a2', '1955', 'allelespecific', '1956', 'credentialing', 'hdgc', 'hgdc', 'trifunctional', 'ksr2', 's226', 'kornev', 'l106', '195', 'f185', 'y164', 'd220', 'a70', 'i196', 'm128', 'e208', 's218a', 's222a', 't286', 'misnomer', 'heterotetramer', 'seg er', 'map1', 'multienzyme', 'boulder', 'withbamhi', 'terrific', 'tb', 'frost', 'mnacl', 'k97m', 's1206c', 'l1122v', 'pclg', 'bicocca', 'monza', 'tettamanti', 'haematology', 'gerardo', 'cecccon', 'cadore', '20900', '0264488362', '0264488363', 'mcecccon', 'gmail', '1541', 'ldk', 'astellas', 'k299ar300a', 'k29

9ar300b', 'k299ar300c', 'k299ar300d', 'm2ar500a', 'm2ar500b', 'm2ar500c', 'm2ar500d', 'supm2', 'npmalk', 'subcultured', 'atcacccctcattcggggtgtggccatggcg c', 'gagacctaaggcttcctccgagagaccgc', 'ctgccccgggttcatcctgatggagctatggcg', 'gaagccctgatcatcagcaaagtcaaccaccagaacattg', 'gttcatcctgctggagttcatggcgggggag ac', 'nm004304', 'codenaturation', 'abbotmolecular', 'tgcattatagtggacagcac', 'ctgtaaaccaggaggccgtac', 'tcagtaacggggacataaa', 'ggggctgtactgcttaaccag', 'tgcaccaccacccgttagc', 'ggcatggactgtggtcatgag', 'purelink', 'qseq', 'deindexe d', 'ceqr', '97c', '31f12', '3lcs', 'drugresistant', 'supm2ar500b', 'k299', 'k299ar300', '1324', 'm2ar500', '9984', '9068', '4491', 'retrotranscriptio n', 'd1203n', 'p1139s', 'singularly', '0222', 'kar300a', '5261', 'kar300b', '3546', 'kar300c', '5653', 'kar300d', '0134', '0182', '0314', '0405', '036 8', '4472t', '4544c', '4316c', '4538c', '4569c', '01165', '1243', '02911', '0 433', '07159', '09735', '1548', '3933', '01798', '01881', '1333', '01787', '1 082', '04105', '2831', '02491', '4224', '08548', '04576', '3552', '06797', '0 1249', '3503', '9623', '3849', '5337', '1645', '7582', '945', '5955', '3762', '1421', '005771', '4325', '4161', '3863', '7426', '1226', '1187', 'foresee', 'supm2ar500a', '4472', '4544', 'aminoacidic', '11122', '1248', 'counterselec ted', 'disadvantageous', 'm2ar300b', 'f1098', 'f317', 's1206', 'chaperon', 'karpasderived', '4575t', '11208p', '4556g', '4367c', '4559g', '4593g', 'r1214 h', '02179', '00002644', '00949', '03651', '01548', '02255', '04739', '9646', 'ported', 'metacore', 'r132s', '000258', 'ira1p', 'ira2p', 'arun', 'gottfrie d', 'r1276', 'resat', 'welti', 'r1391k', 'r1276p', 'methodspasmid', 'prepara tionthe', '19993', '3538', '4512', 'gggacaagttgtacaaa', 'aaagcaggctcatggaagt tctgacaaaaatccttc', 'ggggaccacttgtacaagaaagctgggtcaat', 'tggcagtatcttccagc aac', 'taatacgaactcaactatagg', 'rbgh', 'tagaaggcacagtcgagg', 'mutationssequenc ing', '3610c', 'r1204g', 'krkljus', 'r1204w', '3827g', 'r1276q', 'fahsold', '3902t', '11301r', '3919a', 'i1307v', '3971c', 't1324n', '3980a', 'e1327g', '4007a', 'q1336r', '4067a', 'e1356g', 'trov', '4171a', 'r1391g', '4193t', 'v 1398d', '4225a', 'k1409e', '4235c', 'p1412r', '4306a', 'k1436q', '4388c', 's1 463f', '4173a', 'medgen', 'analysisvariant', 'mutpred', 'novelthe', '000588', '0404', 'novelc', 'trovo', '0412', 'novelfigure', 'slideschematic', 'omnima x', 'promokine', 'nonmutagenized', 'assaycells', 'winglow', 'analysisan', 'analysismissense', 'mort', 'skippy', 'woolfe', 'esr', 'stadler', 'esss', 'fpr', 'grdthe', 'tasser', 'ambrish', '1nf1a', 'pettersen', 'slidethree', 'datathe', 'consurf', 'rhesus', 'panda', 'toad', 'wasp', 'mosquito', 'resultsin', 'slide assessment', 'deem', 'surmised', 'predictionsusing', 'thusberg', 'conservatio non', 'variantsto', 'r1204', 'discussionalthough', 'sterne', 'weiler', 'deplo yed', 'sensu', 'stricto', 'csrd', 'mangoura', 'starinsky', 'elbaz', 'linesalt hough', 'k1423', 'e1437', 'morphs', 'vacic', 'r1391', 'sermon', '1236', '127 9', '1286', '1382', '1430', '1432', '1437', 'microlesions', '171400', 'e768', 'bolino', 'ponder', 'chappuis', 'fagin', 't315', 'py1062', 'py905', 'multidoc king', '114k', 'gorre', '148k', 'ciardiello', 'decker', 'r8441', 'v778i', 's9 04c', 'macclesfield', 'munchen', 'y905', 'salvatore', 'springfield', 'scintil lator', 'refreshed', 'aitl', 'cancers6', 'arg5gln', 'mutations10', 'taste', 'vav1', 'retarding', 'gly17', 'amphotericin', 'mycoalert', 'sureprint', 'pas w', 'collard', 'furrow', 'wtrhoa', '9106', '9402', 'scavenger', 'ridley', 'se r221', 'concentrates', 'stiffness', 'resisted', 's651', 'gln145', 'q145h', 'e longinb', 'pro564', 'pro402', 'unhydroxylated', 'hyp564', 'ser111', 'his115', 'fractionate', 'depolymerization', 'pmal', 'phd1', 'ratcliffe', 'ppb75', 'sub complex', 'amylose', 'maltose', 'ultracentrifuged', 'ti45', 'immunopurify', 'bruick', 'mcknight', 'reticulolysates', 'ketoxoglutarate', 'feso4', 'ubc5 c', 'cdc34', 'electroneutrality', 'hydroxyproline', 'isobaric', 'namd2', 'ther malization', 'xj', 'vh1q145h', 'ubk48r', 'lys48', 'undeterred', 's65c', 's65 p', 's65a', 'nanosecond', 'isothermalisobaric', 'hinging', 'pvhls651', 'isost eric', 'pvhls65a', '163p', 'r64p', 's68w', 'r69c', 'e70k', 'vh1q145h', 'nod d', 'codd', 'p564a', 'p564g', 'lys389', 'lys391', 'lys547', 'leu63', 'arg64', 'arg69', 'glu70', 'hon', 'asn67', 'a69c', 'acetylcholinesterase', 'trp88', 't

yr98', 'trp117', 'hopeful', 'unicellular', 'dot11', 'setd1a', 'setd1b', 'mll5', 'disintegration', 'ptms', 'ptip', 'pou5f1', 'xaa', 'hoxa', 'cgis', 'hcf1', 'glcnac', 'dot1', 'pumps', 'meis1', 'fertile', 'schizophrenia', 'fibromyxoid', 'hepatoblastoma', 'steiner', 'flavopiridol', 'multipurpose', 'mox1', 'wolffe', 'egger', 'krogan', 'uncommitted', 'zygote', 'diaz', 'huntsman', 'engrailed', 'testa', 'flpe', '1554', '1228', 'hoxa5', 'hoxb4', 'hoxb5', 'trembl', 'invokes', 'condensing', 'nutritional', 'retards', 'decayed', 'brachyury', 'hoxb', 'suppressors', 'guenther', 'outgrowths', 'titin', 'gelatinized', 'neurofilament', 'livak', 'schmittgen', 'blebbing', 'misshapen', 'pou', 'blastomeres', 'salvaging', 'hyalinosis', 'd33', 'caldesmon', 'myf4', 'lo26', 'kepten', 'carlbad', 'c00', '84110', '586a2', 'palely', 'ovoid', 'syncytial', '50hpf', 'operatively', 'flares', 'bizarre', 'evolutions', 'committing', '11val560asp', '17asp816val', 'kittrp557', 'lys558del', 'thr670ile', 'preferntially', '9ala502', 'tyr503insalatyr', '11val555', 'leu576del', '13val654ala', '14thr670ile', '17asp820gly', '17asn822lys', 'volunteered', 'neng', 'nhri', 'asp820aal', 'cellytic', '1dscan', 'ircss', '4750', 'steiger', 'ckel', 'ff99', 'glidescore', 'asn822lys', 'validations', 'sawaki', 'cauchi', 'moths', 'pdgfraasp842val', 'val654ala', 'explorative', 'tamborini', '1696', '75k', 't670', 'distributing', 'srinivasan', 'respiration', 'xanthoastrocytoma', 'tmz', 'heterotetramer', 'aberrance', 'r172s', 'r132l', 'trended', 'decarboxylation', 'kilocalories', 'kinetically', 'gls', 'glutaminase', 'reform', 'energetics', 'electrochemical', 'aciduria', 'nmda', 'excitotoxic', 'ldha', 'diverted', 'diversion', 'bistris', 'vdac', '2020', 'uac', 'aconitase', 'preadipocytes', 'fabp4', 'hydroxylases', 'leonardi', 'channeling', 'hnf4', 'hbs', 'hnf1', 'extinguished', 'sox9', 'diethylnitrosamine', 'deploy', '24hrs', 'casette', 'inefficiency', '005896', 'lys4', 'dh2o', 'rgb', 'plugin', '1ml', 'homogenate', 'meoh', 'dasl', 'described36', '24h', 'p450', 's1h', 's1i', 's1p', 'gf', 'caught', 'koivunen', 'yilmaz', 'irreparable', 'patrick', 'bach', 'arg132', 'disabled', 'stereoisomer', '9q343', '350t', '1117p', '539t', '1180p', '572t', '1191h', '671t', 'm224r', '737g', 'r246k', '913g', 'g305r', 'g305w', '1526g', 'r509q', 'g1035s', '3290g', 'r1097h', '381deltgt', '128delv', '595delact', 'n198f199delinsi', '2075c', 'r692x', 'stagnated', 'anoxia', 'spastic', 'tsc24', 'parkinson', 'echodensity', 'prenatally', 'vacterl', 'tracheoesophageal', 'r692x3', 'antagonise', 'g305w15', '2074c', 'foetus', 'friedlander', 'doubtful', 'destroying', 'anthracyclines', 'ory', 'avidly', 'cbfa2', 'theaml1', 'paroxysmal', 'nocturnal', 'conray', 'datp', 'competitions', 'explosion', 'manufacturing', 'poison', 'war', 'cytogenetical', 'raebt', 'busulfan', 'gpa', 'survivor', 'lys83', 'arg177', 'cbfa1', 'cleidocranial', 'al37', 'anisotropy', 'photobleaching', 'lipidation', 'hvr', 'partitioning', 'natively', 'c118s', 'maleimid', 'dioleoyl', 'glycero', 'phosphoethanolamine', 'atto', '6his', 'stokes', 'dopc', 'allostery', 'jena', 'urbana', 'polarizing', 'diode', 'generalizations', 'radation', 'stothers', '1191', 'dispersive', 'nsf', 'atm75', '15513a02', 'atm78', '1609a01', 'atm81', '05079a02', 'atm82', '13601a02', 'dpp', 'investi', 'gations', 'gv41411', 'drill', '8019756', '8400574', 'geophysical', 'coding', 'electroperu', 'geodetic', 'moseley', 'huaynaputina', 'cerro', 'trapiche', 'ance', 'analy', 'bolzan', 'comments', 'paskie', 'vitch', 'doddroe', 'tope', 'contribu', 'posi', 'subre', 'odc', 'phyto', 'chrome', 'droxymethylglaryl', 'ductase', 'phytochrome', 'recog', 'nizing', 'caseins', 'seins', 'compilations', 'quenced', 'howev', 'outnumbers', 'hydrophilicity', 'asymmetries', 'i986', 'uous', 'drophilic', 'nized', 'hydropa', 'p730', 'glutaryl', 'aminotrans', 'ferase', 'helyldvtapedpnecavsqifpd', 'vmlavqegidlltfppapgspepph', 'heagfppsddegeeefvldyveh', 'rtcgmfvyspvsepepepepepar', 'recnsstdscdsgpsntppeih', 'rnyldldydsqvpyfycdeeen', 'fyqqqqselqppapsekd', 'kfellptplpsr', 'rgndggggsfstadqlemtellggdm', 'vnqsficdpddetfik', 'hsvcstsslylqdlasaasecidpsvff', 'pyplndssspk', 'kscasqdssafspssdsslstesspqg', 'speplvlh', 'heetpptfssdseeqedeiedvvsvek', 'rsesgspssaggh', 'hcmddlllpqdvveffgpsealr', 'rvsgapaaqdptvetpgvapapa', 'tpwplssfvpsqk', 'hsvvvpyeagseyttih', 'hspadsfssmgspvnaqdfct

dlav', 'ssanfiptvtaistspdlqwlvqpa', 'lvssvapsqtr', 'kveqlspeeeek', 'kipddlgfp
 eemsvasldltgg', 'pevatpeseeafatlplndpepk', 'ktepfdfflpassr', 'hsgs1gmmpmate
 leplctpvvtc', 'tpsctaytssfvtpeadsfpscaaah', 'kgsssneppssdsllssptlial', 'keqpg
 sddeedesneqtfmyyvndgvyg', 'sfncilydh', 'hgfppeveeqddgtlpmscaqesgmdr', 'kveqegy
 pnessk', 'hgclpeesaspar', 'hlqymenmnsiaslvmavvvneneed', 'deaeseqpaqqk', 'rtt
 psyvaftdser', 'ketaeaylgesitdavitvpayfnfsqr', '445', 'ktfstysdnqpgvsiqvyeger',
 'klssveeepgvsqdr', 'rtqeleielpsepr', 'rpneelqilesaek', 'hfpefendvefter',
 'kemeaesissseeivpnsvqek', 'rqfyqldaypsgawyyvplgtqytda', 'psfsdipnpigensek',
 'releelnvpgeiveslssseesitr', 'kypvqpfetesqsltldvenlh', '365this', 'fri', 'ut
 call', 'ela', 'poooo', 'phyt', 'drophobic', 'fea', 'tern', 'iden', 'tifies',
 'secre', 'tory', 'parted', 'whit', 'ney', 'pyk', '155dhf', 'hydrophobicities',
 'polyglycine', 'portable', 'valley', 'proteolyt', 'degrad', 'reticulo',
 'cyte', 'polyglutamic', 'resi', 'dues', 'phorylation', 'teases', 'calpains',
 'plm', '234this', 'calpain', 'conju', 'gation', 'tain', 'macromolecules', 'in
 tracellu', 'phy', 'tochrome', 'iiij', '41z1', '300o', 'ot', 'lld', 'chy', 'fe
 r', 'abp', 'cab', 'ald', 'gpd', 'dhf', 'tpi', 'myo', 'ppa', 'cpa', 'adh', 'so
 d', 'ygr', 'vhem', 'parvalbumin', 'trypsinogen', 'tase', 'chymotrypsinogen',
 'arabinose', 'anhydrase', 'triose', 'myoglobin', 'carboxypeptidase', 'hem',
 'oxyhemoglobin', '367this', 'htpr', 'deh', 'linger', 'schimke', 'behav', 'io
 r', 'latory', 'hypothe', 'specu', 'rect', 'expects', 'attaching', 'flipi', 'b
 ouska', 'okosun', 'pinyol', 'relander', 'gse40989', 'memo', 'gse37088', 't283
 a', '2124', 'musgrove', 'naderi', 'v287d', 'q276', 'elenitoba', 'umcn', 'sola
 l', 'ligny', 'unmc', 'hsi', 'llanos', 'egle', 'oricchio', 'vavbcl2', 'shp16',
 'shrb', 'peanut', 'nbaek', 'dh16', 'oltersdorf', 'wshu', 'ccnd', 'nct0111118
 8', 'obatoclax', 'sbl', 'pathogenesis5', 'eb1', 'hivbl', 'mechanisms7', 'intr
 athecally', 'setting8', 'cases9', 'psnvs', 'dgyw', 'wrch', 'work10', 'dlbcl
 9', 'dlbcl13', 'potency15', 'ddx3x', 'psnv', 'dpf1', 'act16a', 'bl16', 't28
 3', 'p284', 'i290', 'r271fs', 't283p', 'oncogenicccnd3', 'stability18', 'p284
 1', 'i290r', 'phase19', 'c10stop', 'y49stop', '154v', 'v55', 'p56t', '164f',
 '170p', 'q71stop', 'v82', 'q100', 'd93', 'a102d', 'v557e', 'd561e', 'm572k',
 'development21', 'n551k', 'v557', 'd561', 'm572', 'binding23', 'l64', 'kink
 s', 'shtcf3', 'p56s', 'e68', 'c47', 'q63fs', 'biotag', 'nncacctg', 'ggcagct
 g', '81e', '03e', 'dlbcl5', 'responses25', 'pou2af1', 'ltb', 'chains22', 'lin
 es13', 'dlbcl28', 'mln120b', 'expected29', 'pathway32', 'trials34', 'mir17h
 g', 'expression36', 'methods5', 'transcription38', 'responses16', 'viability3
 0', 'signaling43', 'reaction16', 'unsafe', 'delivery8', 'ethic', 'beadstudi
 o', 'fhit', 'antiviral', 'vignali', 'whitehouse', 'baf60a', 'aliases', 'super
 ose', 'gajiwala', 'burley', 'immunomodulatory', 'cri', 'crs', '120k', '76k',
 'scheffner', 'mccarthy', 'caveolin1', 'localizer', 'fancn', 'synchrony', 'ph
 2ax', 'eufa1341', 'p2f1', 'n42', '1314', '1557', '1364', 'cc2', 'coiped', 'k1
 4a', 'l21a', 'y28a', 'l35a', 'e42a', '1424', 'hcc937', 'coimmunostaining', 'm
 acrocomplexes', 'rap80', 'sowter', 'ashworth', 'venkitaraman', 'narod', 'lowe
 nfels', 'maisonneuve', 'simard', 'daniels', 'lomonosov', 'kraakman', 'zwet',
 'fuks', 'marston', 'braf35', 'marmorstein', 'usp11', 'schoenfeld', 'hussai
 n', 'bccip', 'coats', 'pellegrini', 'tarsounas', 'brh2', 'esashi', 'flj2181
 6', 'loc79728', 'immunobeads', 'netn300', 'w300', 'iped', 'op95', 'wi38', 'pa
 lb2n200', 'tp15', 'postnuclear', 'pnl', 'ns100', 'ns420', 'np420', 'mnase',
 'pmn', 's420', 'p420', 'netn100', 'brca2f9', 'coiumunoprecipitate', 'chodos
 h', 'netn420', 'codepleted', 'permeabilization', 'refocused', 'postdamage', '6
 954', 'brca2f8', 'palb2f4', 'preextracted', 'coimmunostained', 'refocusing',
 'unfeasible', 'nonfocal', 'preextraction', 'enaction', 'rk5c1', 'b2n60', 'w3
 1r', 'w31c', 'g25r', 'r18h', 'wtbrca2', 'i27v', 'f12v', 'guise', 'insures',
 'nordling', 'scds', 'multisite', 'molded', 'bioessays', 'riskof', 'fitzroy',
 'heierhorst', 'princes', '3065', 'jheierhorst', 'svi', 'bies', '20204', 'dnad
 ependent', 'esc4', 'fanc', 'lcd1', 'mdt1', 'mrc1', '53kda', 'smc1', 'betterch
 aracterized', 'tqs', 'dnadamage', 't68q', '2609', '2648', 'scd1', 'scd2', '13
 69', 'scd3', '1716', 'claspin', 'rad26', 'rad17', '967', '1113', 'asciz', 'pr

4otein', 'crb2', 'overrules', 'pt68', 'fha1', 'fha2', 'autoactivate', 'damage induced', 'pondr1', 'faciliate', 'ptq', 't11q', 'atrlike', 'intial', 's345q', 's317q', 't312q', 's354q', 's363q', 't603q', 'pt603q', 's1387q', 's1423q', 's1457q', 's1524q', 's6q', 's25q', 's29q', 's784q', 's278q', 's343q', 's397q', 's615q', 'subclusters', 't634q', 's637q', 't645q', 's599q', 's614q', 's604q', 's864', 's895', 't906q', 's934', 's635q', 's645q', 'rad1', 'damagedependent', 'ubiquitinylation', 'sq5', 's222q', 's1404q', 's957q', 's966q', 'agein g', 's15q', 's37q', 's395q', 's664q', 's745q', 's1073q', 's1078q', 's1096q', 's1116q', 'brietta', 'nora', 'tenis', 'carolyn', 'mcnees', 'lindus', 'conlan', 'scdencompassing', 'brctand', 'overruling', 'phosphosq', 'ofs', 'damagere sponse', 'embark', 'reassure', 'q581', 'infertile', 'polyglutamine', 'infertility', '1647', 'xq11', 'oulu', 'gynecomastia', 'fsh', '157q', 'primates', 'ais', 'hypospadias', 'micropenis', 'dsd', 'urethral', 'perineal', 'prism310', 'applera', 'mamld1', '000044', 'pantherdb', 'psvar0', 'kpn', 'pcr1', 'krefel d', 'androgendb', 'gonads', 'audi', 'chorionic', 'luteinizing', 'llerian', 'remnants', 'ria', 'milwaukee', 'ramensky', 'rearing', '15t', 'changeb', 'surg e', 'ferlin', 'deeb', 'concurring', 'turkish', '1809', 'xpo5', 'd1709', 'carc inosarcoma', '1709', 'stalls', '34g', 'gly12ser', 'gly12arg', 'gly12val', '38g', 'gly13asp', 'dtpp', '37g', 'judge', 'assurance', 'burt', 'sigmoid', 'pgem teasy', 'g382d', 'exhaustion', 'his1047', 'glu545', 'glu542', 'nish2', 'sail', 'boat', 'hull', 'navy', 'topologies', 'asp560', 'asn564', 'asn345', '2rd0', 'boar', '1e8x', '865', 'lys729', 'lys867', 'arg349', 'lys410', 'arg412', 'lys413', 'lys416', 'arg38cys', 'arg38his', 'arg88gln', 'arg38', 'arg88', 'gln738', 'asp743', 'asp746', 'framed', 'glu453gln', 'cys378arg', 'lys379', 'arg340', 'his701', 'gln661', 'leu956', '1050', 'phosporylated', 'met1043ii', 'wortmaninn', 'k802', 'wormannin', 'p850', 'gly376arg', 'glu365', 'immunoprecipitate', 'exon6', 'r195', 'kle', 'ser118', 's118a', 'e47k', 'g75r', 's80r', 'p94a', 'm117v', 'r121q', 'd140g', 'greb1', 'abca3', 'fulvestrant', 'catalouge', 'masatoshi', 'hagiwara', 'pcin4', 'genechem', 'genepharma', 'spop1', 'gg augauguaaaugagcaa', 'ggacagcgactctaatct', 'acagacuuccggaguaccug', 'ab137537', 'ab72999', 'ab134175', '8644', '6652', 'mm5', '101r', 'tiangen', 'spurious', 'somers', 'phosphodiester', 'benomyl', 'pericentromeric', 'wake', 'oppositely', 'ccr4', 'opted', 'phylogeny', 'naei', 'affy', 'schiffman', 'lengthening', 'kaneda', '500kb', 'verhaak', '27k', 'thalamus', 'pons', 'striatum', 'foxg1', 'hypermethylator', 'volcano', 'thalamic', 'telencephalic', 'ligon', 'gibbon s', 'dkfz', 'russia', 'rbzburg', 'capper', 'dia', 'pfister', 'heaphy', 'thalasaemias', 'classification8', '278k', 'pannets', 'described22', 'previously20', 'h3k36me2', 'r283c', 'prototypes', 'r283p', 'unrecognizable', 'y126c', 'vatting', 'koptk1', 'nash', '2516', 'dns', 'pf382', 'hrt1', 'mislocalizes', 'mrk', 'dnd41', 'tall11', 'rpmi8402', 'hsb2', 'phosphodegrons', 'esis', '2512', 'sup t13', 'shq1', 'subverting', 'tmpyp4', 'quadruplexes', 'quadruplex', 'myopathy', 'bmj', 'innervated', 'trophic', 'r181c', 'denaturating', '1tup', 'histidines', 'contours', 'geometrical', 'novabiochem', '1uwh', '1aq1', 'aromat', 'beg', 'antileukemic', 'mutation19', 'post docking', 'methylphenyl', 'unactivate d', 'helping', 'appendage', 'wnt3', 'hamburger', 'hindlimb', 'intercross', 'fgf8b', 'electrophoreses', 'amphibian', 'otic', 'nilsson', 'farnham', 'cter', 'ylp', '1480', 'huen', 'starita', 'meza', 'ad1', 'nonsignificant', 'nup133', 'humanmutation', 'straina', '1755', 'e33a', 'c44y', 'szabo', 'coyne', 'phelan', 'jackknife', 'multicopy', 'pesc', 'westmoreland', 'kaganovich', 'kopito', 'meriin', 'ssa1', 'sarkosyl', 'nucleoporin', 'nup49', 'doye', 'makio', 'c47f', 'm1r', 'analyzeda', 'studyf', 'nomenclatureb', 'bicc', 'brca1d', 'iarce', 'ablank', 'bhgvs', 'cbic', '5559c', '5558', '5559insa', 'dumd', 'elovd', 'fsm all', 'met1', 'met18thr', 'glu33ala', 'cys39tyr', 'cys44tyr', 'cys47phe', 'cys61gly', 'met1652ile', 'ala1708glu', 'pro1749arg', 'met1775arg', 's1841a', 'ser1841ala', 'tyr1853x', '54atg', '40atg', '49atg', 'tyedmers', 'outeiro', 'lindquist', 'reorienting', 'delocalized', 'holt', 'founded', 'congress', 'chengdu', '2039', '1287', 'drawbacks', 'q120e', 'a159t', 'g216r', 'g248v', '1272

f', 'p380r', 'del571', 'ins581', 'karger', 'amenorrhea', 'swellings', 'underv
irilised', '2q12', 'heemers', 'mcewan', 'jenster', 'skel', 'inen', 'kazemi',
'esfarjani', 'arntd', 'q571', 'giwercman', 'pgre2', 'knoke', 'ppem', 'barbul
escu', 'glycylglycine', 'mgso', 'nanolight', 'xs3', 'invistrogen', 'peroxidas
econjugated', '000369266', 'hygromycinb', 'lavery', 'shimadzu', '1501', 'wave
lengths', 'trimethylamine', 'tmao', 'solutes', 'clitoromegaly', 'pt5', 'undes
cended', 'labial', 'perineo', 'scrotal', 'bifid', 'scrotum', 'pt6', 'peno',
'tilley', 'tadokoro', 'cuccaro', 'gluc', 'digitised', 'uugcugcugcug', 'cagca
gcagcaa', 'yeap', 'gre2', 'endogluc', 'chymo', 'fqnlf', 'mage', 'rap74', 'bet
ney', 'a236t', 'e238g', 'tramp', 'ara160', 'ara70', 'helsen', 'g248v1', 'zucc
arello', 'ala159', 'pstls1', 'asn235', 'gly248', 'pro380', 'hancock', 'underv
irilisation', 'idkowiak', 'e2k', 'arango', 'humanised', 'metabolise', 'beberm
eier', 'utilisation', 'interspaced', 'androgenresponsive', 'srcap', 'ep400',
'3238', 'rspo3', 'pharmacogenomics', 'c106', 'c125', 'c70', 'c84', 'c99', 'h
ct8', 'ls180', 'ls411', 'ls513', 'h716', 'h747', 'skco', 'c2b', 'sw1417', 'dc
td', 'colo678', 'gp2d', 'ht115', 'cck81', 'cocm', 'hsrrb', 'vaco10', 'vaco4
s', 'mcbain', 'kclb', 'hdc101', 'hdc54', 'hdc82', 'hdc87', 'hdc90', 'hca46',
'hra19', 'kirkland', 'postgraduate', 'rw7213', 'calabresi', 'roger', 'provid
ence', 'qmir', 'agrf', 'beadchips', 'gse55832', 'oncosnpv2', 'methylation', 'w
eisenberger', 'cacna1g', 'ngn1', 'pmr', 'mismapping', '937', 'meth', 'nn3',
'nshp', 'edm', 'subchromosomal', 'a2bp1', 'macrod2', 'bcl91', 'fat1', 'trend
line', 'therapeutics1', 'agtr1l1', 'aplnr', 'lphn3', 'gnao1', 'tumorigenesis
1', 'technology10', 'her21', 'deubiquitinases', 'dubs', 'sorting10', 'abundan
ce12', 'itis', 'somaticmutations3', 'ofmutation', 'cosmic13', 'permb', 'fusio
n14', 'approach16', 'hr1', '3420', 'expressway', 'wartik', '16802', 'nature09
208', 'cmklr1', 'mas1l', 'ptgfr', 'metabotropic', 'grm1', 'bai2', 'r265', 'os
teodystrophy20', 'r243h', 'signalling21', 'reported22', 'nek2', 'tlr4', 'acvr
1b', 'alphabetical', 'thegpcr', 'g49', 't118', 'c189', 't420', 'a442', 'w46
1', 'g586', 'c819', 'i869', 'a1024', 'r1050', 'r1124', 'c1148', 'm1258', 'f13
78', 'g1404', 'n1475', 'd1449', 'y1479', 'p1510', 'm206', 'w215', 'i574', 'i5
75', 'k669', 's691', 'a808', 'g523', 'a282', 'v150', 'l76', 'grm3', 'grm4',
'grm5', 'grm6', 'grm7', 'kcf', 'anf', '7tm', 'tsp', 'gps', 'ncd3', 'dagger',
'eif3s3', 'edd1', 'mc3r', 'trim11', 'usp21', 'nlrp3', 'fbxw7npy2r', 'mchr2',
'npy5r', 'trim37', 'crkrs', 'r201h', 'r265h', 'q226', '1cs4', 'others4', 'n2
34', 'studies2', 'prkd1', 'dimerization28', 'dimerization29', 'cancers30', 'g
lioblastomas5', 'proteinaltering', 'nrg3', 'reported2', 'samples6', 'a572v',
'r724h', 'chiron', 'biscistronic', 'geneeditor', 'lactogenic', 'ha1', 'tripu
re', 'mjak1', '21r', 'tyr931', 'incb018424', 'phe116', 'aut', 'chromaspin',
'ggagt', 'gcagtatctccctct', 'ccatgc', 'ccaggcactcatatca', 'mactin', 'gct
g', 'gaaggtggacagttag', 'ctctggctcc', 'tagcaccatgaag', 'myiq', 'pmegix', 'py1
034', '3331s', 'py1007', '3771s', '9131s', 'py694', '9351s', '91011', '3332',
'24b11', 'cmp6', 'microfiltered', 'canberra', 'meriden', 'epigenic', 'f635v',
'y652h', 'v658i', 'tyr652', 's1043i', 'f958v', 'f958c', 'f958l', 'd895h', 'e8
97k', 't901r', '1910q', 'glu71lys', 'arabinoside', 'ara', 'f958', 'p960', 'ac
tive26', 't970m', '1783f', 'oxidants', 'surrounds', 'mrp', 'glucuronosyltrans
ferase', 'g430c', 'g364c', 'sourced', 'wellesley', 'eisen', 'depc', 'nl20',
'h1435', 'h292', 'pub6', '1921', 'bronchi', 'telecanthus', 'cdk3', 's10c',
'nephroblastoma', 'r960', 'covalab', '2033', 'labia', 'syringomyelia', 'leuk
emia1', 'defects3', '218cright', 'gright', 'arrowa', 'cascade2', 'tttright',
'arrowaaa', 'f71k', 'site9', 'syndrome4', 'glu76', 'collects', 'iaslc', 'y18
5', 't599', '707', '1wt', '4032', 'wimmer', 'baccarini', 'r188', 's365a', 'ba
ljuls', 'heterodimerisation', 's729', 'rafinst', 'ritt', 'r481', 'takt', 'the
rrien', 'recalcitrant', 'blunts', 'f468', 'kubo', 'polybasic', 'exempted', 'p
mig', 'terai', 'hahb', 'hingorani', 'hahbraf', 'criss', 'matos', 'cistronic',
'herr', 'disintegrated', 'uncaged', 'capon', 'shirasawa', 'patho', 'janne',
'accountable', 'sobering', 'imitate', 'sensitise', 'trialled', 'c13', 'resum
ed', 'hspg', 'diii', 'fgfr1b', 'fgfr3b', 'fgf17', 'p0', 'synostosis', 'y766',

'dispensible', 'anteroposterior', 'posses', 'pfeifer', 'midface', 'cys342ser', 'stump', 'tyr375cys', 'intramolecularly', 'gnrh', 'telencephalon', 'tyr177', 'lagged', 'hspgs', 'mfp1', 'mfp3', 'biphosphate', 'differentiation26', 'survival36', 'degradation38', 'cancer48', 'cancer56', 'mmset', 'gammopathy', 'bloodstream', 'oct1', 'cancer111', 'zmym2', 'iap1', 'ang2', 'hyperphosphatemia', 'fv', 'reassuring', 'birrer', 'nontarget', 'xcelligence', 'electricall', 'impedance', 'genecopoeia', 'tnfr1', '517g', '631g', '572a', '520c', '587g', '617c', '1735', '100965', 'unicancer', 'ugg', '000059', 'immortalised', 'mazoyer', 'crcl', 'poulain', 'bioticia', 'caen', 'br2', 'gaccagatctgttatacctttgcctgaga', 'ttta', 'mlu', 'gaccacgcgt', 'gcttgacaccactggactacc', 'pcas2', 'ko1f', 'tgacgtcg', 'ccgcccacac', 'pcas2r', 'attgggttggacttgggttc', 'compatibles', 'gly173cys', 'arg174cys', '522t', 'arg174arg', 'asp191gly', 'ser196asn', 'f7', 'ser206cys', '625c', 'leu209phe', 'val211ile', '7008', 'voluntary', 'tggtatgtggagttg', 'tttc', 'catgacttgcagttctcttt', 'ctgccgtgtaccaatc', 'ttcagatgttttcattt', 'pyromark', 'symbolise', 'br2pyro520f', 'ccaaagtgtgaa', 'br2pyro617f', 'cacc', 'acccaccccta', 'ssf', 'int11', 'splices', '631del', 'gly173serfsx19', 'asp191', 'ile210del', 'val188serfsx19', '518g', 'gly173val', '581g', 'trp194x', '231t', 'thr77thr', '8165c', 'thr2722arg', '7992t', 'ile2664ile', 'cftr', 'culty', 'unforeseen', 'upcoming', 'y591c', 'v579a', 'f590gy591d', 'sscps', 'mononucleated', 'hypaque', 'flt3jm', 'caattagtgaaagccag', 'tgatccttagtacctt', 'ttcatattctctgaaatcaacgtagaagt', 'red640', 't cattatctgaggagccgtcacct', 'hyb', 'metabion', 'martinsreid', 'promocell', '8392', '5441', 'reyeydl', 'cssdneysfyvdreydlkwefprenl', 'glvqvtgssdneysfyvdreye', 'significant', 'v579', 'tina', 'initially', '590g591d', 'encouragment', 'autoactivated', 'f590', 'y591d', 'prohibits', 'offsets', 'r529', 'w561', 'v560a', 'v550a', 'i537a', 'd551a', 'viremia', 'pyxxm', 'cells14', 'formation17', 'notation', '005026', 'lrt', 'cancer22', 'activity22', 'bet', 'degranulation', 'cd57', 's244', 'derangements', 'unrestrained', 'tumorigenesis36', 'described47', 'loads', 'pasli', 'novelty', 'akap13', 'lish', 'snd1', 'cancer28', 'tetra tricopeptide', 'teal', 'phox', 'puts', '15c', 'conserving', 'disregarded', 'prkaca', 'fibrolamellar', 'gtex', 'htseq', 'gdac', 'tale', 'vmp1', 'rb a', 'dba', 'hemangioendothelioma', 'cytoband', 'fluors', 'angiosarcomas', 'tiele1', 'dms', 'pl5', 'd538', 'znf384', 'snb19', 'medications', 'cscl', 'nalgene', 'jersey', 'expresshyb', 'prehybridization', 'cushions', 'gelling', 'sustainable', 'rx', 'primes', 'ros1g2032r', 'preemptively', 'momentum', '6976', 'selectionresultshigh', 'jak3a572v', 'sudh11', 'mgcd', 'jnj38877605', 'shpten', 'v2098i', 'g1971e', 'g2032r', '11947r', '11982f', '11152', '11982', 'c2060g', 'e1935g', 'rosg2032r', 'sectiondiscussionthe', 'accentuating', 'reciprocity', 'awad', 'eder', 'sectionmethodsadditional', 'dmsz', 'hereon', 'methanethiosulfonate', 'kin3f', 'kin1', 'p3tplux', 'smad3n', 'smad3c', 'capitalized', 'with32p', 'and9', 'cand', 'slowest', 'p3tp', 'alk4', 'aurka', 'figure5d', 'figure5e', 'figure6a6a', 'figure6c6c', 'figure7b7b', 'figure7d7d', 'ttk', 'stevens', '151623', 'frequencya', 'durocher', 'oncogenes4', 'cells3', 'pairings', 'd9s149', 'd9s66', 'platt', '1250c', 'thr417ile', 'his68', 'phe158', 'phe216', 'fugu', 'his105', 'his206', 'thr417', 'ser35', 'his68arg', 'his105arg', 'ser35cys', 'his206asp', 'phe158cys', 'phe216asp', 'hcv29', 'phe216ala', '104c', '648t', 's35c', 'h68r', '203a', '473t', '616c', 'benzyl', 'optically', '7j', '1916r', 'e412v', 'aggresomal', 't417i', 'gratuitous', 'nonrearranged', 'dysgeusia', 'dizziness', 'hypophosphatemia', 'hypotension', 'v185g', 'r217c', 'v326a', 'hpms2', 'matchmaker', 'gatc', 'gyrase', 'mlhs', 'mshs', '19hmlh1', 'nfdht', 'hpms1', 'nfht', 'criteriaa', 'statusa', 'mmrb', 'yes24', 'yes25', 'na25', 'yes30', 'yes29', 'no33', 'na33', 'yes31', 'na31', 'yes34', 'na34', 'noc', 'yesc', 'yes35', 'yes39', 'afulfillment', 'criteria22', 'd18s5824', 'd3s1029', 'mfd27', 'mfd4729', '39bthis', 'cpierre', 'dslicing', '2484', 'u07343', 'aac50285', 'alw44i', '67mer', 'kurt', 'ballmer', 'schere', '1116', 'hmh11', '21f10', 'n357', '13mp2', 'hmutl', 'nr9162', 'csh50', 'indoyl', 'colorless', 'assay42', 'pms1p', 'possible29', 'pierre', 'r261h', 'pr

ecaution', 'hpv16', 'fusions7', 'deaminases', 'chd7', 'synthetically', 'episo me', '125b', 'cancer23', 'raloxifene', 'cancers23', 'cells26', 'uroplakins', 'pkm2', 'hsp90aa1', 'allyl', 'demethoxygeldanamycin', 'phosphotransferases', 'radicicol', '1790', 'y791', '062', 'dok4', 'y687', 'y806', 'groot', 'utrech t', 'neurotropic', 'mtcs', 'asp892', 'retm918t', 'families1', 'nfatc2', 'uppercase', 'ssx', 'igfbp3', 'igfbp5', 'dsrct', 'ribominus', 'templated', 'quadrants', 'analytically', 'constitutionally', 'wick', 'monzon', 'subtleties', 'brisbane', 'dracopoli', 'x1', '31f', 'espoo', 'temp', 'btx', 'msec', 'microlitres', 'minigel', 'pont', 'compartmentalized', 'cdc37', 'distorting', 'massague', 'improperly', 'midway', 'radically', 'myxopapillary', 'korbel', 'cytoscape', 'subgrouping', '0068', 'sturm', 'maml2', 'nn', 'hemangiosarcoma', 'p3k', 'glutamates', 'cefs', 'engulfing', 'leukosis', 'linings', 'medicated', 'ithas', 'unequalled', 'myristylated', 'spafas', 'preston', 'mcintyre', 'lakeside', 'candled', 'drilled', 'injuring', 'webs', 'gipf', 'oberfrick', 'gagp19', 'volker', 'khco3', 'kontes', 'grinder', 'homogenizing', 'wings', 'coaggregation', 'nonaggregating', 'depositions', 'amyloidoses1', 'therapy3', 'uiuc', 'domain5', 'unfolded6', 'proposed9', 'radiochemotherapy10', 'unexplained12', 'p2501', 'e258v', 'perinucleus', 'aggresomes', 'integrity15', 'reports16', 'octamers', 'tango17', 'core18', '683', 'findings19', 'polymers', 'esi', 'tango', 'p53e258v', 'p53r175h', 't256r', 'mog', 'ccm', '1376', 'reports20', 'thioflavin', 'i254d', 'r175h13', 'coaggregate', 'i255r', 'i255d', 'heterotetramerization', 'transactivity', 'k305n', 'noxa22', 'tetramers7', '1344p23', 'cells13', 'undetermined25', 'tetramerize', 'independent26', 'hydrophobicity27', 'tap73', 'i324r', 'i274r', 'b28', 'pathways30', 'mutations34', 'autophagy9', 'activity36', 'unifies', 'oligomerization38', 'tetramer39', 'tetramerization7', 'development41', 'interactions42', 'vitro43', 'accumulation44', 'retention45', 'signal46', 'synergies', 'activation9', 'neurodegeneration', 'stochasticity', 'proteostatic', 'tango47', 'literature48', 'else', 'field49', 'ypydvpdy', 'elsewhere50', 'pwo', 'dykdddk', '10838', 'vergata', 'elsewhere13', 'borohydrate', 'alexa594', '18mm', 'nativepage', 'manuals', 'tricine', 'cathode', 'nativemark', 'microfarads', 'h3po4', 'clustalx', 'y627', 'impurities', 'line weaver', 'burk', '1289', 'tyr877', '1834r', 'pjs', '175200', 'extraintestinal', 'giardiello', 'hemminki', 'mehenni', 'intussusception', 'fob', 'dash', '10s', 'd176n', '19f', '898', 'amos', 'ylikorkala', '903delg', 'zuo', 'habit', 'sclcs', 'biologia', 'idibell', 'strad', 'mo25', 'thr172', 'zeb1', 'nr4a2', 'aigar', 'bioenergetic', 'azd', 'zur', 'tiff', 'caski', 'biallelic', 'prism5', 'adenyl', 'tyr185', 'activatory', 'leu242', 'trp308', 'i177', 'edn1', 'vasculogenesis', 'egln1', 'hif2a', 'somatostatinomas', 'eldletlapyipmdgedfq', 'somatostatinoma', 'hif3', 'jafri', 'pollard', 'epas1', 'ala530thr', 'ala530val', 'ala530', 'pro531ser', 'pro531leu', 'ile533', 'pro534del', 'pro534', 'asp536del', 'asp539tyr', 'ssb', 'pgp', 'mre11a', 'mdr1', 'irif', 'h53bp1', 'hpmc', 'ssbs', 'rnf8', 'rnf168', 'acetamide', '2714', '061', 'tecan', 'ware', 'hene', '5a4', 'hemicolecotomy', 'leucovorin', 'oligomerisation', 'burgelab', 'd935n', 'rationalizations', 'bordered', 'ashkenazim', 'athens', 'visualise', '2283', 'hogervorst', 'dyedeoxy', 'confidentiality', '12delt', 'ivs7', 'delt', 'q356r', '5075', '5685', '26t', 'multinational', 'epiphomena', '17aag', 'pccs', 'cystadenomas', 'epididymis', 'endolymphatic', 'amar', 'rotkreuz', 'decremental', 'c634y', 'hlrcc', 'phaeochromocytomas', 'fibroids', 'catalyse', 'myometrium', '3131', 'alterations1', 'cancer4', '1140', 'occured', 'algorithm18', 'citing', 'bok', 'melanoma7', 'h2110', '20q11', 'mice41', 'permuted', 'hirning', 'hatton', 'cgtg', 'grandori', '3004', 'yosef', 'nesbit', 'dildrop', 'moroy', 'stanton', 'spencer', '3005', 'henriksson', 'hann', '3006', 'bls', 'lidereau', 'locker', 'spandidos', 'slot', 'highgrade', 'mxi1', 'steck', 'vlietstra', '3008', 'liaw', 'cristofano', 'finley', 'microcell', 'hajdu', 'subungual', '3009', 'nonlymphocytic', 'jennings', 'mills', 'onethird', 'tsuda', 'abortive', 'prothymosin', 'eilers', 'ering', 'ikegaki', '3010', 'norris', 'tenuous', 'salgia', 'skarin', '3011', 'kawashima', 'mbii', 'mbiv', 'myce'

r', 'triplex', 'neuroepithelioma', 'crossover', 'dnp', 'septic', 'phosphorodi amide', 'avi', 'benzodiazepine', 'sarkar', 'vaccine', 'vaccines', 'berg', '2800', 'hsps', 'bullet', 'marcu', 'bieche', 'isola', 'lombardi', 'lccc', 'g eorgetown', 'antidigoxigenin', 'nonradioactive', '896', 'mcnemar', 'schraml', 'utilised', 'astonishing', 'chronicle', 'myc1', 'leukaemogenesis', 'myelocyto matosis', 'unravelled', 'chimaeras', 'immunoglobin', 'protein13', 'emu', 'dna 27', 'eisenman', 'protein32', 'ensued', 'business', 'topography', 'transformation53', 'cells62', 'region63', 'p67', 'hras1', 'vaux', 'cory', 'frustration', 'overshadowed', 'blackwood', 'amati', 'knoepfler', 'autosuppression', 'asa', 'ziff', 'nfy', 'miz1', 'littlewood', 'ertam', 'exacerbating', 'collided', 'mai', '1beta', 'cul7', 'eischen', 'larsson', 'interprets', 'surprises', 'myc l', 'kohl', 'malynn', 'crossroads', 'canntg', 'amplication', 'fulminant', 'protooncogenic', 'yen', 'spotts', 'sola', 'bends', 'nikiforov', '182kb', 'miz', 'inr', 'seoane', 'lncrnas', 'fbp', 'melts', 'salghetti', 'tansey', 'prevalent ly', 'usf', 'circadian', 'orian', 'zeller', 'rahl', 'gandarillas', 'iritani', 'lin28', 'promiscuously', 'nonproliferative', 'lipogenesis', 'rhythm', 'semen za', 'opt', 'ramsey', 'glucosamine', 'yuneva', 'giuriato', 'interrupt', 'selt zer', 'deoxyglucose', 'fermentation', 'strategically', 'murga', 'medica', '50 80', '6600', '1305', 'gui', 'landes', 'pyrrolidine', '1643', 'phosphoegfr', '1644', 'h2so4', 'heatinactivated', 'laminar', 'acl', 'd837a', '3t3s', '223 2', '2234', 'y317', '9132', '9131', 'm67', 'phrl', 'luminoskan', 'eberhard', 'bftc909', 'cal29', 'u3', '639v', 'lucc1', 'lucc3', 'preetrosuper', 'p110a', 'panakt', 'luminata', 'forte', 'a1066v', 'sgk3', 'v136i', 'outs', 'healed', 'mm1', 'yoshioka', 'clostridium', 'exoenzyme', 'ribosyltransferase', 'rhoc', 'tufts', 'ix70', 'transblotted', 'rhogdi', 'ribosylated', 'legion', 'filopodial', 'cdc42hs', 'dominantnegative', 'youth', 'military', 'shaanxi', 'ges', 'underlay', 'traversed', 'twothirds', 'tiation', 'borrmann', 'humangastric', '6 244', 'endurance', 'chemosensitivities', '6245', 'difficile', 'women1', 'arias', '3180', 'cncr', 'v115', 'souza', 'biosci', 'boureux', 'mellor', 'annu', 'fritsch', 'abo', 'haematol', 'merajver', 'sz', '9002', 'hauser', 'mackinno n', 'smggds', 'phila', '1940', '0216', 'amoeboid', 'pharmacol', '1227', 'cerv icitis', 'skimmed', '1616', 'fujisawa', '2727', '1317', 'workman', 'biotechno l', 'planting', 'chaudhary', 'johansson', '1302', 'egawa', 'mh', 'wl', 'proto c', 'supt1', 'lumi', 'gearray', 'lrp5', 'campomelic', 'sox', 'sry', 'femora', 'kyphosis', 'tracheostomy', 'ventilation', 'hips', 'communicative', 'andhindi ii', 'pt7', 'ppumi', 'pmli', 'synthesizer', 'col2a1', 'bent', 'deoxyribose', 'tyr15', 'arg4', 'arg7', 'hydrophobically', 'c16', '00005', 'pakistani', 's udbeck', 'asp816val', 'omniscrypt', 'gdnas', '1031', 'histochemistry', 'cyclo sporine', 'complaining', 'msc', 'epicentre', '2724', '2607', 'horny', 'analge sics', 'btk', 'kantarjian', 'ctcae', 'lipase', 'hermine', 'diphtheria', 'val5 60gly', '17a', 'radicular', 'decrement', 'dept', 'gutkind', '8600', 'interfac ed', 'adoptively', 'notorious', 'untreatable', 'cux1', 'cytomorphology', 'mul tiparameter', '1614', 'rated', 'socalled', 'kasumi', 'd1246', 'd1246h', 'd124 6n', 'erasmusmc', 'dermatologist', '1205', 'restarting', 'emesis', 'palpitati ons', 'myalgia', 'fear', 'rat2', 'unveils', 'phophorylated', 'ship1', 'regula rized', 'icam', '103k', 'mkk4ala', 's257a', 't261a', 'sek1', 'whitmarsh', 'mapkk1', 'osmotic', 'shocks', 'atfi', 'ganiatsas', 'roulston', 'ferrigno', 'pr unier', 'rivat', 'hilberg', 'nishina', 'hemipterous', 'glise', 'hepatogenesi s', 'teng', 'beddington', 'ramirez', 'mmk4', 'draper', 'nakano', 'okabe', 'wo bus', 'mmkk4', 'snl76', 'subcultures', '293k', 'subculture', '63k', 'approx 2', 'p116', 'p46', 'thr183', 'obliterates', 'pp54', 'pp46', 'pp38', 'kieran', 'jnks', 'p38alpha', 'brunet', 'cardone', 'mcpherson', 'goldenberg', '46plusmi nus0', '58plusminus0', '08plusminus0', 'junaa', 'burdon', 'jirmanova', 'obliterate', 'duriez', 'apoptosis', 'schultz', 'waf', 'analysisisthe', 'ddb1', 'fails afe', 'powerpointfig', 'biomeda', 'j2', 'j3', 'proteininsthe', 'discussionthe', 'al55', '916', '886', 'paque', '1103', '1670', 'tth', '1503', '1090', 'repeti tions', 'bern', 'karyotyped', 'consolidated', 'auer', 'rods', 'infidelity',

'scl', 'gfi1', 'calreticulin', 'smmhc', 'ser21', 'airspace', 'pcsk6', 'gemtuzumab', 'pragmatic', 'antineoplastic', 'cyano', 'doors', 'a18d', 'colon2', 'target3', 'developed6', 'frequencies10', 'properties15', 'supraphysiologic', 'plk416', 'unstarved', 'manner19', 'stress20', 'egfrhigh', 'pretherapeutic', 'conditions16', 'polarisation', 'transition21', 'growth22', 'genotypes13', 'subtype23', 'krt6b', 'krt16', 'krt17', 'itgb4', 'lamb3', 'lamc2', 'anxa8', 'col17a1', 'collagens', 'col18a1', 'ctsc', 'plau', 'plaur', 'serpine2', 'outcome25', 'vitro26', 'g12d13', 'tumourgenesis27', 'homogenates', 'cxcr2', 'burden29', 'outcome30', 'inhibition31', 'plaur31', 'iip', 'dichlorophenyl', 'electrophile', 'binding21', 'brefeldin', 'microfluidizer', 'trays', 'lct', 'dimethylsulphoxide', 'uplc', 'nap', '3676', '2f2', '4970', '2c10', 'rapla', 'ads', 'pmei', 'monotransgenic', 'superovulated', 'icycler', 'exsanguination', 'inflated', 'vectamount', 'alcohols', 'iscript', 'gcagggttggccat', 'atgcgtcgccaatgtaaat', 'rasg12d', 'caaggacaagggtacagttatgtact', 'ggcatctgcctgctttg', 'cuboidal', 'septa', 'logs', '6022', 'defensive', '173c', 'undiagnosed', 'rab4', 'phosphinositol', 'pik3r3', 'leeds', 'ptag1', 'ptag2', 'ptag3', 'pt1g1', 'pt1g2', 'pt1g3', 'pt2g1', 'pt2g2', 'pt2g3', 'ptx', 'bftc905', 'jmsu1', 'luc4', 'lucc5', 'lucc6', 'lucc7', 'lucc8', 'blc1', '181523', 'r274a', 'e137k', 'r162', 'e218', 'r262t', 'k288q', 'n564d', 'cantley', 'immunoprecipitation', 'kidlington', 'carbayo', 'nhu', '1519gc', '1670gc', 'r557p', '652gt', '862ac', '1519c', '1670c', 'r358', 'q579', 'y580del', 'n377k', 'flvradas', 'r358a', 'r358x', 'n377', 'g376', 'p85nish2', 'e365', 'i566', 'd578del', 'n441i', 't654i', 'flvres', 'k142fs', 'i177n', 'e217k', 'n285h', 'e297k', '1550', 'm176', 'v181', 'immunoprecipitation', 'w237', 'y242del', 'r262', 'k288', 'interhelix', 'k48', 'minella', 'exhausted', 'auxin', 'bolstered', 'deepened', 'xrcc4', '9165', 'alcian', 'pipc', 'unexcised', 'flinn', 'kamura', 't58a', '099', '163v', 'g505s', 'bre', 'itch', 'relb', '9297', 'mercaptopurine', 'pcr1i', 'ubiquitin', '0147', 'nav2', 'tcf7l1', 'radiotherapeutic', 'see3', 'crc3', 'analyses4', 'mutsig11', 'signaling12', 'niche13', 'slc9a9', 'differently17', 'two17', 'described18', 'algorithm19', 'q6', '10p25', 'vti1a', 'translocation4', 'usp12', 'trastuzumab20', 'fate23', 'crc23', 'crc26', 'pathogenesis28', 'method29', 'tcf7l14', 'ttc28', 'growth30', 'dkk', 'wilms', 'tumor31', 'fzd10', 'database33', 'platform35', 'scn5a36', 'apol637', 'hindgut', 'transcription39', 'analyses6', 'promise40', 'translocations43', 'humanmethylation27', 'flexibilities', 'thermus', 'aquaticus', 'sideways', 'nbds', 'nbd', '1108', 'authorizes', 'nma', '1e3m', '2o8b', 'cmap', 'vibrant', 'c33a2', 'ksik', 'hjk', 'kth', 'hik', 'jth', 'rmsfxray', 'rigidification', 'levers', '1281', 'meaningless', 'motional', 'uncertainties', 'reordered', 'wagging', 'basepairing', 'flap', 'phe596', 'phe650', 'phe1108', 'closes', 'reopening', 'rebind', 'mutations20', 'vasculitis', 'function20', 'mir30', 'h4230', 'described48', 'domain20', 'piccart', 'gebhart', 'romond', 'monsey', 'larson', 'esteva', 'nahta', 'scaltriti', '9e11', 'dihua', '1m14', '3bbt', 'arnold', '2jiv', 'sbkr3', 'atpbased', 'toxiclight', 'electrochemoluminescence', 'boulbes', 'dsu', 'fluoview', 'olympusfluoview', 'fticr', 'hydroxypropylmethylcellulose', 'e804d', 'n842i', 'thirteenth', 'i654v', 'breyer', '1792f', 'g857e', 'v794m', 'g785s', 'd808n', 'r838q', 'm887i', 'disfavoring', 'methylsulfonylaminol', 'breach', 'grape', 'undephosphorylation', 'rexer', 'conserves', 'lehrer', 'mahlknecht', 'rafidi', 'crop', 'pt66', 'ouyang', 'underphosphorylated', 'hsp90a', 'partnership', 'pacey', 'babchia', 'koboldt', 'santarpia', 'herter', 'sprie', 'tikhomirov', 'affordable', 'n642h', 'sichuan', '610041', 'elizabeth', '4division', 'toxicological', '7department', 'groupe', '14department', '15department', 'clinique', 'naturecommunications', 'wholeexome', 'activation5', 'a573v', 'snt', 'celiac', 'study11', 'b72', 'b39', 'il2r', 'wtand', 'scatchard', 'w385c', 'a496s', 'f866y', 'f824w', 'r105s', 'n456s', 'g774', '1239r', 'i638f', 'h2286', 't654m', 'shrna2', 'k608e', 'cotreated', 's768r', '3020', 't685s', 't692n', 'spondylometaepiphyseal', 'haukeland', 'bergen', 'd125y', 'pwzblast', 'trcn000121117', 'cccatgcctatgccactccat', 'atggagtggcataggcatggg', 'tttttg', 'trcn00012111

21', 'ccctggagggttccatcattta', 'taaatgatggaacctccaggg', 'pwzl', 'lanthasree n', 'y417', '4490', 'ml3', 'ddrs', 'npxy', 'shca', '1k3a', 'mineralization', 'ras1', 'par3', 'par6', 'rab5a', 'collagenases', 'giannini', 'mk410', '236 8', 'interspecific', 'h1421y', 's1512i', 'ndividuals', 'domain6', 'znf2f', '5 9gcggaattcatgg', 'atttatctgctctt', 'bral', '59atagtcacttccagccca', 'tctgttat gt', '59cggaatt', 'ccagagggataccatgcaa', 'lx21dm', '59gcggtcga', 'catctgtggc atgtt', '1455', 'ux14', '59cggaattcactcacagaaaagttagt', '59gc', 'ggatcctcagta gttggctgtggggat', '59cggaattccagagggataccatgcaa', 'ux12', '59cggaattcgg', 'tg aagcagcatctggg', '11b24', 'ggaattccagaggg', 'ataccatcaatataacc', 'ctgatcaagc cccagca', 'cctgctgggcttgcac', 'g1706af', 'gacactgaaatatttctagcaattgccggc g', 'gaaaatgg', 'g1706ar', 'ccatttccggcgaattt', 'ctagaaaatatttcagtg', 'a 1752pf', 'ccaa', 'ggtccaaagcgacccgagaatcccaggac', 'a752pr', 'gtcctgggattctcg aggtcgcttggaccttgg', 'dpn1', 'gttagaacag', 'tatgggagccagccttct', '59agaaggct ggctcccatactgttcta', 'ggtacatgcaca', 'tatgctctgg', 'ccagagcatatgtgcatgtac c', 'gcagaatctgcccagagttcagctg', 'cagcagctgaac', 'tctggcagattctgc', 'gccg gctataatgcaacggaaagaaagtgtga', 'gcagg', 'm128tr9', 'cctgctc', 'acactttctccgtt gcattatagccggc', 'm1628vf', 'gccggctataatgcagtgaaag', 'aaagtgtgaggcagg', 'm162 8vr9', 'cctgctcacacttctccactgcattatag', 'ccggc', 'gtttgtgtgtgaacgggactgaaa tat', 'atatttcagtgcccg', 'ttcacacacaaac', 'gctatttctgggtggccaggctattaa', 'tt aatagactggccaccc', 'agaaatagc', 'gctgtgtgcttctgtgg', 'ccacagaagcaacacacag', 'ccaattgtttgtcag', 'ggctgcacaa', 'acacaattgg', 'ggtgaccggagagcgggtgtt', 'c caacacccgctctggg', 'pj103', 'microarray30', 'gal4dbd', 'apic', 'gal1p', 'de ng44', 'amie', 's1512', 'comparisons23', 'bilat', 'f92', 't334', 'brca65', 'm 792', 'othca', 'homcs', 'disgerminoma', 'ovcs', 'txn', 'iv1', 'pdg', 'co11', 'c4323t', 'hdynk', 't4339c', 'c4380t', 'g4654t', 'sgca', 'cases34', 'a4956 g', 'controls38', 'a5001g', 'mvsr', 't5002c', 'c5173t', 'g5236c', 'c5242a', 'mutation39', 'a5277g', 'tivs', 'g5373c', 't5443g', 'g5482t', 'history41', 'g5544t', 'vil', 't5628c', 'disease42', 'c5677a', 'ay077732', 'ay428536', 'm egalign', '1interspecific', '11number', 'paircoil', '1alpha', 'aggt', '34t', 'ivs1', 'cggt', 'luh', 'donation', 'moskaluk', 'c149', 'bstni', 'gruis', 'dr hodamine', 'hollenberg', 'pds', 'molten', 'blunting', 'haplodeficiency', 'red undancies', 'isobutyl', 'qscript', 'c1000', 'cfx', 'dmirb', 'endogeneous', 't riglyceride', 'aib', 'pki', 'immunofluorescence', 'grp1', 'improvision', 'open lab', 'ch3oh', 'waste', 'nh4oh', 'curtailed', 'scrutinize', 'endorsement', 't aki', 'mtg8', 'hner', '1072', '1283', 'beghini', 'worobec', 'pignon', '1104', 'nongastric', 'karyotypically', 'val617phe', 'methodspatients', 'allo', '132 5', 'biomed', '1928', 'igvh', 'estimations', 'resultsdetection', 'ozogamici n', '2125', 'item', 'fare', 'eecs', 'neecs', 't576', '3tag', 'pik3r1r348', 'introductionspecific', 'resultspik3r1r348', 'pik3r1l370fs', 'inhibitorshaving', 'counterscreen', 'ic25', 'pik3r1e160', 'rapamycin', 'informer', 'bi78d3', 'aeg3482', 'dkrmns560', 'quayle', 'pik3r1wt', 'optionsto', 'ovk18', 'linesmek', 'as601245', 'sb203580ecc1', 'nahec108', 'nahec265', 'nasngii', 'naovk18', 'naic25', 'micromoles', 'optionspik3r1r348', 'activationwe', 'l370fs', 'skut 2', 'jnkmapks', 'kinasesmapk', 'rijard', 'mkk2', 'mlk3', 'subcellular', 'mlk 2', 'dlk', 'mlk', 'nucleusnext', 'pik3r1n324', 'n299', 'n324', 'e212', 'q23 5', 'e266', 'jnkp85', 'tolias', 'manneville', 'optionsalong', 't17n', 'signal ingbecause', 'cytosolto', 'vivoto', 'gdc0973', 'cuevas', 'bachelot', 'hellye r', 'pandey', 'prohibit', 'lanning', 'buongiorno', 'lagana', 'chadee', 'kyria kis', 'stankiewicz', 'jip1', 'dickens', 'jip', 'inder', 'dhanasekaran', 'comu tations', 'untoward', 'procedurescell', 'reagentsendometrial', 'gulf', 'coast', 'assaythe', 'prestoblue', 'blottingcytosol', 'fractionprep', 'immunohisto chemistrythe', 'mdacc', 'broaddus', 'miceall', 'analysisall', 'tirkkonen', 'visakorpi', 'lx', 'gnas1', 'postfixed', '11q14', 'p31', '12q21', 'recombinatio ns', 'krebs', 'subregions', 'vac', 'mately', 'pst', 'pbr322', 'antisheep', 'cerbb', 'histomorphological', 'tlrs', 'pneumocystis', 'pansorbin', 'aliq uotted', 'pyogenic', 'candidiasis', 'concanavalin', 'tlr2', 'apcs', 'resynthe sis', 'dissociating', 'posttransplant', 'mcdonald', 'serology', 'ssp', 'a11',

'cw4', 'serologically', 'b15', 'abi3130x1', 'officially', 'dqb1', 'sbt', 'ipd', 'v6', 'serological', 'deoxyribonuclease', '2711', '1046', '724', 'w7', '1655', '19n', '11n', 'idls', 'boland', 'tobe', 'acharya', 'palombo', 'encountering', 'bindingto', 'thecorresponding', 'asdescribed', 'veri', 'theimmunoprecipitate', 'studythe', 'wereincubated', 'thepresence', 'tothe', 'resultsexpression', 'materialsand', 'invitro', '5721', 'para', 'mutationsand', 'cancerassociated', 'isassociated', 'arenot', 'consistentwith', 'andthat', 'associatedwith', 'e011', '345541', 'easysep', 'il5', 'jacques', 'ser536', 'surefire', 'vacuolization', 'cc12', 'lentivirally', 'sulfonic', 'ultralink', 'electroporating', 'egr', '31st', 'folfiri', 'nonstop', 'immunofluorescence', 'antioxidative', 'h2dcfd', 'tp53inp1', 'r189x', 'vault', 'mvp', 'missense', 'alluded', 'unreasonable', 'supercompetent', '6h2', 'genosys', 'wullschleger', 'proud', 'sin1', 'guertin', 'molino', 'robb', 'urano', 'ohne', 'm135t', 'mtors2215y', 'mtorr2505p', 'mtorc', 'reiling', 'toschi', 'edinger', 'r8131i', 'f218i', 'kiyokawa', 'areddy', 'ducommun', 'cdc25b', 'kristjansdottir', 'rudolph', 'koldner', 'wijnen', 'parc', 'vos', 'pyatt', 'superscripttm', '1896', 'basepairs', 'somewhere', 'carcinomatous', 'theories', 'punches', 'misclassification', 'fat2', 'pdl', 'ivs3', 'e705k', 'claims', 'xrcc2', 'rbbp8', 'kms', 'exon1', 'q4', 'entertained', '1344', 'al26', 'motokura', 'buckley', 'pej6', 'sarcevic', '1891', 'armed', 'uncle', 'c81', 'poisoning', 'chiefly', 'comparability', 'mammographic', 'appropriateness', 'abnormalities1', '26k', 'byrd', 'rahman', 'bakkenist', '6t', 'maillet', 'spurdle', 'tamimi', 'multiplicatively', 'proteosomally', 'vitroubiquitylation', 'proteasomally', 'buys', 'napco', '7001', 'jouan', 'radioisotopic', 'carbobenzoxyl', 'leucinyl', 'norvalinal', 'aggagcacacatgggcttctg', 'gcagaattcggttcacaagctagctaatctggatcatgttatccatccatccatccatgtgtggc', 'theecori', 'havhl', 'haeii', 'andbgl', 'cggtatcccaggcatctc', 'cgtcaacattgagagatgg', 'vhsequence', 'pgal', 'epas', 'irp', 'maneuvers', 'hankinson', 'liebold', 'ig32', 'pab419', '190b', 'ca630', 'acylamide', 'affiniti', 'aminopurine', 'weight35s', 'isopeptidase', 'ubial', 'meub', 'unprejudiced', 'ascrivable', 'sv40t', 'nonspecifically', 'seepart', 'comparable', 'r82p', 'p86h', 'n90i', 'q96p', 'subsequences', 'vitroassay', 'p86n', 'vitrobinding', 'thein', 'arntta', '682', 'vitrowith', 'sited', 'unconnected', 'cobaltous', 'overactivity', 'mullighan', 'kuiper', 'pam', 'tibshirani', 'conserting', 'chapiro', 'episomal', '15x', 'aall0232', 'p9906', 'downing', 'prkdcscid', 'xv', 'xvi', 'p185', 'anesthesia', 'pallor', 'ntc', 'arg93', 'gly106', 'lys111', 'glu453', 'r93w', 'g106r', 'g106v', 'k111e', 'delp449', '1455', 'hormonally', 'biosample', 'oncomatrix', 'pfastbac', 'yardena', 'r93q', 'e453a', 'e365k', 'g364r', '15hr', '10min', '000xg', 'r108h', 'adb', 'r88', 'r93', 'g106', 'k111', 'e453', 'e81k', 'dele109', 'g118d', 'v344a', 'delli5', 'v22inshpi', 'a222v', 'l445', 'f667l', 't1025n', 'm1040l', 'm1004f', 'aktthr308', 'dutt', 'gymnopoulous', 'glu454', 'glu348', 'v344r', 'music', 'tbx3', 'myh9', 'map4k4', 'r166q', 'lymphoma27', 'arcs', 'vntr', 'genego', 'pathscan', 'shc1', 'kdm4a', 'kdm5b', 'letrozole', 'anastrozole', 'exemestane', 'remark', 'pindel', 'lumina', 'uscs', 'lumens', 'ultrastructural', 'kanai', 'bländ', 'sidhar', 'subjective', 'copied', 'rtpcr', 'cancer14', 'overwhelm', 'tumours16', 'alexafluor', 'ultravision', 'ecodry', 'described7', '20mm', 'h36p', 'p44a', 'degraded', 's41', 's33c', 's37c', 's37f', 'i35s', '20071', '30600', '24138', '14865', '21217', '6700', '29923', '28356', 'af009674', 'pbatem', 'bcata', 'hiii', 'pt81', 'hin', '2146', 'proofed', 'fao', 'audard', 'trustable', 'colnot', 'benhamouche', 'cki1', 'alanin', 'gsser', 'gaunitz', 'igfs', 'igfbp', 'bingen', 'pbsk', 'skirball', 'rybp', 'h2ak119', 'ubh2a', 'hgu133', 'bcorl1', 'cebpe', 'monoubiquitylation', 'pug', 'relapse3', 'spi1', 'malignancy5', 'chromosome6', 'apoptosis1', 'origin3', 'ikzf13', 'all3', 'haematopoiesis4', 'tumours13', 'myelodysplasia18', 'sequenced1', 'c1408y', 'p9251', 'domain22', 'nrid', 'taz1', 'q1500p', 'rts13', 'q1500', 'r1446h', 't1070fs', 's1761', 'i2101', 'reduplication', 'crebbp25', 's1887', 'i2101m', 'ncbd', 'dsrna', 'outcome26', 'acetylation28', 'cells29', 'h13k18', 'mefs23', 'analysis

3', 'at1', 'ctv', 'tlx3', 'mllt7', 'trad', 'kopt', 'ttg2', 'molt3', 'mkb', 'mutz5', 'op1', 'lmo1', 'ttg1', 'sd1', 'skw3', 'tcl1', 'tom', 'picalm', 'uocb1', 'af6', 'skno', 'olaf', 'heidenreich', 'dario', 'campana', 'recommendations31', 'op132', 'replig', 'cogenics', 'dbsnp33', 'phred34', 'snpdetector36', 'indeldetector37', 'pipeline38', 'consed40', 'extracts41', 'a22', 'coa22', '2210', 'pdb3biy', 'pymol43', 'previously25', 'ab1191', '2574', 'crebbpflox', 'ep300flox', 'schaumburg', 'mkb1', 'endocardial', 'tie2', 'atrioventricular', 'hearts', 'valves', 'infused', 'pyrazolo', 'tetramethylbenzidine', 'embryological', 'multisystemic', '294002', 'nakalai', 'tesque', 'dpc', 'immunosuppressant', 'worthington', 'oocytes', 'halfway', 'parasternal', 'diastolic', 'systolic', 'systole', 'rapa', 'decapentaplegic', 'wap', 'abscusses', 'transdifferentiation', 'inhibin', 'keratinized', 'ru486', 'intrastrand', 'lodish', 'bsteii', 'xe', 'ar3', '0033', 'micrograph', 'villi', 'microadenomas', 'desmoplasia', 'bmpr', 'smits', 'mcn', 'interrelationship', 'bardeesy', 'takaku', 'bierie', 'moses', 'subramanian', 'kawaguchi', 'brevity', 'bmp4', 'ipmn', 'muc2', 'hruban', 'finetek', 'carraway', 'denuded', 'pearl', 'psmad1', 'cancerization', 'tamura', 'tortuous', 'grady', 'myeroff', 'oncor', 'ascnar', 'ame liorating', 'smarca5', 'epicardial', '70kda', 'cin85', 'lowry', 'ods', 'qstar', 'carbamidomethylation', 'keratins', 'rasal2', 'nonhybridized', 'tied', 'foxl2', 'r130', 'fumihiko', '3815', '3814', 'pcmv3', 'hawkins', 'babraham', 'k227e', 'nonepithelial', 'tag2a', 'py2', 'procaspases', 'nonfixed', '10417', 'ag1478', 'gw583340', 'regenerated', '1658p', 'winmdi', 'k708', '1206', 'varian', 'unconditional', 'mcap', 'polydactyly', 'ectopia', 'nosology', 'shunt', 'chiari', 'cals', 'midbrain', 'svt', 'tachycardia', 'hypermobility', '6m', '8w', '7m', 'valproic', 'gait', 'transfers', 'expressive', 'febrile', 'vpa', 'waves', 'speaks', '2922', 'overweight', 'underreported', 'comfort', 'parasagittal', '5q35', 'fissure', 'handicap', 'gerd', 'postzygotic', 'development16', 'seattleseq', '75mm', 'plaines', 'dyskerin', 'dkc1', 'h446', 'h209', 'haddad', 'mdv', 'prooncogenic', 'c228t', 'c250t', 'mccs', 'ok', 'leukemogenicity', 'al115', '1906', 'hpa', 'ref52', 'comigrating', 'i87syn', 'h58n', 'k83n', 'r80c', 'r177ter', 'intensified', 'supershifting', 'r177', 'progressions', 'hypernormal', 'reservation', 'varieties', '601399', 'quinacrine', 'k83e', 'd21s1413', 'd21s65', 'pebp2beta', 'langabeer', 'roumier', 'vegesna', 'taketani', '235k', 'r139', 'r174', 'rodenhiser', 'betap1', 'betab', 'betae', 'betaf', 'beta p2', 'tahirov', 'mundlos', 'mappings', '307k', 'iergeau', 'higuchi', 'cities', 'victims', 'soldiers', 'residents', 'aboveground', 'weapon', 'nevada', 'bikini', 'atoll', '1940s', '1950s', 'textbooks', 'liked', 'childbirth', 'kostman', 'l29s', 's21syn', 'g42r', 'transplantaion', 'buijs', 'acquirement', 'backcrossing', 'bxh2', 'spi', 'nk2', 'domain14', 'binding15', 'lineages25', 'lichen', 'c618', 'c620', 'e805k', 'activity49', 'oestrogens', 'tumours151', 'patients7', 'somatotrophs', 'teens', 'sword', 'cancers80', 'neuropathic', 'neuroprotective', 'amyotrophic', 'neuroprotection', 'forwards', 'refused', 'accredited', 'bolus', 'epinephrine', 'norepinephrine', 'microinvasive', 'furthest', 'acgt', 'glutamines', 'practitioners', 'gallen', 'daejeon', 'tabulation', '0039', 'tumorspecific', 'lovejoy', 'al24', 'guarding', '100a', 'conductance', 'invoking', 'synthesised', 'rluc', 'dax1', 'teaa', 'soc', 'fallen', 'cumbersome', 'colinear', 'powers', 'jpages', '7794', 'suppmat', 'gang', '536a4g', 'dele', 'commence', 't172', 'a173', 'n7', 'ht1080', 'inbred', 'nontranslocate d', 'sedivy', 'ladders', 'intuitively', 'mycs', 'friable', 'ichthyosis', 'supraorbital', 'angulated', '000011', '000012', '178gright', 'arrowc', 'cfc73', '458aright', 'cfc8', 'cfc91', 'asp154', 'evolutionally', 'proteins9', 'dots13', 'apoptosis8', '56k', 'cancers9', 'g469e', '1406gright', 'cfc76', 'cfc81', 'cfc94', 'cfc114', 'n581d', '1741aright', 'arrowg', 'cfc95', 'cfc110', '1485f', '1455gright', 'cfc83', 'k499e', '1495aright', 'cfc79', 'e501k', '1501grig ht', 'cfc77', 'e501g', '1502aright', 'cfc90', 'cfc105', 'a246p', '736gright', '770aright', 'cfc96', 'cfc103', '00433', 'phrlnull', 'rla', '1485s', 'condition2', 'nevi9', 'cancer9', 'activity11', 'activity9', 'syndrome15', 'ganglione

uroblastoma', 'carcinoma8', 'leukemia15', 'c114', 'hmuts', 'ighv', '37k', 'bl
ok', 'abel', 's2n', 'ps6k', 'peripherally', 'rafi', '88x', 'unifiedgenotype
r', 'patientswith', 'conspire', 'hyperthermic', 'iphc', 'endure', 'creighto
n', '131550', 'ensg00000146648', 'hrgfp', 'w731l', 'e734q', 't785a', 'c797y',
'y801h', 'l831h', 'e868g', 'lxt', 'inflow', 'r2a', 'r2b', 'r2c', 'resectabili
ty', 'w731', 'e734', 't785', 'y801', 'r831', 'e868', 'figure33', 'figure9a9
a', 'hyperthermia', 'figure9b9b', 'scanarray', 'nmi', 'ssca', '97k', '1832g',
'provisions', '0196', '0103', 'speeding', 'uch', 'autoubiquitinated', 'gfps',
'dgfp', 'anita', 'pnls1', 'pnls2', 'methylcoumarin', 'qiaexpressionist', 'ami
nco', 'slidebook', 'virapowertm', 'sastry', 'tattooing', 'grubb', 'studentize
d', 'aseptically', 'facscaliburtm', 'externalization', 'a95d', 'g178v', 'c9
1', 'c91a', 'krkkfk', 'rrkrssr', '8bp', 'c19a', '6844', 'herpesvirus', 'haus
p', 'mallery', 'a161t', '1237', 'cpt', 'preanalytix', 'w303', 'kvdmr1', 'conc
eived', 'hyaluronidase', 'resumption', 'igf2r', 'magi3', 'binding13', 'cancer
6', 'leukemia21', 'tumours10', 'qualification', 'eminence', 'periventricula
r', 'krieg', 'c342y', 'k656e', 'moosa', 'meninges', 'tuj1', 'hebert', 'dience
phalic', 'anthony', 'fgfr2c342y', 'alterations2', 'progression3', 'morbiditie
s4', 'pedbrain', 'nf15', 'event6', 'disease2', 'astrocytomas9', 'medulloblast
oma10', 'fusion16', 'brafins599t', 'alteration9', 'rnf130', 'clcn6', 'gnai1',
'pa65', 'insvrl', 'homodimerization18', 'stimulation19', 'brafinsvrl', 'pa11
7', 'pa142', 'r73m', 'cancer20', 'others22', 'cancer24', 'tumors25', 'nacc2',
'activation27', 'k656', 'breast28', 'lung29', 'genes33', 'database34', 'embry
os35', 'potential36', 'vitro37', 'tissues40', 'fgfr1wt', 'pa89', '5kb', 'leuk
emia41', 'monomyelocytic', 'activation42', 'pa84', 'pa166', 'patients46', 'mu
tations50', 'mice51', 'pa69', 'gbm5', 'gliomas5', 'membrane54', 'gbm39', 'ren
ewal55', 'development56', 'loss57', 'astrocytomas61', 'trials62', '2327', 'ap
ovldlii', 'd538g', 'y537s', '1469v', 'v422del', 'y537d', 'd5358g', 'e380q',
's463p', 'azd9496', 'mbc', 'y537', 'mbcs', 'aporeceptor', 'degraders', 'ser
d', '00897702', 'y537c', 'y537n', '1536h', 'serine118', '1536r', 'src3', 'rst
er', 's432l', 'v534e', 's118', 's167', 'nrd', 's436p', 'whatssoever', 'presatu
rated', 'k529', 'serds', 'ic99', '58668', '0810', 'stc2', 'pik3can345k', 'e17
76', 's705fs', 'antiestrogens', 'yy1', 'crossroad', 'tropism', 'ndicate', 'r5
020', 'scap', 'tccttcgtggccttc', 'cctacctgaccagtcgtccctg', 'tccttgccgggt
ccccaaactgcacc', 'gctctcgctctgctgc', 'gctctctgcctcgccggag', 'cctcgccggagcg
tgccgactctgct', 'ctgccagccccctcac', 'acgctcacgaattcagtcggagg', 'cagcgcattc
taaaggtaaggacg', 'cgccctcaggatggctt', 'tttggctctgcctcgctgctc', 'ctcgctgc
tcccggcgttggcgcc', 'metzger', 'marburg', 'pxp20', 'nordeen', '4s46', '4s48',
'pt81luc', 'pmluc', 'kmno4', 'reip', 'ultrasonics', 'a350g', 'reimmunoprecipi
tation', 'reimmunoprecipitated', 'gctgaaaactaattgatctggag', 'ccattgttaagcccta
agtcc', 'aacaaaaccaatttagaacacctt', 'atttccttcattgtcctct', 'tatgaaaacccggactac
agg', 'ctgttgttaagcaaagatcaaag', 'ccaggcctacaatttcattat', 'agggatctgagattcaga
aag', 'ctcacaggacaaggatgcgggttgca', 'tggcggagattagcagagaggatct', '1603',
'abbondanza', 'r6775', 'ach4', 'lys5', 'lys8', 'c100', 'bromocinnamyl', 'iso
quinolinesulfonamide', '2hcl', '36b4', 'junb', 'nonmammary', 'herber', 'erg
e', 'antireceptor', '3d3d', 'justifying', '4e4e', 'metivier', '5d5d', 'detach
es', 'atgcaaat', 'reenter', 'trophoblasts', 'keratinoytes', 'antioncogene',
'chn', 'alky1604', 'stat3y705', '2t202', 's6pt240', 'baf73611', 'bluesky',
'worcester', 'lir3', 'tgi', 'c1156', 's1206r', 'd1268', 'v1180', 'r1181', 'e
1210', 'e1241', 'd1203', 'f1241', 'l11', 'columbus', '3figure', 'wtegfr', 'be
tacellulin', 'trf', '115k', 'ptprf', 'ptprg', 'ptpn3', 'ptpn13', 'dsps', 'lmp
s', '1375', 'nrptp', 'fn3', 'postsynaptic', 'occludans', 'ptpc', 'predecesso
r', 'nonselected', 'r1212w', 'r1346l', 't1368m', 'difluoro', 'methylumbellife
ryl', 'difmup', 'nickelaffinity', 'menton', 'r632x', 'valosin', 'vcp', 'p97',
'phosphatome', 'phophatas', 'e1129', 'negligibly', 'tryptophans', '1124',
'chromophores', 'nificantly', 'comets', 'ume', 'tartu', 'estonia', 'bryostat
in', 'pacc', 'overarching', 'chr7', 'chr16', 'sor', 'ns1', 'mutations7', 'tar
gets11', '1875', 'gun', 'nonclinical', 'motivates', 'meti', '7q32', 'bliss',

'additivity', 'drc', 'centering', '9433', '3127', '4051', 'compliant', 'velvet', 'gse22208', 'bisected', 'tudor', '3522', 'codeletions', 'karls', 'pilomyxoid', 'cleanseq', 'diagnostically', 'neuenheimer', 'feld', 'lately', 'oliva', 'akbari', 'hist1h3b', 'pangenomics', 'centroids', 'molybdate', 'informatique', '55c', 'kitor', 'gistdevelop', '31patients', 'gistwho', 'were13', 'whowere', 'nonresistant', 'withprimary', 'and15', 'afterinitial', 'of15', 'whomhad', 'mutationinkitexon11', 'gistcommonly', 'kitv558d', 'mrsg', 'cce', '106841', 'ca102774', 'dk55748', 'besmer', 'po1', '47179', 'ca94503', 'ca102613', '3203', 'antonesc', 'f2005', 'z90', '70jc', 'gtctgatccactgaagctg', '50jc', 'accccatgaactgcctgtc', '51jc', '54jc', '55jc', 'wiz', 'isothiocynate', 'gccgatccttcgtatacgtagct', 'ggcgaattctgtatacacagttgaaaatg', 'chloramphenicolresistant', 'penzburg', 'snapfrozen', 'tyr99', '20jc', 'pepsinhcl', 'fixated', '68jc', 'airdried', '4183', '4184', '122', '11vyidptql569', '576del', '2ac', '11p551hmy552', '554del', '120', 'ins502ay', '3bc', '11vqwkv555', '11k558n', 'ins559q', '11wk557', '19', '11ngnnyyvidptql563', 'kitamplification', '8del', '4185', '14bc', '11kvvee558', '562del', '1of', '111576p', '17c', '127', '113', '11v559d', 'ins559p', '114', '24a', '11v560del', '24b', '11ins574tqlpyd', '11ngnnyyvidptqlpy564', '578del', 'primr', '11py577', '31c', 'ctested', 'individual', '4186', 'exon11mutations', 'kitimatinib', '4187', '4188', 'gistc', 'monthsbt', '11of', 'ccurrent', 'bchronic', 'perioperative', 'extirpation', 'avert', 'rfg', 'tk3', 'hybaid', 'omnigene', '738', 'glycyl', 'histidyl', 'phenylmethyl', 'pn1a', 'polygonal', 'bioview', 'function22', 'suicide', 'researched', 'cripples', 'undertakes', 'motives', 'rasgef', 'motivation', 'wtp53', 'minn', 'cxcl2', 'metallopeptidase', 'shcon', 'p21waf', 'brij', '1188', 'methacarn', 'lauren', 'hypodermic', 'micromanipulator', 'radioisotope', 'claus', 'struewing', 'roa', '10u', 'niece', 'skolnick', 'shattuck', 'eidens', 'gwas', 'v12061', 'v1238i', 'v1110i', 'h11121', 'h1124d', 'y1248d', 'embodied', 'hos', 'k252a', 'idaho', '1268', 'hr10', 'stoker', 'bellusci', 'c3223t', '3223', 'inoculum', 'underhill', 'rainin', 'oefner', 'hongyo', 'pmb11', 'pmb1', 'sp260', 'metastasizes', 'byproducts', 'p56', 'woodland', 't185', 'diffuses', 'ix81', 'juxtacrine', 'sulfhydryl', 'ladc', 'necropsied', 'accidentally', 'ladcs', '289k', 'bellon', 'phosphosites', 'jackman', 'm1268', 'jagadeeswaran', 'am7', 'mf1', 'unrecombined', 'cyclind1', '1700t', '95k', 'translesion', 'monoubiquitylated', 'capan', 'f11', 'f12', 'xrcc3', 'takata', 'k6', 'fanco', 'fancp', 'rad52', 'icelandic', '999del5', 'mice16', 'lessened', 'polygenic', 'heptamer', 'nonamer', 'epha6', 'structure10', 'e317k', 'r544w', 'e563k', 'e452k', 'e836k', 'e872k', 'k751m', 'y1162', 'y1284', 'erbb4s', '6ap', '31t', '63t', 'vect', 'errb4', 'erbb411', 'important20', '1126', 'dip', '005235', '8327667', 'pcdf', 'mcs2', 'lipfectamine', 'pfiv', '34n', 'todd', '60hr', 'walkersville', '650nm', 'softmax', '24hr', 'hema3', 'trypsinizing', 'myelodyplastic', 'homozygotic', 'methodscomplementary', 'crnas', 'gtaaaacgcaggccagtcctgagcaaggagcgggt', 'cagggaaacagctatgaccaggattttcttataatgt', 'gtaaaacgcaggccatgtaaactgaaatacgcacgttg', 'cagggaaacagctatgaccagttaaagtctcttagcg', 'gtaaaacgcaggccatgtatgctctaaaagaata', 'cagggaaacagctatgaccacatctattctgtgctgaag', '1exon', 'dnaprotein', 'rcm', 'discussiongenomic', 'pvs', '2mutations', 'cancerview', '3polymorphisms', 'ptpn11mutations', 'e69v', 'r289g', 'asp61', 'gly503', 'e76v', 'v451', 'hcc1171', 'n58s', 'val45', 'asn58', 'r138q', '2h10', 'detoxication', 'pf02341066', 'l1198p', 'v3a', 'rogier', 'versteeg', 'racemic', 'anderk2', '1282', 'alkkin', 'ggcatcatgattgttacgg', 'tctttttgggggttctctg', 'blipi', '2xp2', 'schroedinger', 'alkl1196m', 'h3122em14', 'alkg1269s', 'g1123s', 'g1123d', 'y1278h', 'crossresistance', 'e1197', 'k1267', 'p1198', 'sulfonated', 'gly1123', 'his1124', 'immunobiology', 'bldg', '21702', '5946', '1110', 'v157i', 'shu', 'h1112y', 'gxgxxgxx', 'y1248h', 'eble', 'borden', 'pmex', '2346', 'weterman', 'm1149', '11213', '5468', '5434', 'l1213v', '5161', '5976', '5456', 'bigner', 'mcclellan', 'kimberly', 'suzanne', 'livolsi', 'government', 'publisher', 'multiadaptor', 'newark', 'bertrand', 'vinylidene', 'mgb', 'trisomic', 'smallcell', 'price', '40c', 'carboxyterminal', 'preincubating',

'nickoloff', 'r1811', 'johannes', 'gutenberg', 'seromed', 'msv', 's602', 'ze
bisch', 'mapkkk', 't1799a', 'melanosomes', 'attrition', 'dimri', 'stressful',
'bauer', 'knauf', 'exploitable', 'extraosseous', '291x', 'bbrc', 'rosette',
'exon9', 'taf15', 'tafi68', 'prkc', 'extraskelatal', 'homeodomains', 'oct
3', 'patz1', 'nr4a3', 'atf1', 'angiomatoid', 'hyalinizing', 'hccc', 'endobron
chial', 'creb3l1', '16p11', 'ewsr', 'anteriorly', 'mapview', 'eagi', '314k',
'1052', '48k', 'infectants', '64k', 'mobilizes', 'jeon', '1588', 'ecorl', 'i
nsofar', 'e1af', 'censoring', 'glycan', 'cbfa2t1', 'preleukaemic', 'syndrome1
7', 'fdp', 'preleukaemia', 'designate', 'enforces', '19k', '142k', 'megakaroc
ytes', '77k', 'binding35', 'thpo', 'leukaemogenic', 'e1317q', 'thc', 'scripp
s', 'shear', 'conveniently', 'sensogram', 'folfox', 'administreated', 'mario
n', '1263', 'transformability', 'enthusiasm', 'pry', 'passes', 'xuewu', 'biop
hysics', 'cleavable', 'ptriex', 'enzcheck', 'mesg', 'mmt', '4obe', 'pdb2pqr',
'propka', '4g0n', 'dismissed', 'warhead', 'progrowth', 'sml', '161', 'crystal
ize', '4epr', 'pried', 'podbids', '1n6l', '1n6k', 'n61', '4l9w', 'difi', 'crc
2', 'intermixed', 'granda', 'aiei', 'finds', '633c', '1106', 'oligodendrogli
a', 'unimpeded', 'fir', 'wz', 'aee', 't798m', 'v773', 'n857', 't862a', 'h87
8', 'l755p', 'nmumg', 'v773a', '2ity', '2ito', '2itz', '2itt', '2tip', '2it
p', 'aee778', 'helga', 'bernhard', 'tyr1248', '2jiu', 'extrasalivary', 'scler
osing', 'behaviourally', 'mls', 'emc', 'afh', 'ccsltgt', 'ppms', 'hidradenom
a', 's100pr', 'zsg', 'tcf12', 'cuff', 'chondromyxoid', 'fviiirag', 'hyalinisi
ng', 'eswr1', 'znf444', 'creb3l2', 'calretinin', 'paravertebral', 'thoracopul
monary', 'askin', 'viscera', 'rosettes', 'adamantinoma', 'nse', 'lipoblasti
c', 'anastomosing', 'haemosiderin', 'fibrovascular', 'muscularis', 'propria',
'submucosa', 'patternless', 'myoid', 'paratesticular', 'featureless', 'irregu
larly', 'nesting', 'dscrt', 'multivacuolated', 'lipoblasts', 'chondrocytic',
'lobulated', 'hypovascular', 'haemorrhagic', 'ultrastructure', 'nr43', 'ss1
8', 'ssx2', 'fascia', 'aponeuroses', 'ankle', 'slender', 'touton', 'microcyst
ic', 'melanotic', 'pseudopapillary', 'ultrastructurally', 'fatally', 'nonet
s', 'disuse', 'apocrine', 'eccrine', 'osteochondroid', 'syringomas', 'calponi
n', 'actins', 'hidradenomas', 'ossifying', 'e600', 'circumferential', 'macroc
ystic', 'erectile', 'distension', 'serosa', 'ae3', 'misdiagnosis', 'r280t',
'hema', 'lifr', 'sirolimus', 'iq5', 'yxxm', 'bayascas', 'carracedo', 'pandol
fi', 'rudd', 'burke', 'headgroups', 'pik', 'headgroup', 'waterhouse', 'wif',
'1088', 'ascibe', 'anionic', 'fc2', 'guerrero', 'cpbla2', '365k', '204k', 'm
arone', 'alkylphospholipid', 'dennis', '3400', 'na2co3', 'mutation1', 'extra
chromosomal', 'tnfr', 'gangliosides', 'neuroblastic', 'unscreened', 'immunogen
ic', 'mapk10', '5q32', 'h1963', 'h1450', 'pseudocolor', 'sox4', 'segundo', 'c
la', '10q24', 'clumped', 'elapsed', 'sknbe', '21n', '7s', 'hppt1', 'brcal',
'chronological', 'banerjee', 'c57b6', 'c57bl6', 'hematogenous', 'glomerulu
s', 'reported15', 'm90i', 'a77p', '4666', 'a77s', 'r1221', 'leukemia22', 'f82
1', 'q791', 'q401', 'q40', 'syndrome25', 'malignancies23', 'cosmic26', 'pc6',
'shrit1', 'shlacz', 'gtpase24', 'rit2', 'rit', 'stat2', 'stat4', 'pik3ip1',
'heterotopic', 'azacytidine', 'transposons', 'dnmts', 'tedious', 'cyclindepe
ndent', 'sfrp', 'cpeb1', 'sparc', 'garden', 'pitx2', 'ionising', 'nes1', 'cor
ey', 'lbds', 'subnanomolar', 'pet15b', 'tation', '711', 'mineralocorticoid',
'mnk', 's298', 'piggybac', 'zuckerman', 'ptpro', 'transposase', 'inroads',
'mekk', 'uim', 'endometroid', 'induration', 'taxotere', 'al3', 'tcn', 'cells
8', 'cancers12', 'genes10', '2242', '4754', 'cmf', 'analytes', 'matthew', 'bu
reau', 'minors', 'cvad', 'copynumber', 'nr3c2', 'btla', 'analysis19', 'retak
e', 'combo', 'artist', '33p9', 'cells23', 'ottawa', 'helios', 'nonconfirmed',
'actuarial', '130k', 'erg1', 'nmc', 'described42', 'rpl13a', 'sictrl', '6c5',
'wit', 'bioanalytical', '3d4', 'robotic', '001210', 'faller', 'taz2', 'k56',
'c646', 'writers', 'erasers', 'progression4', 'mitsubishi', 'cancer30', 'dra
ining', 'plausibility', 'ace', 'l344a', 'nes2', 'ibid', 'described8', 'cdr',
'cdrs', 'amd', 'rps24', '60s', 'cd123', '1254', 'd5s721', 'd5s23', 'caltec
h', 'rpci', 'bari', '5q11', '2x2', 'mesencephalon', '8n', 'reticulocytosis',

'asci', 'courtesy', 'cj7', 'described30', 'china7', 'hereinafter', 'integrators', '0030', 'security', 'yat', 'bgi', 'kdm5d', 'suv420h1', 'ttrap', 'proto r', 'morevoer', '4jsn', 'ccrccc', 'trd3', 'grabiner', 's6ks', 'k1452n', 'a1519t', 'immunoproprecipitates', 'pld1', 'l1433', 'k1452', 'y1463', 'y1463s', 'l1460', 'a1519', 'q64', 'brrs', 'lipomatosis', '7k', 'pulsing', 'times100', '33k', 'toto', '997', '1001', 'mutations3', 'andthe', 'syndrome3', 'deactivatin g', 'itk', 'agammaglobulinemia', 'laura', 'a13', 'pyrimidin', 'nima', 'bm', 'transphosphorylates', 'formate', 'conduit', 'r544', '146k', 'reverts', '164k', 'maulik', 'sugen', 'chemie', 'zncl2', 'rag2', 'scratches', 'graveel', 'zygotic', '1599', 'myoma', 'expenses', 'erp', 'papillomata', 'eyelashes', 'hypergammaglobulinemia', 'otitis', 'hyperuricemia', 'multicycle', 'ung', 'fgfr2s252w', 'tunicamycin', 'lawrenceville', 'acidification', 'strepavidin', 'serums tarved', 'interfollicular', 'treg', 'ats', 'pdc', 'foxp3', 'inflamed', 'corces', 'mukhopadhyay', 'berx', 'airways', 'rasgrf1', 'r377c', 'organizer', 'hurtado', 'myelofibrosis22', 'addenbrooke', 'belfast', 'n542', 'e543del', 'f537', 'k539delinsl', 'h538qk5391', 'methylmethacralate', 'h26', 'pixera', 'pro150e s', 'mscviresgfp', 'ecopak', 'mutation4', 'k537', 'e543', 'myelofibrosis14', 'k539ltransduced', 'oropharyngeal', 'hpv18', 'prox1', 'caat', '7800', 'keyence', 'nucleoporins', 'sch772984', 'processes9', 'activation22', 'tofacitinib', 'p105', 'deployment', 'fadu', 'dihedral', 'pruning', 'tanimoto', 'nhd', 'nha', 'talon', 'aspartates', 'mapkk1', 'ando', 'nup153', 'ieg', 'reexpressing', 'iegs', 'prediposition', 'fountain', 'loughran', 'reznikoff', 'yeager', 'ambiguously', 'tevelev', 'yarbrough', 'arap', 'fairman', 'ragione', 'appaisal', 'a20p', 'insr113', 'v95a', 'h98y', 'g122s', 'a132p', 'a118t', 'sandig', 'frizelle', 'temperamental', 'faith', 'vagaries', 'fueyo', 'spillare', 'ahrbo m', 'splaying', 'e119', 'n71', 'r103', 'g101', 'phi', 'd84', 'e88', 'm53', 'zulueta', 'komiya', 'olshan', 'neutralise', 'a132', 'sheds', 'a127', 'a127v', 'kyritsis', 'a127e', 'konishi', 'a127p', 'betticher', 'v95', 'hraeus', 'h98p', 'h98', 'xalkoritm', 'g1128s', 'c1156f', 'n1178h', 'e1210k', 'abound', 'alki', 'k299r1', 'k299r2', 'k299r3', 'supm2r1', 'supm2r2', 'supm2r3', 'k299r', 'supm2r', 'ultradeep', 'tables22', 'confronted', 'l1196q', '2500x', '4mkc', 'asp1203', 'g1202', 'nonsynonimous', 'hemmer', 'g254', 'ile1171', 'tae628', 'tables33', 'cys1156', 'glu1210', 'years1', 'involuting', 'ebna', 'described37', 'a147', 'carboxylated', 'absorptiometry', 'separator', 'wald', 'domestic', '5632', 'v1838e', 'variation1', 'uvs', 'ck14', 'ck17', 'classifications13', 'immunotypes', '9345', '5506', 's1760a', 'leu1790', 'phe1761', '1837', '1871', 'n550h', '3743', 'occurrent', '3031', 'f4861', '2878', 'v920a', '3827', '2048', 'k607t', '8801', 'q2858r', 'bubr1a', 'filamina', 'bcc', 'nonfounde r', 'published17a', 'informativeness', 'analysis14', 'variants14', 'p10991', '1054', 't2515i', 'a2717s', 'understanding13', 'imas', '3045', 'jengelman', 'chugai', 'sciencetranslationalmedicine', '120ra17', 'scitranslmed', '3003316', 'transformative', 'resensitizes', 'mgh011', 'mgh020', 'mgh021', 'precrizotinib', 'g340w', '1151t', 'stoked', 'diaminophenylpyrimidine', 'alkpositive', 'phosphoalk', 'alkexpressing', 'nextgeneration', 'phosphortk', 'phosphoerb3', 'treatmentna', 'mgh0nz', 'kitamplified', 'mgh018', 'mekerk', 'modelers', 'ser1206', 'morpholine', 'inconvenient', 'tkina', 'nne', 'incumbent', 'bagro dia', 'rosenzweig', 'camidge', 'doebele', 'tanizaki', 'lovly', 'gandhi', 'wei ckhardt', 'gainor', 'crizotinibnaive', 'gadgeel', 'v1180l', 'ignatius', 'bb b', 'certinib', 'snu2535', 'mgh056', 'alkg1269a', 'sqstm1', 'alkg1202r', 'alki1171t', '3240', 'fs2f', 'emax', 'mager', 'ceff', 'ceffs', 'yamazaki', 'ec90', 'pf06463922', 'mgh051', 'dogs', 'bloodbrain', 'gaussia', '06439015', '06469322', 'penetrable', 'pupil', 'locomotor', 'nct01970865', 'pglycoprotein', 'subq', 'gi50s', '750mg', 'mgh045', '1196', 'alkoxy', 'l1152p', 'k1150', 'jf cr021', 'jfcr', '440g', 'variant1', 'variant3', 'bertholdtechnologies', '1459', 'trihydrate', 'hewish', 'aspx', 'drost', 'desch', 'e541k', 'g857a', 'rs2228006', 'rs1802683', 'e41a', 's46i', 's46n', 'c843y', 'q205p', 't511m', 'g750d', 'r20q', 'g207e', 't485k', 't511a', 'contrariwise', 'nups', 'nephrotic',

'srns', 'podocytes', 'nup205', 'exportin5', 'bmp7', 'eskd', 'glomerulosclerosis', 'fsgs', 'podocyte', 'mapping12', 'a1671', 'a1626', 'a2241', '014669', 'gly591val', 'tyr629cys', 'sequencing11', 'a2403', 'a3256', 'a1394', 'mesangial', 'dilation', 'a1733', '015135', 'phe1995ser', '020750', 'val552ile', 'f1092', 'rnp', 'synaptopodin', 'lys442asnfs', 'addback', 'diic18', 'ab414', 'arg388trp', 'nup20531', 'importin7', 'profibrotic', 'apoptosis', 'nup188', 'nphs11', 'nphs12', 'nphs13', 'allegro42', 'npl', 'clc', '250bp', 'mgc', '21106', '4750923', '28230', '3991335', '168237', '9020614', '74566', '5492046', '8602', '2961238', '34440', '5221801', '8960', '3906006', 'ipo7', '175386', '100014508', 'dest2', 'psirenretroq', 'moin', 'quarterly', 'chromotek', 'glomaxt', 'cim', 'rtca', 'floated', 'fv1000', 'theerthagiri', 'psiren', 'retroq', '374400', '7966', 'nb100', 'mab414', 'nup9349', 'hpa024574', '9526', 'coverglasses', 'sp5x', 'dm6000', 'p210bcr', 'others10', 'asbestos', 'q42', 'diaphragm', 'm8', 'refgene', '32k', 'alex', 'gxl', '10q26', 'i195t', 'guard', 'fmc', 'c242s', 'k0', 'guda', '26g', '004936', '256g', 'd86n', 'insolubility', 'asp86asn', 'hrpt2', 'pro40thr', 'ala23glu', 'infl', 'wolverhampton', 'midlands', 'barts', 'liverpool', 'nihr', 'pesquisas', 'rachou', 'funda', 'oswald', 'belo', 'horizonte', 'eamonn', 'cb2', '2qq', '01223', '746715', '746777', 'erm1000', 'medschl', 'exemplifi', '144700', '193300', 'rs11894252', 'rs7105934', '103g', 'glu35stop', '68c', 'a23e', '118c', 'p40t', 'callithrix', 'jacchus', 'otolemur', 'gametti', 'myotis', 'lucifugus', 'e35x', '2015cancer', 'plstire', 'ala23', 'asp86', 'pro40', 'mcsm', 'duet', 'suffi', 'a23', '165', 'd39', 'k29', 'a19', 'a17', 'r89', 'd86', 'r31', 'i19', 'skrc47', 'ktcl26', 'giris', 'asn41asp', 'leu64arg', 'parafi', 'bromin', 'enlarges', 'cycleave', 'igfir', 'h3t6', '4494', 'mast2', 'graff', 'yb', 'psf', 'greaves', 'bignell', 'cube', 'enumerating', 'meek', 'overestimation', 'interdependency', 'swaps', 'hatch', 'ravi', 'abruptly', 'reconstruct', 'p220', 'supressor', 'hect', 'weissman', 'hai11', 'ibi', 'lane1', 'ser226', 'y134c', '590885', 'exsite', 'daf', 'tras', '3920', 'ctbp1', 'igg2', 'expressiona', 'aher2', '704', 'intratumoural', 'whsc1', 'topographic', 'ubr5', '15q12', 'hu133', 'idelalisib', 'viably', 'nolan', 'rt2', 'naci', 'grb10', '969', 'travel', 'cotranslated', 'aspirin', 'extracellularly', 'pericyte', 'assoian', 'ruffles', 'connexin', 'bifida', 'inductive', 'injuries', 'chondrogenesis', 'shelf', 'phagocytic', 'balloon', 'sano', 'beds', 'fyve', 'rhee', 'wiskott', 'thiery', 'robo', 'delhommeau', 'thol', 'mutations8', 'prcs', 'benzonase', 'cancelled', 'kusabira', 'fertility', 'cynthia', 'tisch', 'tumors10', 'r206h', 'g328e', 'g328v', 'patients19', 'inha', 'o4', 'cancers22', '74x', 'multiexperiment', 'sigclust', 'blingen', 'stainer', 'incisional', 'alk2', 'alp', 'msx2', 'genbanktm', 'hisa', 'bgh', 'bmpr1b', 'acvr2', 'bmpr2', 'solubilize', 'pgcs', 'complexation', 'unexpressed', 'immunophilin', 'endocytosed', 'g356d', 'ldn', '193189', '1b6c', 'arp', 'warp', 'alk6', 'g328w', 'lys27met', 'jinghui', 'majewski', 'arg206his', 'gly328glu', 'gly356asp', 'ventralization', 'dorsalized', 'signaling14', 'gly328val', 'synergic', 'laurent', 'masterpure', 'nv', 'm32', 'irs4', 'years6', 'khong', '131g', 'ionotropic', 'ma1', 'alberto', 'tennessee', 'ventralized', 'jmjd1c', 'prd9', 'smyd3', 'sirt7', 'kdm3b', 'kdm2a', 'uhrf1', 'ncoa1', 'rd54b', 'smarcd2', 'chd2', 'mmp16', 'ml1t10', 'nmu', 'aff1', 'cdh2', 'ptn', 'ap3b1', 'melk', 'nrip1', 'reln', 'sb2f', 'kdm6b', 'hasbeen', 'identifiedas', 'propertiesof', 'digestedwith', 'cesium', 'cellswere', 'wasused', 'atroom', 'wereanalyzed', 'genomiccdna', 'thec', 'didnot', 'thetwo', 'wval', '5and', 'whilethe', 'hwth', 'v11', 'lucap35', 'topota', 'equality', 'ce1', 'ce2', 'orc', 'hiectomy', '2069', 'akr1c3', '5323', 'arbss', 'arbs', 'phosphoryation', 'umekita', 'ise', 'actd', 'tra2', 'ligate', 'invites', 'methodsethics', 'controlthe', 'ar5', 'arr2', 'tmepai', 'hoxb7', 'hijack', 'p69', 'autoregulate', 'transrectal', 'euthanization', 'ah', 'jae', 'disassociated', 'incubators', 'd1s1586', 'd3s1765', 'norwalk', 'kohno', 'igfiir', 'coagulative', 'pl6', '1528', 'bacterium', 'ps3', 'i0', 'bongarzone', 'amv', 'brook', 'sozzi', 'miranda', 'minoletti', 'described4', 'unpurified', 'gh1', 'hcd45', 'hcd34', 'somaticsnipe'

r', 'binned', 'breakdancer', 'graubert', 'visconte', 'przychodzen', 'outgroup', '100m', 'adml', 'kinked', 'quesada', '001810', 'unstranded', 'miso', 'langmead', 'trapnell', 'wagenmakers', 'goseq', 'reichert', '3001', 'nontemplate', 'uridine', 'microbead', 'uag', 'mnd', 'topro', 'znf1', 'znf2', 'hirano', 'crossovers', 'mcam', 'equine', 'bz', 'ishida', 'cajcl', 'icelanders', '2040', '2q31', 'julius', 'lauri', 'kari', 'reykjavik', '955', 'bcas3', 'tbx4', 'athe', '1107', 'recql4', 'mychis', 'xb', 'xar', '139c', 'ser990', 'unaccounted', 'ampflstr', 'identifiler', 'py1604', '1708e', 'mislocalisation', 'al38', 'snupe', 'pzeosv', 'snupe1708f', 'snupe1708r', 'pgal4b', 'equating', 'campos', '1710', '1708v', 'al62', 'abdel', 'wahab', 'bejar', 'expenditure', 'persuasive', 'enforce', 'miyata', 'lyse', 'fwhm', 'torrance', 'policlinico', 'pavia', 'prsv', 'vwf', 'piezo', 'csu', 'deconvolved', 'kinases3', 'accommodation', 'e181', 'k78i', 'stereochemical', 'microsomal', 'nei', 'prbgp2', 'kazuto', 'sayama', '8511', 'kindai', 'snpdb', 'pqclin', 'wis', '045a', '1367', '1370', '1330', '0038', '4363', 'e330a', 'theta', 'erickson', 'students', 'quentin', 'yvelines', 'sponsors', 'marco', 'diary', 'amelogenin', 'hyperextensibility', 'rasg60e', 'rast74p', 'rasa146t', 'a31', 'ecopack', 'krasa146t', 'elsewhere17', 't74p', 'nrasd60e', 'krast74p', 'rasa146', 'hrasa146v', 'nitroso', 'methylurea', 'radiations', 'maine', 'donna', 'vav3', 'rnaase', 'fancl', 'fancm', 'pmep4', 'c307', '1274p', 'thermostable', 'pdbviewer', 'changea', '098', 'diseasecausing', 'a230t', 's387y', 'lasergene', 'g361', 'raji', 'aldo', 'b10', 'm21', 'stricter', 'philips', '5500', 'antioncogenic', 'holden', 'y751', 'proneness', 'isogenomic', 'decamer', 'delitto', 'perfetto', 'igal', 'trp5', 'pgal11', 'hygror', 'rheostatable', 'counterselectable', 'klura3', 'kanmx4', 'hygros', 'ypda', '2ml', 'stationary', 'biospec', 'beater', 'bartlesville', '5415r', 'batavia', 'optiplates', 'victor2', 'pharmenigen', 'lineberger', 'unc', '9819', 'multiinstitutional', '150007', 'p151a', 'm237i', 'p53r2', '1130v', 'a138v', 'c141w', 'p151h', 'r174k', 'r174w', 'p1901', 'h214r', 'l1194p', 'k305m', 'l1194r', 'delaunay', 'tessellations', '134vi', 'sum149pt', 'c242f', 'c242y', 'partakes', 'lukashchuck', 'sestrins', 'sesn1', 'sens2', 'tigar', 'carbazole', 'phikan083', 'kelvin', 'coughing', 'resumes', 'axilla', 'm48', 'conformity', '1303', 'inb', '0045', '3542', '3248', 'ncc', 'h2126', 'f100', 'r415g', 'r272c', 'inp', '293tcells', '1307', '1308', 'garnered', 'slidethe', 'and13', 'enjoyed', 'fluorometric', 'shrunk', 'apoptosis14', 'hou', '2900', 'snuh', 's39', '4472395', '4474004', 'fup', '4471250', '4469496', 'jk1', 'sh1', '1858ra', 'aincludes', 's117', 's129', 's211', 's153', 't1574a', 'r342c', 'g245v', 's61', 's196', 'r82h', 'i179v', 's254', 'rosell', 'wjtog3405', 'hematopathology', 'deacetylated', '9g10', 'grooves', '0a', 'cmr', 'tle', 'mecs', 'sims', 'initiators', 'xa', 'fom', 'chir258', 'septicemia', '350n15', '8f1', 'belloni', 'cardiotoxicity', 'innovations', 'ricerca', 'absorbances', 'egfinduced', 'y1230c', 'nct01639508', 'antennapedia', 'eaiyaapfakk', 'ultrascan', 'remodeler', 'gata6', 'arrayer', 'smarcc2', 'ignoring', 'downes', 'r161', 'k163', 'k164', 'dgfdldl', 'bonneau', 'longy', 'f21a', '294k', 'duerr', 'f21', 'gfdldl', 'd24y', '123f', 'd24n', 'q17a', 'e18a', 'd19a', 'd22a', 'l25a', 't26a', 'g20a', '01228', '00037', '00092', '00034', 'k31e', 'devreotes', 'trotman', 'hbocs', 'q14091', 's1473p', 'e1586g', 'r1589h', 'y1703s', 'w17181', 'catalan', 'pon', 'p38398', 'thursberg', 'disagree', 'y1703', 'h1746', 'k1702', 'w1718', 'g1770', 'benchmarked', 'w1718s', 'wxxxxs', '5720', '5723delctct', 'l1409', 'processivity', 'illuminates', 'adenosines', 'dimmer', 'xpf', 'eta', '00014', 'pneumoniae', 'novobiocin', 'adpnp', 'cripple', 'gpc', '47k', 'nurses', 'personnel', 'codominant', 'apolipoprotein', 'tbrii', 'disuccinimidyl', '558c', '2708c', 'pkd', 'resorption', 'presynaptic', 'can1', 'gliobastoma', 'transcriptomics', 'biorepository', 'emr', 'tapestation', 'icahn', 'casual', 'alemtuzumab', 'reimbursement', 'miaoli', '3341', 'c26g7', '9663', '11644807001', 'w4', 'w5a', 'w5c', 'bellacosa', 'y1021', 'inacti', 'qa', 'tfiiie', 'superstructure', '2m9', 'mikura', 'uhnres', 'utoronto', 'princess', '37232', '3990', 'tafs', '3991', 'idrs', 'disopred2', '3dw', '3biy', '1tot', '1u2

n', 'cr2c', '2lqh', '2lqi', '3io2', 'globule', '2kkj', '2l14', '3992', 's626', '118c', 'cited2', '1r8u', '1p4q', '2ka4', 'pkid', '1kdx', 's111', '1sb0', 'ad2', '2agh', '1hbz', 'tax', '2rny', '1jsp', '2kje', '2k8f', '2ka6', '3p57', '3993', 'bm11', 'pcabinding', '1zoq', '1kbh', '3994', 'ser133', 'y658', 'myeloidlymphoid', 'pcet', 'hbz', 'c4hc3', '3995', 't18', 'encircles', 'n803', 'eklf', 'hcts', '3997', 'uncoloured', 'tfbss', '3998', 's142', 'camkii', 's121', 's436', 'r754', 'theorell', 'snakes', 'e1alike', 'eid', 'pbx1a', 'pbxb', 'diferuloylmethane', 'spice', 'tumeric', 'ischemin', 'dimethylisoxazole', 'chetomin', 'epidithiodiketopiperazine', 'itad', 'depside', 'sekikaic', 'aryl sulfonamide', 'kcn1', 'technicians', 'burge', '5074', '8t', '4987', '5t', '7007g', '7805g', '9501', '9t', '12t', 'exemplify', '0211', 'bonnet', 'berard', '0050', 'leukemia11', '3table', 'pyknotic', 'serosal', '14q22', 'nielsen', 'gmi', 'discontinuations', 'reichardt', 'grossshadern', 'yves', 'meal', 'measurability', 'allocation', '0109', 'antiemetics', 'kaplanmeier', 'omental', 'underdosing', 'granulocytopenia', 'pleuritic', 'lethargy', 'hyclon', 'koreans', 'isel', 'thatcher', 'tokumo', 'dimensionally', 'koutsopoulos', 'm3563', 'cheshire', '155273123', '18ex', '155348122', 'tggctcacaggaccactgatt', '154838u22', 'tcagagcctgtttctacaa', '154899u20', '55634u24', '19ex', 'aaataatcagtgtgattcgtggag', '156027120', 'tgtggagatgagcagggtct', '156107122', 'gaggcca gtgctgtcttaagg', '155750u20', 'gtgcatcgctggtaacatcc', '173160122', '21ex', 'cagctctggctcacactaccag', '173076119', '172656u22', 'gcagcgggttacatcttcttc', '172747u19', 'collet', 'dova', 'argiris', 'comportment', 'tkts', 'delly', 'cmin', 'secretions', 'jt', 'charlestown', '02129', 'farms', 'hd11', 'v03', 'm73', 'californica', 'stefan', '46k', '34k', 'anticyclin', '7o', '53c', 'zang', 'disseminations', 'd18s46', 'g3pdh', 'synthetized', '2023', '2944', 'arthur', 'razor', 'lb953', 'r175a', 'sorensen', 'refractility', 'ifb', '1229', 'rob', 'haras', '958c', '588c', '728c', 'kouhara', 'hawley', 'salomon', 'strehl', 'cip4', 'allergies', 'aspirations', 'dust', '1609', '1afigure', 'meterr', 'aswell', 'combinationwith', '671', 'titled', 'a1a', 'a1b', 'tyr772', 'narctic', 'spurring', 'gatzemeier', 'nondisruptive', 'portends', 'chlorambucil', 'cd274', 'pdcd1lg2', 'observed27', '1578', 'ortega', 'ventrally', 'kitling', 'studies11', '147k', '161k', 'wks', 'r38', '34th', 'kepler', 'noreuil', 'pgk1', 'tumorportal', 'boehm', 'retesting', 'oxnard', 'auy922', 'activity1', 'q5', 'milligrams', 'dpm', '51cr', 'microinvasion', 'ab2', 'richards', 'matsumura', 'phages', 'neuroepithelium', 'judkins', 'ex7', 'bruising', 'invert', 'unrearranged', 'embo', 'ins20', 'del18', 'dele709', 't710insd', 'ic90s', 'azd9291', 'co1686', 'pile', '5305', 'kitakyushu', 'tetsuya', '2366', '0221', '7161', 'mitsudom', 'firstgeneration', 'gpires', 'gfpbased', 'onecell', 'multiplate', '5306', 'a882g', '5307', 'ctrough', 'thoracoscopic', 'rashes', '5308', '5309', 't710insx', 'skip', '5310', '5311', '1850', 'msto', '211h', '9169', 'growthinhibitory', 'ordinating', 'adnexal', 'repeats5', 'vnkr', 'ligase11', 'fold10', 'knobs', 'inholes', 'structbio', 'dpb11', 'rod', 'repeat12', 'proteins13', '1853', '1854', 'interrepeat', 'o23', 'ctip15', 'deacetylas', 'es17', 'transcription18', 'bic19', 'database19', 'repacking', 'core20', 'cyclobutane', 'picopure', 'h1155', 'microfollicles', '1363', 'tyr42', 'scirrhous', 'reports9', 'scicclone', 'phosstop', 'intercepted', '60a', 'multibase', 'methods46', 'latissimus', 'dorsi', 'hypermetabolic', 'k710n', 'retk710n', 'c121', 'azd624459', 'n122d', 'y130h', 'previously59', 'proteins64', 'pathway80', '0099804', 'cells73', 'intercepting', 'procure', 'pck412', 'affirmation', 'h508', 'gog', 'detections', 'abingdon', 'q58delinsl', 'a2a', 'a2b', 'hurdles', 'noncanonical', '7n', '10t', '13t', '16t', '18t', '19t', '23t', 'immunophenotypes', 'braf2', 'stec', 'finelli', 'malgeri', 'opm', 'francomano', 'fracchiolla', 'g384d', 'otsuki', 'translocationwe', 'significative', 'fgfr3s', 'murge', 'linesthe', 'afgf', 'suramin', 'givol', 'u266wt', 'u266y373c', 'cellsto', 'mesenchimal', 'scotet', 'houssaint', 'stenberg', 'lofts', 'plowright', 'nambaa', 'hybridizationtotal', 'dt5', 'aggagctggaggctgac', 'dt3', 'ggagatcttgacacggtgg', 'uf2', 'tggagttccactgcaaggtga', 'plclb', 'hntg', 'constructsthe',

'erytroleukemic', 'kms18', 'p1813', 'aaggacctggtgtcctgtgc', 'p2156', 'agcttga agagctcctcca', 'transfectionsthe', 'muf', 'assaystwenty', 'lymphomagenic', 'interphases', 'raidd', 'cradd', 'iap2', 'reported7', 'me10', 'autoproteolytic', 'rearrangement12', 'elam', 'wotherspoon', 'dixit', 'mathew', 'teddington', 'cells4', 'misinterpreted', 'chr4', 'chr12', 'chr6', 'eleventh', 'myeloblastosis', '080', 'pd11', 'ido1', 'cxcl9', 'tlr7', 'infectivity', 'douville', 'adjuvants', 'immunostimulatory', 'cytotoxicity', 'ssgsea', 'salzberg', 'annette', '45122', 'cryopreservation', 'pecy7', 'culminated', 'sucker', 'mfi', 'interferons', 'vaccination', 'premrna', 'guu', 'uaa', 'ficient', 'allison', 'carbra', 'natali', 'downmodulated', 'jager', 'cormier', 'unmanipulated', 'iec s', 'inos', 'evoking', 'niu', 'ido', 'coinhibitory', 'arginase', 'letterio', 'gorelik', 'steinman', 'intertumoral', 'cohn', 'attenuator', 'burgeoning', 'renaissance', '8755', '9168', 'granada', 'mica', 'illustra', 'adcmv', 'nondisjunction', 'thecell', 'tumorcell', 'bythe', 'tothose', 'supplementedwith', 'dq', 'fractionof', 'obtainedfrom', 'slab', 'describedby', 'denhardt', 'restictionendonucleases', 'concentrationof', 'mediumsupplemented', 'finalconcentration', 'ofa', '1cells', 'ric', 'cwith', 'bglli', 'analyzedby', 'cytofluorometry', 'atthe', 'analysiswas', 'anincrease', 'incubatedunder', 'wereharvested', 'demonstratedthat', 'ofexon', 'thesecound', '453a', 'identifiedin', '1dn a', 'characterizationof', 'llnl', 'lean', 'svf', 'posttest', '8226', 'chondrodysplasias', 'al5', 'tlx1', '3e2', '02114', 'mitani', '131k', 'intermingling', 'veprintsev', 'decameric', 'balagurumoorthy', 'tidow', 'r181e', 'wiman', 'pca2', 'conductivity', 'forcefield', 'hic', 'fgfr2mutant', 'v565i', 'fgfr2n550k', 'fgfr2c383r', 'pamela', 'musk', 'qut', '10016', 'nyumc', 'sylvia', 'postdoctoral', 'de13686', 'ca016087', 'hedge', 'chancellors', 'cdf2', '8002', '121106', 'nct01058434', 'nct00790426', 'nct01379534', 'thirdgeneration', 'pp58', 'dovitinibresistant', 'reauthentication', 'goodfellow', 'en1078d', 'genesearch', 'arundel', 'scoresby', 'byron', '680lt', 'jomar', 'kensington', 'pef1a', 'genelute', 'atgcccagccccacatccag', 'gactggaagccgcattggtg', 'cccaagggccggtcacccg', 'ggcatggtccctgctcagtg', 'pef1', 'fgfr2by376c', 'fgfr2bn550k', 'lycor', '3oy3', 'p459', 'e769', 'm536i', 'i548v', 'n550s', 'e566g', '1618m', 'bdgalactopyranoside', '1762', 't822', 'y770', 'y780', 'y784', 'y813', 'autoflex', 'y770fsx14', 'n550', 'n550t', 'e566a', 'k642n', 'e566', 'benzimidazolequinolinones', 'fortifying', 'i565', '2pwl', '2py3', 'v565', 'm536', 'i548', '1618', 'w3a', 'w3b', 'ligandinduced', 'pd174074', 'fgfr2wild', 'fgfr2y376c', 'reiterates', '2374', '3572', '8749', '1933', 'k660', 'mfe319', 'fgfr2n550kexpressing', 'cognizant', 'adenopathies', '2611', 'ans', 'fj', 'preplanne d', 'listeria', 'hm27', 'hm450', 'mitogenicity', 'urothelia', 'aberrancy', 'huntingdon', 'loo', '3q27', 'dicale', 'spigelman', 'd19s886', 'xeek1', 'g163d', '127c', 'eqkliseedl', 'a14', 'subsection', '1648', 'k178', 'al60', 'wustl', 'stricture', 'leukoplakia', 'fingernails', 'lae', '60c', 'cellfree', 'solgent', 'signagen', 'fma', 'asahina', 'trusolino', 'phosphoproteomics', 'bowen', 'nk4', 'cirrhotic', 'nugc', 'prod', 'f508', '2388delg', 'd17s1322', '1028', '1029', 'rpph1', 'cyclophilin', 'cse', 'syntheses', 'immunostainings', 'seeming', 'maximization', 'posted', 'adaptator', 'immunokinase', 'k78a', 's126', 'ser428', 'ags3', 'psg', 's79', 'exgen', 'euromedex', 'polyvinylidenedifluoride', '1069', 'foxb2', '2afigure', 'decloaking', 'd10s197', 'immunodeficiencies', 'apds', 'bronchiectasis', 'aicd', 'influenzae', 'pax6', 'cells11', 'osteopontin', 'streptococcus', 'osteopenia', 'enfants', 'malades', 'fungus', 'pakistan', 'pneumococcal', 'papulopustular', 'varicella', 'zoster', 'molluscum', 'contagiosum', 'haemophilus', 'sinusitis', 'pararectal', 'osteoarthritis', 'prognathism', 'phlebitis', 'dress', 'osteoclastogenesis', '1610', '1612', '1865c', '1617', 'minegishi', 'dkc', '1618', 'tnfsf11', 'rankl', 'afebrile', 'mice23', 'previously8', 'cytofix', 'cytoperm', 'phosflow', 'perm', 'algorithm13', 'interpolated', 'mmhcc', 's6rp', 'immunochemicals', 'brcaj', '3079', 'brcai', 'hosking', 'l7q21', 'epithehal', 'conresearch', 'lo3unselecteddue', 'intr5', '5296de14', '2576delc', 'intr2', 'ii11', '2569delg', 'q1420x',

'r1835x', '2o8oinsa', 'e730x', '7264', '6170', '3875dei4', 'd173g', 'a1641t', 'r1347g', 's15', 'endometrioid12', 'f667y', 'thral', 'ctgcgc1tfgcactatfgg', 'thra', 'cgggcagcatagcatfgcct', 'agtcctgttagacaaaacctg', 'cag1tfcataccaagu', 'ccf', '1322r', 'ctaggctggcaacaaacga', '17s', 'l322f', 'gcaggaagcaggaatggaa', 'c', '855r', 'acacagactfgtcactgcc', '855f', 'ggatggct1ttagaaaagtgg', '12mm', 'smm', 'c145', '45s', 'defmite', 'l7q', 'd16s767', 'dl9s6ol', 'd3s', 'ntnt', '17s1322', 'racewhite', 'gradeborderline', 'pathoresearch', 'tcp', 'costaine', 'd', 'immunosurveillance', 'progenies', 'capsular', '1193', 'zr75', 'ser240', 'ser244', 'stables', 'hdx', 'py740', 'py751', 'rigidifying', 'arhgdia', 'pri', 'mate', 'signaling13', 'winooski', 'nonreciprocal', 'idealibrary', 'dauer', 's', 'tripes', 'ser380', 'fashions', 'lys558', 'try557', 'studies30', 'activity35', 'bury', 'angular', 'arg181', 'cys4', 'steering', 'ofthese', 'samplesfrom', 'dog1', 'biorobot', 'ofwild', 'aspreviously', 'inaddition', 'caseswith', 'inou', 'r', 'observedin', '50y', '2ui', 'suprapharmacological', 'ic2kitv814', 'sti57', 'issaad', 'penni', 'zermati', 'immunomagnetic', 'milteny', 'ggatgacgagttg', 'gccctaga', 'gtagaacttagaatcgaccggca', 'rnable', 'eurobio', '4times106', '5tim', 'es106', 'times40', 'g559', 'ic2', 'tsujimura', '1times104', 'f3kit', 'beslu', '206k', 'stis71', 'kottaridis', 'p21cdkn1a', 'sperimentale', 'molecolare', 'c', 'nr', 'rejoining', 'jbw301', 'multipolar', 'tams', 'kowalski', 'unlinked', 'tg', 'taaaacgacggccagt', 'avected', 'diverent', 'a16', 'nonresponsiveness', 'snu1', 'vwa', 'mutators', 'thesame', 'vent', 'arrowed', 'borr', 'elastin', 'dermis12', 'suvmax', 'pyd577', 'wkvve', 'pills', 'ostomy', 'hyperglycemia', 'gna1119', 'thoms', 'percents', 'hypokalemia', 'hyponatremia', 'stoma', 'anastomotic', 'zlj9991036390003', '840110', 'alexa598', 'hopman', 'rati', 'figure44', 's380', 'biochain', 'troppmair', 'thauvin', '5535', 'hyperbolic', 'gietz', 'schiestl', 'ind', '308c', 'norit', '1and', '2708', 'hydroxylapatite', '2712', 'houses', 'ivtt', 'i2191', 'p281', 'c22', 'xgal', 'psipred', 'r100p', 's193p', 'e23d', 'c77r', 'k84e', 'i25f', 'n64s', 'g244v', '1653r', 'urological', 'patients2', 'chr10', 'crizotinib16', '19th', 'empire', 'z2', 'preceiver', '2132', '4432', 'affection', 'oncologie', 'vbp', 'grp94', 'oklahoma', 'klausner', 'antenna', '1820', 'arac', 's7h', 'kpn1', 'prestained', 'inhuman', 'agitated', 'brattleboro', 'macprobe', 'xenotransplants', 'sakurada', 'themll', 'combinatorially', 'hoxb3', 'prdm16', '13table', 'sct', 'ciabbreviations', 'przybylski', 'ribonucleic', 'krivtsov', 'watts', 'karuman', 'jishage', 'tiainen', '9th', 'boudeau', 'fated', 'lime', 'oft', 'airc', 'galli', 'laboratorio', 'terapia', 'genica', '56125', 'o50', 'tux', 'ycpgal', 'rs112', 'd98', 'ilv1', 'his3d50', 'prs6', 'his3d30', 'd200', 'pregrown', 'phenylmethysulphonyl', 'chemostatin', 'kimmerly', 'carratore', 'brca1p', 'agr3pk', 'rsy6', 'n132', 's1164i', 'a1789', '3610g4t', '3491g4t', '5484g4a', '5365g4a', '5416t4g', '5295t4g', '1767a4c', '1648a4c', '515c4a', '396c4a', '4654g4t', '4535g4t', '477g4a', '358g4a', 'd120n', '1118a4g', '999a4g', 't333a', '1575t4c', '2576t4c', '2457t4c', 's819p', '3147c4t', '3028c4t', 'p1010s', '5057g4a', '4956g4a', '5123', 't4g', '5324', 'pi403', 'pi397', 'pi258', 'pi222', 'pi340', 'seconddegree', 'e1172x', 'pi432', 'varianta', 'polyphend', 'y179cf', 'f4861f', '1456t4c', 'n550hf', '5075g4a', 'u14860', 'billack', 'gradea', 'dci', 'pi98', 'pi226', 'pi335', 'ccrcal', 'bloomrichardsongradeforbreastcancer', 'collavoli', 'intr', 'aand', 'e143x', 'cultivate', 'ura31', 'mgrrg', 'noninduced', 'experiments7stan', 'dard', 'heteroallellic', 'kupiec', 'earlytruncated', 'r143x', 'hyperrecombinat', 'ion', 'durant', 'ordinated', 'mercurio', 'rock2', 'lmo', 'tal2', 'dn1', 'dn2', 'mistakenly', 'ces', 'polyinosinic', 'polycytidylic', 'repopulated', 'neuraxis', 'archetypal', 'gamou', 'vwrpy', 'affective', 'chelated', 'flowthrough', 'hammerman', 'modulations', 'conscious', 'acetylcholine', 'neurotransmitters', 'hz', 'opioid', 'hypothalamic', 'sorlie', 'type1', 'hospitalar', 'mrf', 'rss', 'antitrypsin', '15q24', 'qmusp', '2745', 'immortalisation', 'xho', 'tlr9', 'l265p', 'virchow', 'mantovani', 'ouerfelli', 'cys64gly', 'deubiquitylation', 'bergman', 'multiformes', '1711', 'p4ebp1', 'gene1', 'ptprk', 'opacity', 'repair1', 'pymt', '71', 'santarosa', 'c3hc4', '1231', 'bogden', 'z9503

098', '1x22', 'z9403097', 'z9503100', 's9600330', 'gcggatcctgcatggaagccattgtc', 'cgggatccacctgtgccaagg', 'gcggatccatctgtggcatgttgg', 'cggaattccaactggaa tggatggta', 'cggaattccaggacagaaaagatc', 'uae', 'z9600205', 'z9600204', 's96002 07', 's9600206', 'lmk', 'lmt', 'lme', 'umk', 'umt', 'umv', '11853', 'cccatttt cctccctcaattcctagaaaa', 'ttttcttaggaattgagggagggaaaatggg', 'gatctttctgtcgctggga ttcttcttgctc', 'gagcaagagaatcccagcgcacagaaaagatc', 'cagttgatctgtggcctgttggtaag g', 'ccttcaccaacaggcccacagatcaactg', 'cagttgatctgtggcgtgttggtaagg', 'ccttcaccaacaagcccacagatcaactg', 'cagttgatctgtggcacgttggtaagg', 'ccttcaccaacgagcccacagatcaact g', 'cagttgatctgtggcacgttggtaagg', 'ccttcaccaacgtgcccacagatcaactg', 'gcggat ccaggttaggtgtcc', 'pcl1', 'cotransformation', 'galctopyranoside', 'psg424', 'ptashne', 'pg5e1bluc', '1636', 'spuriously', 'hybridizes', 'okano', 'select in', 'thymi', 'skew', 'emek', 'ror', 'slam', 'arrayexpress', 'mexp', 'adsorption', 'rasgefs', 'vernier', 'i3', 'romero', 'dance', 'e64', 'aki', 'fruman', 'hbond', 'fondazione', 'mq', 'dhap', 'carmustine', 'mice11', 'pmaip1', 'tony', '500x', 'defects5', 'amacr', 'map3k3', 'allowance', 'booster', 'sloane', 'jbc', 'radiologists', 'kruskall', 'complained', 'al34', '35a', 'unplanned', 'egfrtkis', 'innately', 'tosei', 'scintigram', 'acidlocked', 'fukihara', '000 357129', 'cused', 'shukuya', 'sungkyunkwan', 'soonsoo21c', 'silkahn', 'skku', 'chemother', 's00280', '1876', 'phe723leu', 'ser720phe', 'leu747pro', 'val786 met', 'ser719ile', 'ser729phe', '0101', 'asp842val', 'series24', 'operable', 'val561asp', 'lys642glu', 'noncategorized', 'trp557', 'tyr503dup', 'leu576pr o', 'trp557arg', 'val560del', 'val559asp', 'val560asp', 'val559del', 'val559ala', 'trp557gly', 'asp842tyr', 'val559gly', 'val560glu', 'val560gln', 'trp557 lys', 'asp842', 'covariable', 's492r', 'd38b1', '15g', 'e203', 'benzoxadiazol', 'lapointe', 'polymerisation', 'src1', 'amegakaryocytic', 'purpura', 'ercc 4', 'xrcc9', 'xp2yo', 'cells42', '5277', '59dup12', 'immobile', 'neuropathies', 'radioisotopes', 'skypaint', 'triplicated', 'chambon', 'sedimented', 'esh ap', 'methylprednisolone', 'd68f8', 'transaminitis', 'rads', 'competitiveness', 'schiller', 'ptprm', 'omicron', 'hptp', 'extracellular', 'memprin', 'heterophilic', 'extracellular', 'stablizing', 'f74s', 'a209t', 'k218t', 'f248s', 'y2 80h', 'meditates', 'k788e', 'r790i', 'r49', 'hs683', '033360', 'centuries', 'xii', 'rieger', 'teething', 'hypogammaglobulinemia', 'bologna', 'ivanov', 'clade', 'maize', 'tyr905', 'microassay', 'hw', 'ttaggg', 'resensitize', 'genes22', 'polyscan', 'taira', 'reluctant', 'osteolytic', '46f', 'atomistic', '3791', 'pox', 'ector', 'and5b', 'hungarian', 'meric', 'cd80', 'fdxr', 'tp53 i3', 'intradimer', 'interdimer', 'e180', '859kb', 'p53dbd', '2ady', 'wt344a', 'ee344a', 'rr344a', 'e180r', 'r180', '876kb', 'bh123', '0r', '1889', 'm0027 2', 'acatgtctgaacatg', 'gcwtgt', '917kb', 'p21cdnk1a', '768kb', 'iaspp', 'r18 1h', 'eh', 'r181p', 'unnatural', 'activatability', '849kb', 'catg', 'caag', 'cttg', 'ctag', 'h14', 'functionmodel', '458kb', 'symbolizes', 'orazi', 'sykes', 'petechiae', 'cancer26', 'lat', 'medac', 'described32'])

In []: #####

```
In [51]: dict_list = []
# dict_list =[] contains 9 dictionaries each corresponds to a class
for i in range(1,10):
    cls_text = train_df[train_df['Class']==i]
    # build a word dict based on the words in that class
    dict_list.append(extract_dictionary_paddle(cls_text))
    # append it to dict_list

# dict_list[i] is build on i'th class text data
# total_dict is buid on whole training text data
total_dict = extract_dictionary_paddle(train_df)

confuse_array = []
for i in train_text_features:
    ratios = []
    max_val = -1
    for j in range(0,9):
        ratios.append((dict_list[j][i]+10)/(total_dict[i]+90))
    confuse_array.append(ratios)
confuse_array = np.array(confuse_array)
```

```
In [52]: #response coding of text features
train_text_feature_responseCoding = get_text_responsecoding(train_df)
test_text_feature_responseCoding = get_text_responsecoding(test_df)
cv_text_feature_responseCoding = get_text_responsecoding(cv_df)
```

```
In [53]: # https://stackoverflow.com/a/16202486
# we convert each row values such that they sum to 1
train_text_feature_responseCoding = (train_text_feature_responseCoding.T/train_text_feature_responseCoding.sum(axis=1)).T
test_text_feature_responseCoding = (test_text_feature_responseCoding.T/test_text_feature_responseCoding.sum(axis=1)).T
cv_text_feature_responseCoding = (cv_text_feature_responseCoding.T/cv_text_feature_responseCoding.sum(axis=1)).T
```

```
In [54]: train_text_feature_responseCoding[0].sum()
```

```
Out[54]: 1.0
```

```
In [55]: # don't forget to normalize every feature
train_text_feature_onehotCoding = normalize(train_text_feature_onehotCoding, axis=0)

# we use the same vectorizer that was trained on train data
test_text_feature_onehotCoding = text_vectorizer.transform(test_df[ 'TEXT' ])
# don't forget to normalize every feature
test_text_feature_onehotCoding = normalize(test_text_feature_onehotCoding, axis=0)

# we use the same vectorizer that was trained on train data
cv_text_feature_onehotCoding = text_vectorizer.transform(cv_df[ 'TEXT' ])
# don't forget to normalize every feature
cv_text_feature_onehotCoding = normalize(cv_text_feature_onehotCoding, axis=0)
```

```
In [56]: test_text_feature_onehotCoding
```

```
Out[56]: <665x53213 sparse matrix of type '<class 'numpy.float64'>'  
with 1040910 stored elements in Compressed Sparse Column format>
```

```
In [57]: #https://stackoverflow.com/a/2258273/4084039
sorted_text_fea_dict = dict(sorted(text_fea_dict.items(), key=lambda x: x[1] , reverse=True))
sorted_text_occur = np.array(list(sorted_text_fea_dict.values()))
```

```
In [58]: # Number of words for a given frequency.  
print(Counter(sorted_text_occur))
```

Counter({3: 5266, 4: 3562, 6: 2877, 5: 2801, 8: 2216, 9: 1822, 10: 1689, 7: 1
 649, 11: 1275, 12: 1201, 13: 1104, 14: 894, 16: 884, 15: 790, 18: 663, 20: 62
 4, 24: 579, 17: 572, 21: 569, 19: 481, 22: 441, 32: 387, 30: 385, 23: 374, 2
 5: 367, 26: 347, 28: 339, 29: 338, 27: 333, 38: 321, 33: 271, 51: 265, 36: 26
 5, 31: 254, 34: 246, 40: 231, 35: 227, 39: 225, 42: 216, 48: 197, 37: 194, 4
 7: 189, 43: 188, 45: 175, 41: 175, 44: 170, 55: 159, 56: 153, 52: 153, 54: 15
 1, 46: 150, 50: 147, 60: 141, 57: 134, 53: 131, 49: 126, 62: 125, 69: 124, 6
 4: 123, 63: 120, 61: 116, 66: 115, 58: 113, 72: 112, 67: 108, 70: 102, 59: 9
 8, 65: 97, 80: 93, 78: 88, 68: 87, 84: 80, 76: 77, 74: 76, 73: 75, 94: 74, 8
 6: 74, 81: 74, 77: 74, 71: 73, 83: 72, 90: 71, 102: 69, 75: 68, 87: 65, 96: 6
 4, 85: 64, 98: 62, 97: 62, 79: 62, 93: 61, 88: 59, 118: 58, 95: 58, 82: 57, 1
 29: 56, 108: 56, 121: 55, 92: 55, 99: 53, 120: 52, 104: 51, 116: 50, 115: 50,
 105: 50, 91: 50, 89: 49, 109: 48, 106: 48, 122: 47, 100: 47, 114: 46, 107: 4
 6, 110: 45, 111: 44, 103: 44, 101: 44, 117: 43, 133: 42, 113: 41, 112: 41, 15
 0: 40, 127: 39, 153: 38, 138: 38, 119: 38, 157: 37, 145: 37, 126: 37, 123: 3
 7, 132: 36, 128: 36, 124: 36, 192: 35, 167: 35, 158: 34, 146: 34, 142: 34, 13
 5: 34, 130: 34, 197: 33, 156: 33, 131: 33, 140: 32, 139: 32, 173: 31, 164: 3
 1, 152: 31, 137: 31, 136: 31, 134: 30, 125: 30, 177: 29, 155: 29, 151: 29, 14
 4: 28, 179: 27, 172: 27, 149: 27, 148: 27, 175: 26, 170: 26, 160: 26, 147: 2
 6, 208: 25, 166: 25, 165: 25, 162: 25, 247: 24, 213: 24, 196: 24, 182: 24, 15
 4: 24, 223: 23, 212: 23, 191: 23, 186: 23, 180: 23, 169: 23, 168: 23, 143: 2
 3, 260: 22, 217: 22, 209: 22, 202: 22, 161: 22, 159: 22, 294: 21, 284: 21, 22
 2: 21, 220: 21, 194: 21, 193: 21, 189: 21, 237: 20, 221: 20, 207: 20, 204: 2
 0, 171: 20, 141: 20, 326: 19, 240: 19, 234: 19, 228: 19, 201: 19, 198: 19, 19
 5: 19, 185: 19, 183: 19, 181: 19, 174: 19, 288: 18, 245: 18, 233: 18, 231: 1
 8, 227: 18, 190: 18, 176: 18, 272: 17, 267: 17, 264: 17, 236: 17, 229: 17, 21
 9: 17, 211: 17, 203: 17, 200: 17, 184: 17, 163: 17, 302: 16, 271: 16, 249: 1
 6, 239: 16, 224: 16, 216: 16, 206: 16, 199: 16, 188: 16, 318: 15, 306: 15, 28
 9: 15, 287: 15, 280: 15, 275: 15, 274: 15, 270: 15, 266: 15, 246: 15, 241: 1
 5, 230: 15, 215: 15, 210: 15, 187: 15, 359: 14, 308: 14, 304: 14, 303: 14, 29
 1: 14, 290: 14, 286: 14, 283: 14, 268: 14, 243: 14, 235: 14, 232: 14, 226: 1
 4, 390: 13, 356: 13, 341: 13, 324: 13, 305: 13, 277: 13, 263: 13, 257: 13, 25
 5: 13, 254: 13, 253: 13, 244: 13, 225: 13, 214: 13, 205: 13, 178: 13, 491: 1
 2, 400: 12, 342: 12, 339: 12, 314: 12, 297: 12, 282: 12, 256: 12, 248: 12, 23
 8: 12, 218: 12, 406: 11, 392: 11, 373: 11, 348: 11, 344: 11, 340: 11, 317: 1
 1, 312: 11, 307: 11, 293: 11, 269: 11, 265: 11, 262: 11, 261: 11, 433: 10, 41
 7: 10, 405: 10, 397: 10, 370: 10, 367: 10, 331: 10, 328: 10, 323: 10, 301: 1
 0, 300: 10, 299: 10, 298: 10, 296: 10, 295: 10, 258: 10, 251: 10, 250: 10, 24
 2: 10, 550: 9, 545: 9, 446: 9, 438: 9, 395: 9, 387: 9, 379: 9, 364: 9, 363:
 9, 354: 9, 349: 9, 347: 9, 338: 9, 337: 9, 332: 9, 327: 9, 322: 9, 320: 9, 31
 0: 9, 309: 9, 285: 9, 276: 9, 611: 8, 596: 8, 542: 8, 540: 8, 528: 8, 522: 8,
 517: 8, 515: 8, 480: 8, 477: 8, 468: 8, 464: 8, 457: 8, 447: 8, 442: 8, 435:
 8, 432: 8, 431: 8, 424: 8, 423: 8, 418: 8, 412: 8, 404: 8, 401: 8, 386: 8, 37
 8: 8, 377: 8, 372: 8, 362: 8, 351: 8, 350: 8, 333: 8, 330: 8, 325: 8, 279: 8,
 278: 8, 273: 8, 1289: 7, 767: 7, 673: 7, 660: 7, 659: 7, 605: 7, 541: 7, 509:
 7, 507: 7, 492: 7, 488: 7, 486: 7, 485: 7, 452: 7, 416: 7, 407: 7, 403: 7, 39
 6: 7, 393: 7, 388: 7, 382: 7, 375: 7, 374: 7, 369: 7, 361: 7, 353: 7, 352: 7,
 346: 7, 343: 7, 336: 7, 319: 7, 292: 7, 259: 7, 1278: 6, 1012: 6, 925: 6, 86
 7: 6, 759: 6, 725: 6, 685: 6, 675: 6, 654: 6, 646: 6, 635: 6, 599: 6, 584: 6,
 578: 6, 572: 6, 563: 6, 562: 6, 560: 6, 552: 6, 546: 6, 539: 6, 538: 6, 530:
 6, 519: 6, 505: 6, 501: 6, 498: 6, 497: 6, 472: 6, 462: 6, 459: 6, 456: 6, 45
 4: 6, 443: 6, 441: 6, 436: 6, 419: 6, 414: 6, 409: 6, 408: 6, 391: 6, 365: 6,
 360: 6, 358: 6, 357: 6, 321: 6, 315: 6, 281: 6, 252: 6, 1495: 5, 1140: 5, 101
 7: 5, 984: 5, 956: 5, 905: 5, 892: 5, 851: 5, 835: 5, 833: 5, 821: 5, 805: 5,
 804: 5, 802: 5, 797: 5, 790: 5, 776: 5, 775: 5, 772: 5, 718: 5, 694: 5, 687:
 5, 684: 5, 676: 5, 664: 5, 636: 5, 634: 5, 633: 5, 632: 5, 620: 5, 597: 5, 58
 8: 5, 576: 5, 575: 5, 574: 5, 569: 5, 553: 5, 551: 5, 549: 5, 548: 5, 536: 5,

534: 5, 525: 5, 510: 5, 508: 5, 504: 5, 495: 5, 481: 5, 469: 5, 467: 5, 465: 5, 463: 5, 453: 5, 448: 5, 444: 5, 439: 5, 429: 5, 422: 5, 420: 5, 415: 5, 410: 5, 399: 5, 398: 5, 394: 5, 389: 5, 383: 5, 376: 5, 371: 5, 368: 5, 355: 5, 345: 5, 335: 5, 2170: 4, 1812: 4, 1632: 4, 1541: 4, 1462: 4, 1397: 4, 1358: 4, 1350: 4, 1280: 4, 1276: 4, 1234: 4, 1229: 4, 1204: 4, 1148: 4, 1134: 4, 1086: 4, 1039: 4, 985: 4, 979: 4, 972: 4, 964: 4, 936: 4, 934: 4, 896: 4, 893: 4, 889: 4, 865: 4, 862: 4, 861: 4, 858: 4, 849: 4, 844: 4, 843: 4, 831: 4, 822: 4, 817: 4, 803: 4, 800: 4, 793: 4, 786: 4, 781: 4, 770: 4, 769: 4, 766: 4, 761: 4, 749: 4, 735: 4, 729: 4, 723: 4, 720: 4, 717: 4, 706: 4, 705: 4, 704: 4, 700: 4, 697: 4, 696: 4, 669: 4, 668: 4, 658: 4, 653: 4, 650: 4, 648: 4, 645: 4, 644: 4, 643: 4, 630: 4, 628: 4, 623: 4, 618: 4, 616: 4, 612: 4, 610: 4, 609: 4, 607: 4, 606: 4, 601: 4, 592: 4, 581: 4, 568: 4, 555: 4, 529: 4, 518: 4, 514: 4, 513: 4, 512: 4, 506: 4, 503: 4, 502: 4, 494: 4, 489: 4, 487: 4, 483: 4, 479: 4, 478: 4, 475: 4, 474: 4, 471: 4, 466: 4, 461: 4, 460: 4, 455: 4, 450: 4, 434: 4, 430: 4, 426: 4, 411: 4, 385: 4, 366: 4, 334: 4, 316: 4, 313: 4, 311: 4, 2951: 3, 2562: 3, 2525: 3, 2456: 3, 2423: 3, 2348: 3, 2313: 3, 2228: 3, 2180: 3, 2138: 3, 2117: 3, 2091: 3, 2077: 3, 2046: 3, 2026: 3, 1973: 3, 1951: 3, 1933: 3, 1838: 3, 1834: 3, 1777: 3, 1742: 3, 1735: 3, 1687: 3, 1679: 3, 1610: 3, 1602: 3, 1588: 3, 1587: 3, 1557: 3, 1530: 3, 1524: 3, 1510: 3, 1500: 3, 1480: 3, 1470: 3, 1415: 3, 1387: 3, 1377: 3, 1375: 3, 1369: 3, 1334: 3, 1295: 3, 1285: 3, 1270: 3, 1255: 3, 1240: 3, 1237: 3, 1230: 3, 1224: 3, 1223: 3, 1215: 3, 1197: 3, 1188: 3, 1174: 3, 1171: 3, 1168: 3, 1155: 3, 1151: 3, 1130: 3, 1116: 3, 1114: 3, 1110: 3, 1109: 3, 1106: 3, 1103: 3, 1095: 3, 1093: 3, 1091: 3, 1079: 3, 1077: 3, 1062: 3, 1056: 3, 1053: 3, 1050: 3, 1022: 3, 1009: 3, 1005: 3, 998: 3, 987: 3, 962: 3, 955: 3, 954: 3, 948: 3, 939: 3, 938: 3, 909: 3, 906: 3, 903: 3, 902: 3, 901: 3, 881: 3, 872: 3, 871: 3, 854: 3, 847: 3, 841: 3, 839: 3, 837: 3, 832: 3, 829: 3, 828: 3, 825: 3, 813: 3, 810: 3, 807: 3, 789: 3, 782: 3, 780: 3, 762: 3, 758: 3, 755: 3, 752: 3, 751: 3, 737: 3, 736: 3, 731: 3, 726: 3, 724: 3, 715: 3, 713: 3, 703: 3, 698: 3, 695: 3, 686: 3, 682: 3, 677: 3, 672: 3, 671: 3, 667: 3, 666: 3, 665: 3, 661: 3, 657: 3, 656: 3, 655: 3, 642: 3, 639: 3, 637: 3, 619: 3, 598: 3, 595: 3, 594: 3, 593: 3, 589: 3, 587: 3, 583: 3, 582: 3, 580: 3, 579: 3, 577: 3, 570: 3, 567: 3, 566: 3, 565: 3, 561: 3, 558: 3, 556: 3, 547: 3, 532: 3, 531: 3, 524: 3, 523: 3, 520: 3, 500: 3, 499: 3, 496: 3, 473: 3, 470: 3, 458: 3, 449: 3, 440: 3, 437: 3, 427: 3, 425: 3, 421: 3, 384: 3, 381: 3, 380: 3, 329: 3, 7648: 2, 6976: 2, 6775: 2, 6351: 2, 6179: 2, 5963: 2, 5707: 2, 5651: 2, 5362: 2, 5156: 2, 5032: 2, 4948: 2, 4851: 2, 4835: 2, 4532: 2, 4446: 2, 4437: 2, 4345: 2, 4322: 2, 4305: 2, 4171: 2, 4165: 2, 4102: 2, 4023: 2, 3998: 2, 3996: 2, 3944: 2, 3672: 2, 3657: 2, 3602: 2, 3556: 2, 3539: 2, 3496: 2, 3491: 2, 3430: 2, 3400: 2, 3391: 2, 3376: 2, 3374: 2, 3294: 2, 3285: 2, 3253: 2, 3247: 2, 3192: 2, 3190: 2, 3177: 2, 3160: 2, 3154: 2, 3122: 2, 3103: 2, 3043: 2, 2957: 2, 2908: 2, 2864: 2, 2863: 2, 2861: 2, 2859: 2, 2842: 2, 2746: 2, 2734: 2, 2725: 2, 2691: 2, 2683: 2, 2676: 2, 2668: 2, 2644: 2, 2640: 2, 2621: 2, 2600: 2, 2577: 2, 2576: 2, 2546: 2, 2537: 2, 2513: 2, 2489: 2, 2466: 2, 2462: 2, 2438: 2, 2428: 2, 2422: 2, 2407: 2, 2398: 2, 2391: 2, 2388: 2, 2375: 2, 2357: 2, 2355: 2, 2350: 2, 2345: 2, 2314: 2, 2311: 2, 2296: 2, 2292: 2, 2254: 2, 2248: 2, 2215: 2, 2197: 2, 2143: 2, 2135: 2, 2125: 2, 2123: 2, 2119: 2, 2105: 2, 2097: 2, 2081: 2, 2047: 2, 2044: 2, 2042: 2, 2035: 2, 2033: 2, 2027: 2, 2012: 2, 2011: 2, 1998: 2, 1984: 2, 1975: 2, 1974: 2, 1972: 2, 1968: 2, 1966: 2, 1957: 2, 1925: 2, 1908: 2, 1903: 2, 1902: 2, 1901: 2, 1895: 2, 1877: 2, 1861: 2, 1860: 2, 1821: 2, 1810: 2, 1808: 2, 1805: 2, 1804: 2, 1801: 2, 1792: 2, 1787: 2, 1784: 2, 1782: 2, 1781: 2, 1775: 2, 1740: 2, 1727: 2, 1720: 2, 1715: 2, 1691: 2, 1688: 2, 1680: 2, 1674: 2, 1672: 2, 1660: 2, 1650: 2, 1637: 2, 1636: 2, 1635: 2, 1631: 2, 1625: 2, 1621: 2, 1620: 2, 1615: 2, 1614: 2, 1607: 2, 1597: 2, 1592: 2, 1584: 2, 1575: 2, 1567: 2, 1562: 2, 1549: 2, 1539: 2, 1537: 2, 1535: 2, 1533: 2, 1531: 2, 1519: 2, 1514: 2, 1513: 2, 1491: 2, 1484: 2, 1481: 2, 1475: 2, 1460: 2, 1459: 2, 1455: 2, 1442: 2, 1438: 2, 1423: 2, 1421:

2, 1419: 2, 1417: 2, 1416: 2, 1403: 2, 1394: 2, 1390: 2, 1382: 2, 1381: 2, 13
 76: 2, 1372: 2, 1356: 2, 1354: 2, 1349: 2, 1347: 2, 1344: 2, 1343: 2, 1330:
 2, 1327: 2, 1325: 2, 1322: 2, 1321: 2, 1320: 2, 1316: 2, 1314: 2, 1312: 2, 13
 08: 2, 1307: 2, 1303: 2, 1299: 2, 1298: 2, 1297: 2, 1296: 2, 1293: 2, 1291:
 2, 1283: 2, 1281: 2, 1274: 2, 1273: 2, 1272: 2, 1269: 2, 1265: 2, 1262: 2, 12
 54: 2, 1253: 2, 1252: 2, 1250: 2, 1248: 2, 1247: 2, 1245: 2, 1236: 2, 1233:
 2, 1227: 2, 1225: 2, 1218: 2, 1212: 2, 1211: 2, 1209: 2, 1200: 2, 1198: 2, 11
 95: 2, 1193: 2, 1182: 2, 1181: 2, 1177: 2, 1170: 2, 1169: 2, 1163: 2, 1156:
 2, 1152: 2, 1147: 2, 1144: 2, 1143: 2, 1133: 2, 1128: 2, 1122: 2, 1121: 2, 11
 15: 2, 1113: 2, 1096: 2, 1094: 2, 1092: 2, 1089: 2, 1082: 2, 1081: 2, 1076:
 2, 1073: 2, 1070: 2, 1066: 2, 1063: 2, 1059: 2, 1054: 2, 1044: 2, 1041: 2, 10
 37: 2, 1033: 2, 1032: 2, 1029: 2, 1026: 2, 1023: 2, 1019: 2, 1018: 2, 1004:
 2, 1001: 2, 1000: 2, 997: 2, 994: 2, 993: 2, 992: 2, 988: 2, 986: 2, 982: 2,
 981: 2, 980: 2, 978: 2, 974: 2, 963: 2, 958: 2, 953: 2, 949: 2, 947: 2, 946:
 2, 933: 2, 932: 2, 931: 2, 930: 2, 923: 2, 921: 2, 916: 2, 915: 2, 913: 2, 90
 7: 2, 904: 2, 900: 2, 897: 2, 895: 2, 891: 2, 887: 2, 886: 2, 877: 2, 870: 2,
 866: 2, 863: 2, 859: 2, 852: 2, 850: 2, 848: 2, 846: 2, 842: 2, 838: 2, 834:
 2, 826: 2, 824: 2, 819: 2, 816: 2, 814: 2, 811: 2, 806: 2, 801: 2, 799: 2, 79
 8: 2, 796: 2, 792: 2, 788: 2, 784: 2, 783: 2, 779: 2, 778: 2, 777: 2, 756: 2,
 753: 2, 746: 2, 745: 2, 744: 2, 743: 2, 740: 2, 732: 2, 728: 2, 727: 2, 721:
 2, 719: 2, 714: 2, 712: 2, 711: 2, 709: 2, 708: 2, 707: 2, 701: 2, 699: 2, 69
 3: 2, 691: 2, 689: 2, 683: 2, 681: 2, 679: 2, 674: 2, 670: 2, 649: 2, 647: 2,
 641: 2, 638: 2, 631: 2, 627: 2, 626: 2, 624: 2, 622: 2, 621: 2, 614: 2, 608:
 2, 604: 2, 603: 2, 590: 2, 585: 2, 573: 2, 564: 2, 559: 2, 557: 2, 544: 2, 54
 3: 2, 537: 2, 533: 2, 527: 2, 516: 2, 511: 2, 490: 2, 484: 2, 482: 2, 451: 2,
 413: 2, 151773: 1, 120275: 1, 82046: 1, 67428: 1, 66832: 1, 66656: 1, 66451:
 1, 63947: 1, 63611: 1, 54588: 1, 54087: 1, 50974: 1, 50157: 1, 47220: 1, 4635
 2: 1, 43698: 1, 43111: 1, 42742: 1, 42704: 1, 40873: 1, 40761: 1, 40338: 1, 3
 9052: 1, 38613: 1, 37531: 1, 36905: 1, 36681: 1, 35887: 1, 35739: 1, 35259:
 1, 34664: 1, 33859: 1, 33250: 1, 32998: 1, 32568: 1, 32043: 1, 29161: 1, 2841
 6: 1, 27565: 1, 27034: 1, 26519: 1, 26189: 1, 26140: 1, 25918: 1, 24944: 1, 2
 4862: 1, 24525: 1, 24353: 1, 24239: 1, 24220: 1, 24146: 1, 23512: 1, 22761:
 1, 22618: 1, 22410: 1, 22080: 1, 21928: 1, 21921: 1, 21665: 1, 21267: 1, 2124
 8: 1, 21238: 1, 20615: 1, 20378: 1, 20307: 1, 19725: 1, 19457: 1, 19406: 1, 1
 9302: 1, 19119: 1, 19094: 1, 19051: 1, 18803: 1, 18694: 1, 18333: 1, 18177:
 1, 18153: 1, 18074: 1, 18018: 1, 17954: 1, 17828: 1, 17765: 1, 17595: 1, 1750
 7: 1, 17495: 1, 17372: 1, 17317: 1, 17196: 1, 17121: 1, 17120: 1, 16915: 1, 1
 6796: 1, 16763: 1, 16729: 1, 16646: 1, 16627: 1, 15913: 1, 15901: 1, 15870:
 1, 15743: 1, 15591: 1, 15536: 1, 15436: 1, 15321: 1, 15300: 1, 15223: 1, 1513
 1: 1, 15098: 1, 14962: 1, 14888: 1, 14857: 1, 14841: 1, 14608: 1, 14583: 1, 1
 4547: 1, 14536: 1, 14511: 1, 14503: 1, 14336: 1, 14287: 1, 14221: 1, 14114:
 1, 13827: 1, 13704: 1, 13629: 1, 13557: 1, 13527: 1, 13494: 1, 13422: 1, 1321
 3: 1, 13189: 1, 13139: 1, 13048: 1, 13002: 1, 12994: 1, 12965: 1, 12961: 1, 1
 2953: 1, 12829: 1, 12803: 1, 12789: 1, 12762: 1, 12720: 1, 12637: 1, 12608:
 1, 12586: 1, 12533: 1, 12520: 1, 12471: 1, 12370: 1, 12354: 1, 12315: 1, 1230
 8: 1, 12252: 1, 12247: 1, 12227: 1, 12215: 1, 12175: 1, 12130: 1, 12058: 1, 1
 2056: 1, 12055: 1, 12040: 1, 11953: 1, 11821: 1, 11736: 1, 11700: 1, 11688:
 1, 11604: 1, 11504: 1, 11478: 1, 11452: 1, 11442: 1, 11417: 1, 11300: 1, 1128
 4: 1, 11240: 1, 11213: 1, 11107: 1, 11000: 1, 10944: 1, 10911: 1, 10893: 1, 1
 0870: 1, 10819: 1, 10805: 1, 10783: 1, 10723: 1, 10708: 1, 10588: 1, 10535:
 1, 10526: 1, 10450: 1, 10415: 1, 10283: 1, 10241: 1, 10221: 1, 10116: 1, 1010
 7: 1, 10090: 1, 10067: 1, 10054: 1, 10051: 1, 10016: 1, 9952: 1, 9924: 1, 991
 0: 1, 9858: 1, 9857: 1, 9805: 1, 9793: 1, 9652: 1, 9647: 1, 9630: 1, 9546: 1,
 9538: 1, 9530: 1, 9512: 1, 9499: 1, 9415: 1, 9375: 1, 9292: 1, 9271: 1, 9268:
 1, 9263: 1, 9245: 1, 9234: 1, 9223: 1, 9215: 1, 9175: 1, 9128: 1, 9127: 1, 90
 79: 1, 9073: 1, 9063: 1, 9052: 1, 9048: 1, 9020: 1, 8963: 1, 8933: 1, 8890:
 1, 8860: 1, 8811: 1, 8810: 1, 8738: 1, 8708: 1, 8635: 1, 8634: 1, 8621: 1, 86

16: 1, 8608: 1, 8586: 1, 8567: 1, 8564: 1, 8537: 1, 8509: 1, 8473: 1, 8429: 1, 8423: 1, 8369: 1, 8339: 1, 8311: 1, 8294: 1, 8270: 1, 8258: 1, 8162: 1, 8148: 1, 8147: 1, 8142: 1, 8127: 1, 8120: 1, 8114: 1, 8086: 1, 8074: 1, 8059: 1, 8052: 1, 8004: 1, 7974: 1, 7930: 1, 7872: 1, 7853: 1, 7847: 1, 7843: 1, 7835: 1, 7826: 1, 7794: 1, 7759: 1, 7747: 1, 7732: 1, 7690: 1, 7652: 1, 7641: 1, 7614: 1, 7605: 1, 7588: 1, 7579: 1, 7552: 1, 7534: 1, 7483: 1, 7469: 1, 7458: 1, 7414: 1, 7389: 1, 7378: 1, 7334: 1, 7328: 1, 7311: 1, 7305: 1, 7295: 1, 7293: 1, 7268: 1, 7267: 1, 7256: 1, 7248: 1, 7239: 1, 7190: 1, 7182: 1, 7170: 1, 7166: 1, 7163: 1, 7160: 1, 7148: 1, 7143: 1, 7119: 1, 7090: 1, 7085: 1, 7076: 1, 7027: 1, 7000: 1, 6991: 1, 6983: 1, 6981: 1, 6963: 1, 6933: 1, 6930: 1, 6923: 1, 6903: 1, 6888: 1, 6873: 1, 6859: 1, 6851: 1, 6842: 1, 6828: 1, 6822: 1, 6814: 1, 6703: 1, 6699: 1, 6696: 1, 6694: 1, 6688: 1, 6651: 1, 6644: 1, 6631: 1, 6585: 1, 6567: 1, 6565: 1, 6545: 1, 6540: 1, 6533: 1, 6503: 1, 6445: 1, 6441: 1, 6369: 1, 6361: 1, 6347: 1, 6346: 1, 6339: 1, 6323: 1, 6311: 1, 6309: 1, 6304: 1, 6299: 1, 6294: 1, 6249: 1, 6248: 1, 6242: 1, 6241: 1, 6236: 1, 6211: 1, 6195: 1, 6192: 1, 6170: 1, 6169: 1, 6160: 1, 6106: 1, 6102: 1, 6063: 1, 6022: 1, 6020: 1, 6007: 1, 5996: 1, 5995: 1, 5992: 1, 5984: 1, 5983: 1, 5970: 1, 5928: 1, 5922: 1, 5919: 1, 5913: 1, 5868: 1, 5866: 1, 5859: 1, 5843: 1, 5780: 1, 5772: 1, 5744: 1, 5743: 1, 5731: 1, 5711: 1, 5693: 1, 5684: 1, 5678: 1, 5654: 1, 5652: 1, 5645: 1, 5641: 1, 5636: 1, 5617: 1, 5615: 1, 5607: 1, 5599: 1, 5595: 1, 5585: 1, 5569: 1, 5527: 1, 5521: 1, 5513: 1, 5510: 1, 5504: 1, 5494: 1, 5468: 1, 5411: 1, 5408: 1, 5397: 1, 5390: 1, 5383: 1, 5366: 1, 5349: 1, 5346: 1, 5343: 1, 5342: 1, 5321: 1, 5306: 1, 5283: 1, 5240: 1, 5233: 1, 5220: 1, 5219: 1, 5205: 1, 5204: 1, 5181: 1, 5175: 1, 5159: 1, 5127: 1, 5124: 1, 5120: 1, 5119: 1, 5102: 1, 5072: 1, 5050: 1, 5040: 1, 5035: 1, 5015: 1, 5007: 1, 5006: 1, 5004: 1, 5002: 1, 4987: 1, 4964: 1, 4957: 1, 4927: 1, 4920: 1, 4917: 1, 4906: 1, 4901: 1, 4896: 1, 4872: 1, 4864: 1, 4854: 1, 4834: 1, 4831: 1, 4817: 1, 4790: 1, 4784: 1, 4776: 1, 4775: 1, 4763: 1, 4754: 1, 4753: 1, 4709: 1, 4691: 1, 4687: 1, 4684: 1, 4683: 1, 4672: 1, 4665: 1, 4660: 1, 4640: 1, 4628: 1, 4602: 1, 4598: 1, 4596: 1, 4582: 1, 4528: 1, 4527: 1, 4523: 1, 4522: 1, 4520: 1, 4500: 1, 4486: 1, 4474: 1, 4472: 1, 4471: 1, 4469: 1, 4466: 1, 4464: 1, 4460: 1, 4457: 1, 4435: 1, 4432: 1, 4430: 1, 4429: 1, 4428: 1, 4407: 1, 4403: 1, 4401: 1, 4400: 1, 4396: 1, 4395: 1, 4382: 1, 4378: 1, 4372: 1, 4359: 1, 4353: 1, 4352: 1, 4351: 1, 4336: 1, 4333: 1, 4332: 1, 4319: 1, 4317: 1, 4306: 1, 4294: 1, 4285: 1, 4281: 1, 4270: 1, 4267: 1, 4259: 1, 4250: 1, 4247: 1, 4231: 1, 4228: 1, 4217: 1, 4216: 1, 4215: 1, 4212: 1, 4207: 1, 4198: 1, 4194: 1, 4183: 1, 4180: 1, 4164: 1, 4156: 1, 4155: 1, 4146: 1, 4137: 1, 4135: 1, 4127: 1, 4124: 1, 4122: 1, 4109: 1, 4105: 1, 4104: 1, 4092: 1, 4090: 1, 4076: 1, 4074: 1, 4072: 1, 4069: 1, 4068: 1, 4060: 1, 4056: 1, 4031: 1, 4005: 1, 4001: 1, 3993: 1, 3988: 1, 3983: 1, 3978: 1, 3971: 1, 3964: 1, 3963: 1, 3942: 1, 3940: 1, 3937: 1, 3919: 1, 3914: 1, 3895: 1, 3873: 1, 3871: 1, 3870: 1, 3865: 1, 3851: 1, 3843: 1, 3842: 1, 3833: 1, 3832: 1, 3827: 1, 3814: 1, 3805: 1, 3801: 1, 3798: 1, 3796: 1, 3791: 1, 3790: 1, 3786: 1, 3778: 1, 3777: 1, 3775: 1, 3773: 1, 3771: 1, 3768: 1, 3754: 1, 3753: 1, 3751: 1, 3749: 1, 3744: 1, 3743: 1, 3731: 1, 3726: 1, 3723: 1, 3710: 1, 3702: 1, 3698: 1, 3696: 1, 3694: 1, 3688: 1, 3687: 1, 3685: 1, 3684: 1, 3683: 1, 3682: 1, 3676: 1, 3669: 1, 3662: 1, 3654: 1, 3649: 1, 3643: 1, 3641: 1, 3621: 1, 3605: 1, 3599: 1, 3598: 1, 3589: 1, 3578: 1, 3574: 1, 3572: 1, 3561: 1, 3558: 1, 3555: 1, 3544: 1, 3541: 1, 3527: 1, 3526: 1, 3524: 1, 3523: 1, 3514: 1, 3513: 1, 3512: 1, 3510: 1, 3498: 1, 3495: 1, 3488: 1, 3486: 1, 3477: 1, 3470: 1, 3466: 1, 3464: 1, 3463: 1, 3461: 1, 3459: 1, 3457: 1, 3456: 1, 3453: 1, 3452: 1, 3447: 1, 3444: 1, 3441: 1, 3439: 1, 3426: 1, 3423: 1, 3418: 1, 3407: 1, 3406: 1, 3405: 1, 3404: 1, 3398: 1, 3393: 1, 3388: 1, 3382: 1, 3381: 1, 3380: 1, 3377: 1, 3375: 1, 3372: 1, 3370: 1, 3364: 1, 3358: 1, 3346: 1, 3337: 1, 3331: 1, 3330: 1, 3318: 1, 3307: 1, 3293: 1, 3290: 1, 3288: 1, 3280: 1, 3276: 1, 3271: 1, 3267: 1, 3266: 1, 3264: 1, 3260: 1, 3245: 1, 3243: 1, 3232: 1, 3230: 1, 3228: 1, 3227: 1, 3223: 1, 3215: 1, 3212:

1, 3205: 1, 3201: 1, 3197: 1, 3193: 1, 3182: 1, 3180: 1, 3171: 1, 3170: 1, 31
66: 1, 3152: 1, 3142: 1, 3140: 1, 3138: 1, 3135: 1, 3131: 1, 3127: 1, 3116:
1, 3113: 1, 3107: 1, 3106: 1, 3102: 1, 3100: 1, 3099: 1, 3094: 1, 3092: 1, 30
91: 1, 3089: 1, 3088: 1, 3080: 1, 3078: 1, 3074: 1, 3072: 1, 3067: 1, 3066:
1, 3065: 1, 3063: 1, 3053: 1, 3051: 1, 3050: 1, 3042: 1, 3037: 1, 3035: 1, 30
34: 1, 3031: 1, 3029: 1, 3027: 1, 3026: 1, 3021: 1, 3018: 1, 3012: 1, 3002:
1, 3000: 1, 2986: 1, 2983: 1, 2975: 1, 2969: 1, 2967: 1, 2966: 1, 2961: 1, 29
48: 1, 2939: 1, 2928: 1, 2917: 1, 2904: 1, 2902: 1, 2901: 1, 2891: 1, 2887:
1, 2886: 1, 2879: 1, 2873: 1, 2849: 1, 2848: 1, 2841: 1, 2838: 1, 2834: 1, 28
33: 1, 2832: 1, 2817: 1, 2814: 1, 2811: 1, 2798: 1, 2791: 1, 2785: 1, 2771:
1, 2770: 1, 2762: 1, 2757: 1, 2752: 1, 2749: 1, 2740: 1, 2727: 1, 2721: 1, 27
19: 1, 2713: 1, 2710: 1, 2707: 1, 2700: 1, 2698: 1, 2696: 1, 2693: 1, 2692:
1, 2685: 1, 2678: 1, 2673: 1, 2667: 1, 2652: 1, 2649: 1, 2641: 1, 2636: 1, 26
30: 1, 2628: 1, 2623: 1, 2616: 1, 2612: 1, 2611: 1, 2607: 1, 2604: 1, 2599:
1, 2595: 1, 2593: 1, 2583: 1, 2578: 1, 2574: 1, 2573: 1, 2571: 1, 2569: 1, 25
68: 1, 2554: 1, 2552: 1, 2551: 1, 2542: 1, 2540: 1, 2539: 1, 2532: 1, 2530:
1, 2529: 1, 2528: 1, 2522: 1, 2507: 1, 2503: 1, 2500: 1, 2492: 1, 2477: 1, 24
72: 1, 2459: 1, 2451: 1, 2450: 1, 2448: 1, 2447: 1, 2446: 1, 2445: 1, 2443:
1, 2435: 1, 2431: 1, 2424: 1, 2417: 1, 2410: 1, 2403: 1, 2401: 1, 2399: 1, 23
95: 1, 2385: 1, 2381: 1, 2378: 1, 2376: 1, 2373: 1, 2370: 1, 2368: 1, 2367:
1, 2365: 1, 2363: 1, 2356: 1, 2354: 1, 2351: 1, 2341: 1, 2340: 1, 2337: 1, 23
33: 1, 2327: 1, 2325: 1, 2324: 1, 2321: 1, 2320: 1, 2315: 1, 2310: 1, 2307:
1, 2303: 1, 2302: 1, 2301: 1, 2293: 1, 2291: 1, 2290: 1, 2289: 1, 2282: 1, 22
81: 1, 2278: 1, 2267: 1, 2266: 1, 2258: 1, 2246: 1, 2244: 1, 2238: 1, 2235:
1, 2233: 1, 2226: 1, 2217: 1, 2214: 1, 2213: 1, 2211: 1, 2208: 1, 2187: 1, 21
85: 1, 2182: 1, 2181: 1, 2177: 1, 2173: 1, 2169: 1, 2168: 1, 2166: 1, 2165:
1, 2162: 1, 2158: 1, 2156: 1, 2155: 1, 2149: 1, 2148: 1, 2142: 1, 2140: 1, 21
37: 1, 2134: 1, 2132: 1, 2130: 1, 2129: 1, 2128: 1, 2126: 1, 2121: 1, 2115:
1, 2112: 1, 2110: 1, 2109: 1, 2107: 1, 2103: 1, 2101: 1, 2099: 1, 2098: 1, 20
90: 1, 2087: 1, 2084: 1, 2080: 1, 2079: 1, 2076: 1, 2071: 1, 2067: 1, 2064:
1, 2063: 1, 2061: 1, 2056: 1, 2053: 1, 2052: 1, 2051: 1, 2038: 1, 2037: 1, 20
36: 1, 2034: 1, 2032: 1, 2029: 1, 2025: 1, 2023: 1, 2020: 1, 2013: 1, 2005:
1, 2003: 1, 1996: 1, 1995: 1, 1993: 1, 1992: 1, 1991: 1, 1990: 1, 1988: 1, 19
87: 1, 1983: 1, 1980: 1, 1979: 1, 1971: 1, 1970: 1, 1969: 1, 1965: 1, 1962:
1, 1961: 1, 1960: 1, 1956: 1, 1955: 1, 1954: 1, 1952: 1, 1949: 1, 1942: 1, 19
39: 1, 1938: 1, 1931: 1, 1930: 1, 1927: 1, 1926: 1, 1921: 1, 1919: 1, 1918:
1, 1909: 1, 1907: 1, 1905: 1, 1900: 1, 1899: 1, 1894: 1, 1893: 1, 1892: 1, 18
90: 1, 1885: 1, 1880: 1, 1879: 1, 1878: 1, 1876: 1, 1875: 1, 1869: 1, 1868:
1, 1867: 1, 1864: 1, 1863: 1, 1862: 1, 1858: 1, 1857: 1, 1856: 1, 1854: 1, 18
52: 1, 1850: 1, 1845: 1, 1836: 1, 1835: 1, 1833: 1, 1832: 1, 1831: 1, 1830:
1, 1826: 1, 1825: 1, 1823: 1, 1822: 1, 1820: 1, 1819: 1, 1818: 1, 1817: 1, 18
16: 1, 1813: 1, 1807: 1, 1802: 1, 1800: 1, 1797: 1, 1790: 1, 1789: 1, 1788:
1, 1785: 1, 1779: 1, 1778: 1, 1773: 1, 1771: 1, 1767: 1, 1765: 1, 1762: 1, 17
61: 1, 1759: 1, 1758: 1, 1757: 1, 1756: 1, 1755: 1, 1748: 1, 1747: 1, 1739:
1, 1738: 1, 1736: 1, 1734: 1, 1733: 1, 1731: 1, 1730: 1, 1726: 1, 1724: 1, 17
21: 1, 1709: 1, 1706: 1, 1699: 1, 1698: 1, 1697: 1, 1692: 1, 1690: 1, 1686:
1, 1685: 1, 1683: 1, 1682: 1, 1677: 1, 1673: 1, 1671: 1, 1668: 1, 1665: 1, 16
62: 1, 1654: 1, 1653: 1, 1648: 1, 1646: 1, 1645: 1, 1643: 1, 1642: 1, 1639:
1, 1629: 1, 1628: 1, 1627: 1, 1626: 1, 1623: 1, 1622: 1, 1618: 1, 1612: 1, 16
09: 1, 1605: 1, 1603: 1, 1600: 1, 1599: 1, 1595: 1, 1591: 1, 1590: 1, 1589:
1, 1583: 1, 1579: 1, 1577: 1, 1571: 1, 1569: 1, 1568: 1, 1566: 1, 1565: 1, 15
64: 1, 1560: 1, 1559: 1, 1555: 1, 1554: 1, 1553: 1, 1551: 1, 1550: 1, 1547:
1, 1546: 1, 1545: 1, 1544: 1, 1534: 1, 1532: 1, 1529: 1, 1528: 1, 1527: 1, 15
25: 1, 1522: 1, 1521: 1, 1518: 1, 1517: 1, 1516: 1, 1512: 1, 1511: 1, 1509:
1, 1507: 1, 1501: 1, 1498: 1, 1494: 1, 1490: 1, 1488: 1, 1486: 1, 1482: 1, 14
77: 1, 1472: 1, 1469: 1, 1463: 1, 1452: 1, 1447: 1, 1446: 1, 1444: 1, 1443:
1, 1439: 1, 1430: 1, 1424: 1, 1422: 1, 1420: 1, 1418: 1, 1412: 1, 1409: 1, 14

07: 1, 1404: 1, 1401: 1, 1400: 1, 1399: 1, 1398: 1, 1396: 1, 1389: 1, 1388: 1, 1386: 1, 1385: 1, 1384: 1, 1383: 1, 1380: 1, 1370: 1, 1368: 1, 1367: 1, 136: 1, 1365: 1, 1363: 1, 1360: 1, 1355: 1, 1351: 1, 1346: 1, 1342: 1, 1341: 1, 1338: 1, 1337: 1, 1336: 1, 1335: 1, 1333: 1, 1332: 1, 1328: 1, 1324: 1, 1323: 1, 1319: 1, 1317: 1, 1313: 1, 1311: 1, 1310: 1, 1306: 1, 1304: 1, 1301: 1, 1300: 1, 1290: 1, 1287: 1, 1286: 1, 1284: 1, 1271: 1, 1267: 1, 1266: 1, 1264: 1, 1263: 1, 1258: 1, 1251: 1, 1246: 1, 1244: 1, 1243: 1, 1239: 1, 1238: 1, 1235: 1, 1232: 1, 1231: 1, 1228: 1, 1222: 1, 1220: 1, 1219: 1, 1217: 1, 1210: 1, 1208: 1, 1207: 1, 1206: 1, 1203: 1, 1202: 1, 1196: 1, 1192: 1, 1191: 1, 1187: 1, 1185: 1, 1184: 1, 1179: 1, 1176: 1, 1173: 1, 1172: 1, 1167: 1, 1165: 1, 1161: 1, 1160: 1, 1159: 1, 1158: 1, 1154: 1, 1153: 1, 1146: 1, 1137: 1, 1136: 1, 1135: 1, 1132: 1, 1131: 1, 1129: 1, 1126: 1, 1123: 1, 1120: 1, 1119: 1, 1117: 1, 1112: 1, 1107: 1, 1104: 1, 1102: 1, 1101: 1, 1100: 1, 1099: 1, 1098: 1, 1097: 1, 1090: 1, 1087: 1, 1085: 1, 1084: 1, 1083: 1, 1078: 1, 1072: 1, 1069: 1, 1064: 1, 1061: 1, 1058: 1, 1057: 1, 1051: 1, 1047: 1, 1046: 1, 1043: 1, 1042: 1, 1038: 1, 1036: 1, 1035: 1, 1030: 1, 1024: 1, 1016: 1, 1015: 1, 1014: 1, 1011: 1, 1010: 1, 1007: 1, 1006: 1, 1003: 1, 1002: 1, 999: 1, 996: 1, 995: 1, 991: 1, 990: 1, 989: 1, 983: 1, 977: 1, 976: 1, 975: 1, 973: 1, 971: 1, 970: 1, 968: 1, 967: 1, 965: 1, 960: 1, 959: 1, 951: 1, 950: 1, 944: 1, 943: 1, 942: 1, 941: 1, 940: 1, 929: 1, 927: 1, 926: 1, 924: 1, 922: 1, 920: 1, 919: 1, 918: 1, 914: 1, 911: 1, 910: 1, 908: 1, 899: 1, 898: 1, 894: 1, 890: 1, 885: 1, 884: 1, 883: 1, 882: 1, 880: 1, 879: 1, 875: 1, 874: 1, 869: 1, 868: 1, 864: 1, 860: 1, 857: 1, 856: 1, 853: 1, 836: 1, 830: 1, 823: 1, 820: 1, 818: 1, 815: 1, 809: 1, 808: 1, 791: 1, 787: 1, 785: 1, 774: 1, 773: 1, 771: 1, 765: 1, 764: 1, 763: 1, 754: 1, 748: 1, 747: 1, 741: 1, 739: 1, 738: 1, 734: 1, 733: 1, 730: 1, 722: 1, 716: 1, 702: 1, 692: 1, 688: 1, 678: 1, 663: 1, 662: 1, 652: 1, 651: 1, 640: 1, 629: 1, 625: 1, 615: 1, 613: 1, 602: 1, 600: 1, 591: 1, 571: 1, 554: 1, 535: 1, 521: 1, 493: 1, 445: 1})

```
In [59]: # Train a Logistic regression+Calibration model using text features which are
# on-hot encoded
alpha = [10 ** x for x in range(-5, 1)]

# read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generate
# d/sklearn.Linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_i
# ntercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_
# rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochast
# ic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link:
#-----
```

cv_log_error_array=[]
for i in alpha:
 clf = SGDClassifier(alpha=i, penalty='l2', loss='log', random_state=42)
 clf.fit(train_text_feature_onehotCoding, y_train)

 sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
 sig_clf.fit(train_text_feature_onehotCoding, y_train)
 predict_y = sig_clf.predict_proba(cv_text_feature_onehotCoding)
 cv_log_error_array.append(log_loss(y_cv, predict_y, labels=clf.classes_, e
ps=1e-15))
 print('For values of alpha = ', i, "The log loss is:", log_loss(y_cv, predi
ct_y, labels=clf.classes_, eps=1e-15))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
 ax.annotate((alpha[i],np.round(txt,3)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

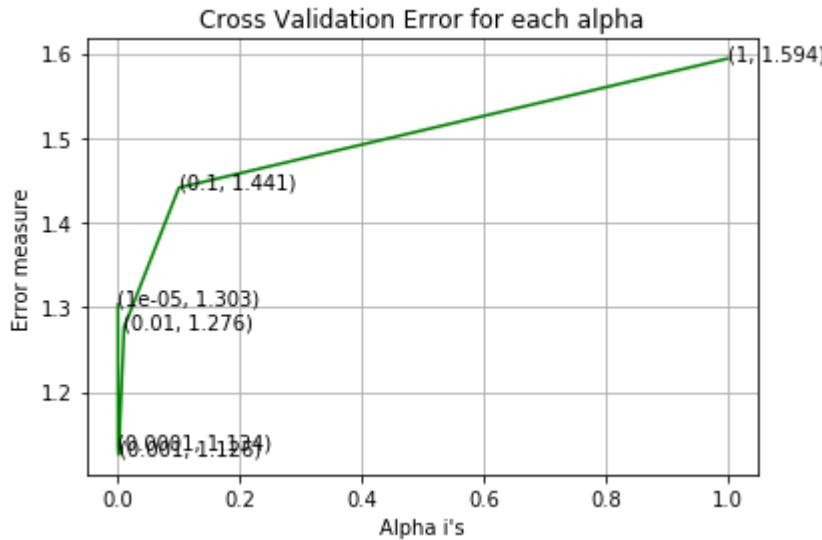
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_
state=42)
clf.fit(train_text_feature_onehotCoding, y_train)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_text_feature_onehotCoding, y_train)

predict_y = sig_clf.predict_proba(train_text_feature_onehotCoding)

```

print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_text_feature_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_text_feature_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
For values of alpha = 1e-05 The log loss is: 1.3032575425463053
For values of alpha = 0.0001 The log loss is: 1.1342166681568358
For values of alpha = 0.001 The log loss is: 1.1260233861479303
For values of alpha = 0.01 The log loss is: 1.2763762931001126
For values of alpha = 0.1 The log loss is: 1.4414277435506773
For values of alpha = 1 The log loss is: 1.593865199933301

```



```

For values of best alpha = 0.001 The train log loss is: 0.7136662535602447
For values of best alpha = 0.001 The cross validation log loss is: 1.1260233861479303
For values of best alpha = 0.001 The test log loss is: 1.178608872613285

```

Q. Is the Text feature stable across all the data sets (Test, Train, Cross validation)?

Ans. Yes, it seems like!

```

In [60]: def get_intersec_text(df):
    df_text_vec = CountVectorizer(min_df=3)
    df_text_fea = df_text_vec.fit_transform(df['TEXT'])
    df_text_features = df_text_vec.get_feature_names()

    df_text_fea_counts = df_text_fea.sum(axis=0).A1
    df_text_fea_dict = dict(zip(list(df_text_features),df_text_fea_counts))
    len1 = len(set(df_text_features))
    len2 = len(set(train_text_features) & set(df_text_features))
    return len1,len2

```

```
In [61]: len1,len2 = get_intersec_text(test_df)
print(np.round((len2/len1)*100, 3), "% of word of test data appeared in train
      data")
len1,len2 = get_intersec_text(cv_df)
print(np.round((len2/len1)*100, 3), "% of word of Cross Validation appeared in
      train data")
```

97.02 % of word of test data appeared in train data
 98.619 % of word of Cross Validation appeared in train data

```
In [1]: import dill
# dill.dump_session('notebook_env.db')
dill.load_session('notebook_env.db')
```

4. Machine Learning Models

```
In [2]: #Data preparation for ML models.
```

#Misc. functionns for ML models

```
def predict_and_plot_confusion_matrix(train_x, train_y,test_x, test_y, clf):
    clf.fit(train_x, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x, train_y)
    pred_y = sig_clf.predict(test_x)

    # for calculating log_loss we will provide the array of probabilities belongs to each class
    print("Log loss :",log_loss(test_y, sig_clf.predict_proba(test_x)))
    # calculating the number of data points that are misclassified
    print("Number of mis-classified points :", np.count_nonzero((pred_y- test_y))/test_y.shape[0])
    plot_confusion_matrix(test_y, pred_y)
```

```
In [3]: def report_log_loss(train_x, train_y, test_x, test_y, clf):
    clf.fit(train_x, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x, train_y)
    sig_clf_probs = sig_clf.predict_proba(test_x)
    return log_loss(test_y, sig_clf_probs, eps=1e-15)
```

```
In [4]: # this function will be used just for naive bayes
# for the given indices, we will print the name of the features
# and we will check whether the feature present in the test point text or not
def get_impfeature_names(indices, text, gene, var, no_features):
    gene_count_vec = CountVectorizer()
    var_count_vec = CountVectorizer()
    text_count_vec = CountVectorizer(min_df=3)

    gene_vec = gene_count_vec.fit(train_df['Gene'])
    var_vec = var_count_vec.fit(train_df['Variation'])
    text_vec = text_count_vec.fit(train_df['TEXT'])

    fea1_len = len(gene_vec.get_feature_names())
    fea2_len = len(var_count_vec.get_feature_names())

    word_present = 0
    for i,v in enumerate(indices):
        if (v < fea1_len):
            word = gene_vec.get_feature_names()[v]
            yes_no = True if word == gene else False
            if yes_no:
                word_present += 1
                print(i, "Gene feature [{}] present in test data point [{}].format(word,yes_no))
            elif (v < fea1_len+fea2_len):
                word = var_vec.get_feature_names()[v-(fea1_len)]
                yes_no = True if word == var else False
                if yes_no:
                    word_present += 1
                    print(i, "variation feature [{}] present in test data point [{}].format(word,yes_no))
            else:
                word = text_vec.get_feature_names()[v-(fea1_len+fea2_len)]
                yes_no = True if word in text.split() else False
                if yes_no:
                    word_present += 1
                    print(i, "Text feature [{}] present in test data point [{}].format(word,yes_no))

    print("Out of the top ",no_features," features ", word_present, "are present in query point")
```

Stacking the three types of features

In [6]: train_gene_var_onehotCoding.shape

Out[6]: (2124, 2194)

In [7]: train_text_feature_onehotCoding.shape

Out[7]: (2124, 53213)

In [5]: # merging gene, variance and text features

```
# building train, test and cross validation data sets
# a = [[1, 2],
#       [3, 4]]
# b = [[4, 5],
#       [6, 7]]
# hstack(a, b) = [[1, 2, 4, 5],
#                  [3, 4, 6, 7]]

train_gene_var_onehotCoding = hstack((train_gene_feature_onehotCoding,train_variation_feature_onehotCoding))
test_gene_var_onehotCoding = hstack((test_gene_feature_onehotCoding,test_variation_feature_onehotCoding))
cv_gene_var_onehotCoding = hstack((cv_gene_feature_onehotCoding,cv_variation_feature_onehotCoding))

train_x_onehotCoding = hstack((train_gene_var_onehotCoding, train_text_feature_onehotCoding)).tocsr()

test_x_onehotCoding = hstack((test_gene_var_onehotCoding, test_text_feature_onehotCoding)).tocsr()

cv_x_onehotCoding = hstack((cv_gene_var_onehotCoding, cv_text_feature_onehotCoding)).tocsr()

train_gene_var_responseCoding = np.hstack((train_gene_feature_responseCoding,train_variation_feature_responseCoding))
test_gene_var_responseCoding = np.hstack((test_gene_feature_responseCoding,test_variation_feature_responseCoding))
cv_gene_var_responseCoding = np.hstack((cv_gene_feature_responseCoding,cv_variation_feature_responseCoding))

train_x_responseCoding = np.hstack((train_gene_var_responseCoding, train_text_feature_responseCoding))
test_x_responseCoding = np.hstack((test_gene_var_responseCoding, test_text_feature_responseCoding))
cv_x_responseCoding = np.hstack((cv_gene_var_responseCoding, cv_text_feature_responseCoding))

train_y = np.array(list(train_df['Class']))
test_y = np.array(list(test_df['Class']))
cv_y = np.array(list(cv_df['Class']))
```

```
In [6]: print("One hot encoding features :")
print("(number of data points * number of features) in train data = ", train_x_
_onehotCoding.shape)
print("(number of data points * number of features) in test data = ", test_x_o
nehotCoding.shape)
print("(number of data points * number of features) in cross validation data
 =", cv_x_onehotCoding.shape)
```

One hot encoding features :
(number of data points * number of features) in train data = (2124, 55407)
(number of data points * number of features) in test data = (665, 55407)
(number of data points * number of features) in cross validation data = (532,
55407)

```
In [7]: print(" Response encoding features :")
print("(number of data points * number of features) in train data = ", train_x_
_responseCoding.shape)
print("(number of data points * number of features) in test data = ", test_x_r
esponseCoding.shape)
print("(number of data points * number of features) in cross validation data
 =", cv_x_responseCoding.shape)
```

Response encoding features :
(number of data points * number of features) in train data = (2124, 27)
(number of data points * number of features) in test data = (665, 27)
(number of data points * number of features) in cross validation data = (532,
27)

4.1. Base Line Model

4.1.1. Naive Bayes

4.1.1.1. Hyper parameter tuning

```
In [10]: # find more about Multinomial Naive base function here http://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html
# -----
# default paramters
# sklearn.naive_bayes.MultinomialNB(alpha=1.0, fit_prior=True, class_prior=None)

# some of methods of MultinomialNB()
# fit(X, y[, sample_weight]) Fit Naive Bayes classifier according to X, y
# predict(X) Perform classification on an array of test vectors X.
# predict_log_proba(X) Return log-probability estimates for the test vector X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/naive-bayes-algorithm-1/
# -----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/naive-bayes-algorithm-1/
# -----


alpha = [0.00001, 0.0001, 0.001, 0.1, 1, 10, 100, 1000]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = MultinomialNB(alpha=i)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
        # to avoid rounding error while multiplying probabilites we use log-probability estimates
    print("Log Loss :", log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(np.log10(alpha), cv_log_error_array, c='g')
for i, txt in enumerate(np.round(cv_log_error_array, 3)):
    ax.annotate((alpha[i], str(txt)), (np.log10(alpha[i]), cv_log_error_array[i]))
```

```
plt.grid()
plt.xticks(np.log10(alpha))
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

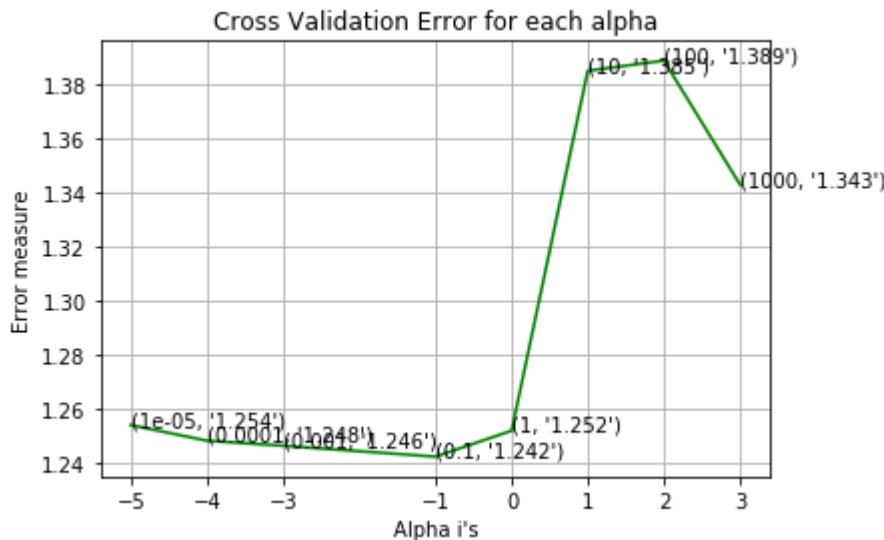
best_alpha = np.argmin(cv_log_error_array)
clf = MultinomialNB(alpha=alpha[best_alpha])
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 1e-05
Log Loss : 1.2540355269224528
for alpha = 0.0001
Log Loss : 1.2482836179707815
for alpha = 0.001
Log Loss : 1.2464012863448808
for alpha = 0.1
Log Loss : 1.2423846899704123
for alpha = 1
Log Loss : 1.2519462059241921
for alpha = 10
Log Loss : 1.3848378778251498
for alpha = 100
Log Loss : 1.3885932868429458
for alpha = 1000
Log Loss : 1.3427493104805197

```



For values of best alpha = 0.1 The train log loss is: 0.9202688483234036
 For values of best alpha = 0.1 The cross validation log loss is: 1.242384689704123
 For values of best alpha = 0.1 The test log loss is: 1.3068184670931153

4.1.1.2. Testing the model with best hyper paramters

```
In [11]: # find more about Multinomial Naive base function here http://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html
# -----
# default paramters
# sklearn.naive_bayes.MultinomialNB(alpha=1.0, fit_prior=True, class_prior=None)

# some of methods of MultinomialNB()
# fit(X, y[, sample_weight]) Fit Naive Bayes classifier according to X, y
# predict(X) Perform classification on an array of test vectors X.
# predict_log_proba(X) Return log-probability estimates for the test vector X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/naive-bayes-algorithm-1/
# -----

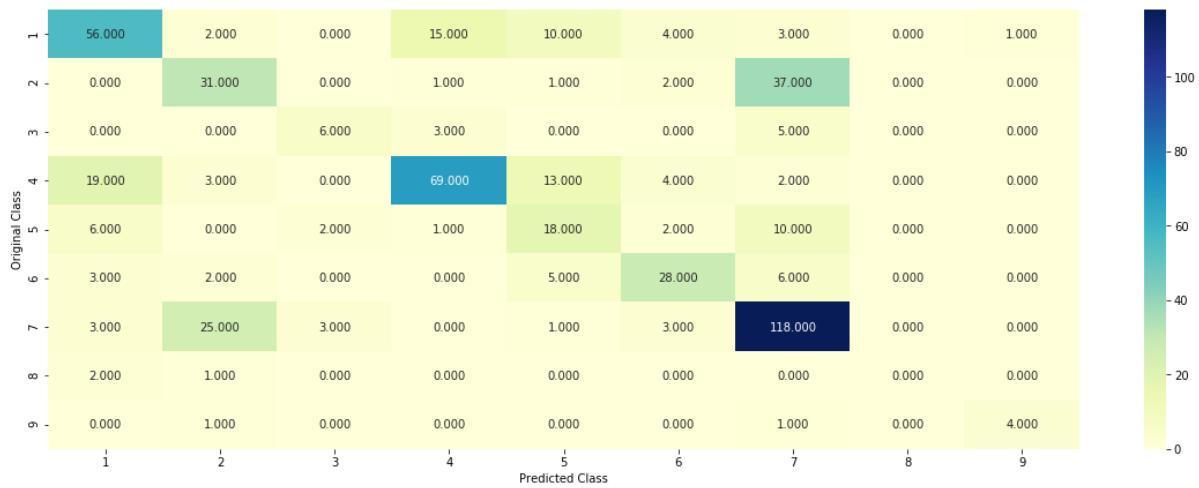

# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
# -----


clf = MultinomialNB(alpha=alpha[best_alpha])
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)
sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
# to avoid rounding error while multiplying probabilit
y estimates
print("Log Loss :",log_loss(cv_y, sig_clf_probs))
print("Number of missclassified point :", np.count_nonzero((sig_clf.predict(cv_x_onehotCoding)- cv_y))/cv_y.shape[0])
plot_confusion_matrix(cv_y, sig_clf.predict(cv_x_onehotCoding.toarray()))
```

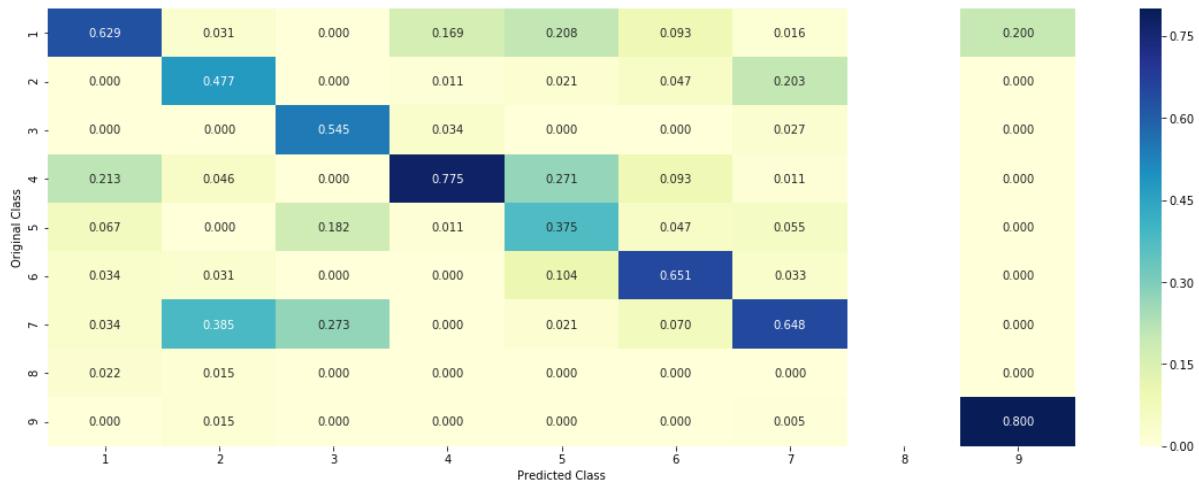
Log Loss : 1.2423846899704123

Number of missclassified point : 0.37969924812030076

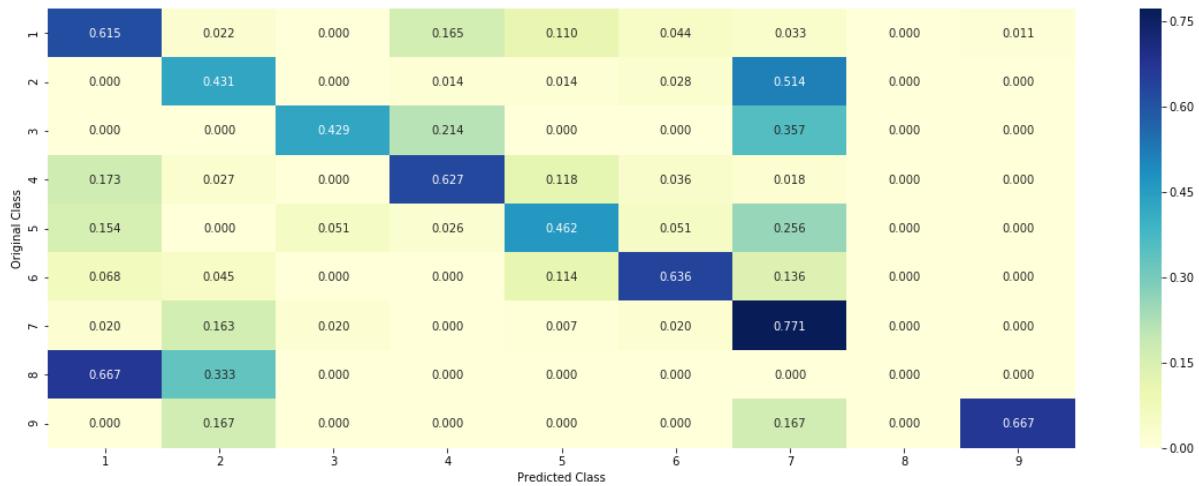
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.1.1.3. Feature Importance, Correctly classified point

```
In [13]: test_point_index = 1
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.coef_)[predicted_cls-1][:,:no_feature]
print("-"*50)
get_impfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

Predicted Class : 4
Predicted Class Probabilities: [[0.0779 0.0748 0.0141 0.6494 0.0341 0.0299 0.
1128 0.0039 0.0032]]
Actual Class : 4

13 Text feature [function] present in test data point [True]
14 Text feature [protein] present in test data point [True]
15 Text feature [mammalian] present in test data point [True]
16 Text feature [proteins] present in test data point [True]
17 Text feature [missense] present in test data point [True]
18 Text feature [retained] present in test data point [True]
20 Text feature [experiments] present in test data point [True]
21 Text feature [suppressor] present in test data point [True]
22 Text feature [activity] present in test data point [True]
24 Text feature [phosphatase] present in test data point [True]
26 Text feature [functional] present in test data point [True]
28 Text feature [pten] present in test data point [True]
29 Text feature [phosphoinositide] present in test data point [True]
31 Text feature [results] present in test data point [True]
34 Text feature [related] present in test data point [True]
36 Text feature [determined] present in test data point [True]
37 Text feature [critical] present in test data point [True]
38 Text feature [whether] present in test data point [True]
39 Text feature [whereas] present in test data point [True]
40 Text feature [transfected] present in test data point [True]
41 Text feature [hamartoma] present in test data point [True]
46 Text feature [ability] present in test data point [True]
48 Text feature [germline] present in test data point [True]
50 Text feature [type] present in test data point [True]
51 Text feature [catalytic] present in test data point [True]
54 Text feature [indicate] present in test data point [True]
55 Text feature [indicated] present in test data point [True]
56 Text feature [tensin] present in test data point [True]
66 Text feature [important] present in test data point [True]
68 Text feature [unstable] present in test data point [True]
70 Text feature [vivo] present in test data point [True]
75 Text feature [purified] present in test data point [True]
77 Text feature [stability] present in test data point [True]
81 Text feature [two] present in test data point [True]
82 Text feature [loss] present in test data point [True]
83 Text feature [affect] present in test data point [True]
86 Text feature [presented] present in test data point [True]
87 Text feature [yeast] present in test data point [True]
88 Text feature [wild] present in test data point [True]
90 Text feature [competent] present in test data point [True]
95 Text feature [incubated] present in test data point [True]
96 Text feature [either] present in test data point [True]
98 Text feature [although] present in test data point [True]
Out of the top 100 features 43 are present in query point

```
In [21]: indices
```

```
Out[21]: array([[ 176,   229,    46,     28,   232,   235,   547,    36,   216,
   165,   200,   144,   210, 23572, 41933, 33241, 41946, 34687,
 44736, 8458, 21877, 49244, 8561, 39030, 40042, 9551, 23578,
 8264, 42225, 40089, 13376, 44708, 45605, 18750, 44270, 25918,
18582, 16754, 54346, 54342, 51414, 26014, 12926, 18581, 46528,
26242, 8216, 51416, 24601, 53156, 52099, 13933, 12066, 18339,
28660, 28661, 50309, 33099,   231,    136,    138,    115,     81,
16475, 50336, 34824, 28465, 15411, 52815, 42409, 53836, 36023,
51708, 50846, 24473, 42375, 18629, 48361, 16225, 16134, 54151,
52076, 32475, 8867, 22156, 46290, 41417, 55148, 54409, 47204,
15816, 40057, 33964, 37375, 28911, 28606, 20616, 5084,  9468,
42938]], dtype=int64)
```

4.1.1.4. Feature Importance, Incorrectly classified point

```
In [22]: test_point_index = 100
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.coef_)[predicted_cls-1][:,:no_feature]
print("-"*50)
get_impfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

Predicted Class : 7
Predicted Class Probabilities: [[0.1219 0.0833 0.0157 0.105 0.0377 0.0332 0.
5954 0.0043 0.0035]]
Actual Class : 1

17 Text feature [kinase] present in test data point [True]
18 Text feature [presence] present in test data point [True]
19 Text feature [activating] present in test data point [True]
20 Text feature [inhibitor] present in test data point [True]
21 Text feature [downstream] present in test data point [True]
22 Text feature [independent] present in test data point [True]
23 Text feature [well] present in test data point [True]
24 Text feature [compared] present in test data point [True]
25 Text feature [recently] present in test data point [True]
26 Text feature [previously] present in test data point [True]
27 Text feature [contrast] present in test data point [True]
28 Text feature [potential] present in test data point [True]
29 Text feature [expressing] present in test data point [True]
30 Text feature [cell] present in test data point [True]
31 Text feature [growth] present in test data point [True]
32 Text feature [shown] present in test data point [True]
33 Text feature [treated] present in test data point [True]
34 Text feature [also] present in test data point [True]
35 Text feature [found] present in test data point [True]
36 Text feature [addition] present in test data point [True]
37 Text feature [cells] present in test data point [True]
38 Text feature [activation] present in test data point [True]
39 Text feature [suggest] present in test data point [True]
40 Text feature [inhibition] present in test data point [True]
41 Text feature [however] present in test data point [True]
42 Text feature [10] present in test data point [True]
44 Text feature [1a] present in test data point [True]
45 Text feature [showed] present in test data point [True]
47 Text feature [similar] present in test data point [True]
48 Text feature [total] present in test data point [True]
49 Text feature [described] present in test data point [True]
50 Text feature [may] present in test data point [True]
51 Text feature [studies] present in test data point [True]
52 Text feature [factor] present in test data point [True]
53 Text feature [obtained] present in test data point [True]
54 Text feature [mutations] present in test data point [True]
55 Text feature [identified] present in test data point [True]
56 Text feature [interestingly] present in test data point [True]
57 Text feature [report] present in test data point [True]
58 Text feature [using] present in test data point [True]
59 Text feature [new] present in test data point [True]
60 Text feature [including] present in test data point [True]
61 Text feature [inhibitors] present in test data point [True]
62 Text feature [without] present in test data point [True]
63 Text feature [observed] present in test data point [True]
64 Text feature [reported] present in test data point [True]
65 Text feature [enhanced] present in test data point [True]
66 Text feature [12] present in test data point [True]
67 Text feature [approximately] present in test data point [True]
68 Text feature [1b] present in test data point [True]
69 Text feature [small] present in test data point [True]
70 Text feature [3b] present in test data point [True]

```
71 Text feature [although] present in test data point [True]
72 Text feature [figure] present in test data point [True]
74 Text feature [respectively] present in test data point [True]
75 Text feature [followed] present in test data point [True]
76 Text feature [various] present in test data point [True]
77 Text feature [15] present in test data point [True]
78 Text feature [phosphorylation] present in test data point [True]
79 Text feature [confirmed] present in test data point [True]
80 Text feature [occur] present in test data point [True]
81 Text feature [consistent] present in test data point [True]
83 Text feature [approved] present in test data point [True]
84 Text feature [three] present in test data point [True]
86 Text feature [due] present in test data point [True]
87 Text feature [either] present in test data point [True]
88 Text feature [proliferation] present in test data point [True]
89 Text feature [suggests] present in test data point [True]
90 Text feature [measured] present in test data point [True]
91 Text feature [two] present in test data point [True]
95 Text feature [3a] present in test data point [True]
96 Text feature [show] present in test data point [True]
97 Text feature [therapeutic] present in test data point [True]
98 Text feature [demonstrated] present in test data point [True]
99 Text feature [furthermore] present in test data point [True]
```

Out of the top 100 features 75 are present in query point

4.2. K Nearest Neighbour Classification

4.2.1. Hyper parameter tuning

```
In [23]: # find more about KNeighborsClassifier() here http://scikit-Learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html
# -----
# default parameter
# KNeighborsClassifier(n_neighbors=5, weights='uniform', algorithm='auto', leaf_size=30, p=2,
# metric='minkowski', metric_params=None, n_jobs=1, **kwargs)

# methods of
# fit(X, y) : Fit the model using X as training data and y as target values
# predict(X):Predict the class labels for the provided data
# predict_proba(X):Return probability estimates for the test data X.
#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/k-nearest-neighbors-geometric-intuition-with-a-toy-example-1/
#-----

# find more about CalibratedClassifierCV here at http://scikit-Learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
#-----
# video link:
#-----
```

```
alpha = [5, 11, 15, 21, 31, 41, 51, 99]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = KNeighborsClassifier(n_neighbors=i)
    clf.fit(train_x_responseCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_responseCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_responseCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
    # to avoid rounding error while multiplying probabilites we use log-probability estimates
    print("Log Loss :",log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
    ax.annotate((alpha[i],str(txt)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
```

```
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

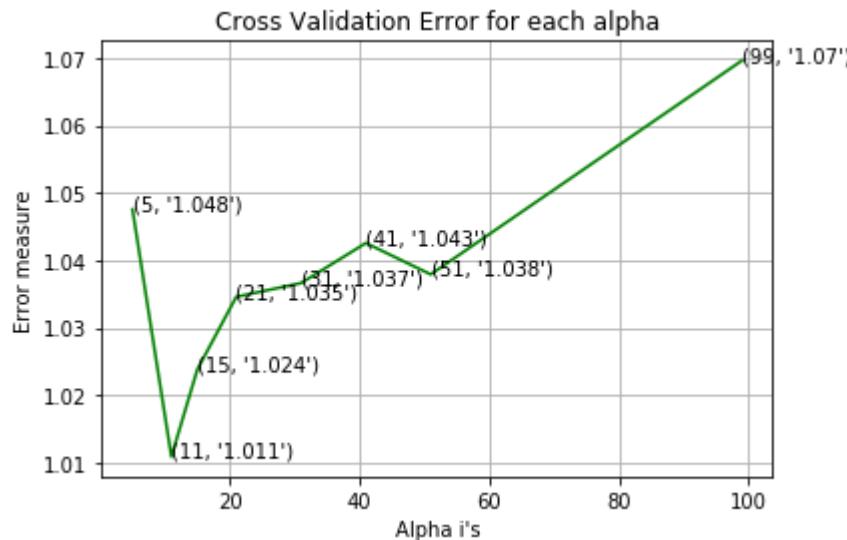
best_alpha = np.argmin(cv_log_error_array)
clf = KNeighborsClassifier(n_neighbors=alpha[best_alpha])
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_responseCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_responseCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_responseCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 5
Log Loss : 1.0475475703145196
for alpha = 11
Log Loss : 1.0109009426519564
for alpha = 15
Log Loss : 1.0238128217601776
for alpha = 21
Log Loss : 1.0345791946880367
for alpha = 31
Log Loss : 1.0366171768925598
for alpha = 41
Log Loss : 1.0425559330633891
for alpha = 51
Log Loss : 1.0379068777214795
for alpha = 99
Log Loss : 1.0695904061224548

```



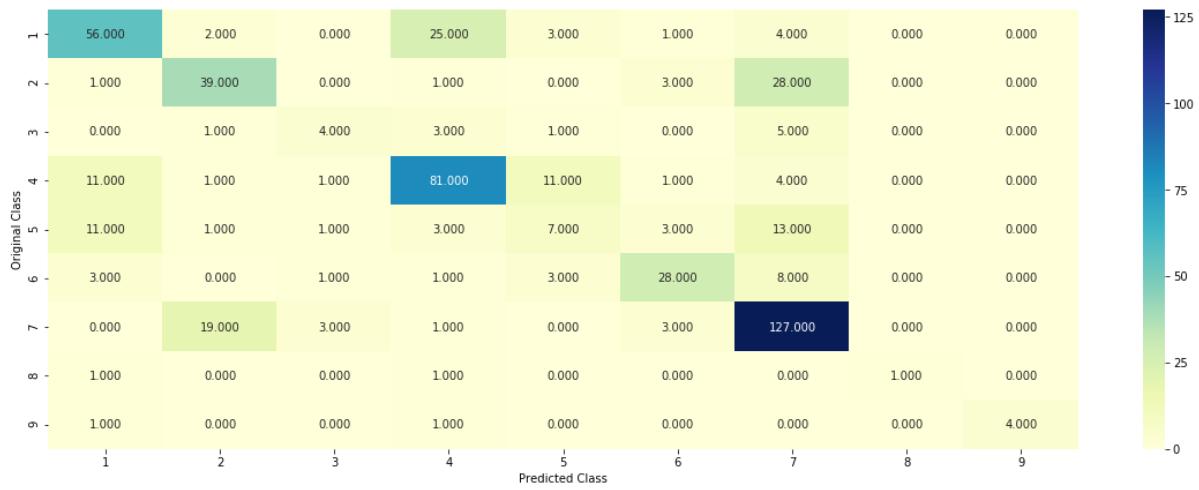
For values of best alpha = 11 The train log loss is: 0.6391579061604918
 For values of best alpha = 11 The cross validation log loss is: 1.0109009426519564
 For values of best alpha = 11 The test log loss is: 1.0614495556346983

4.2.2. Testing the model with best hyper parameters

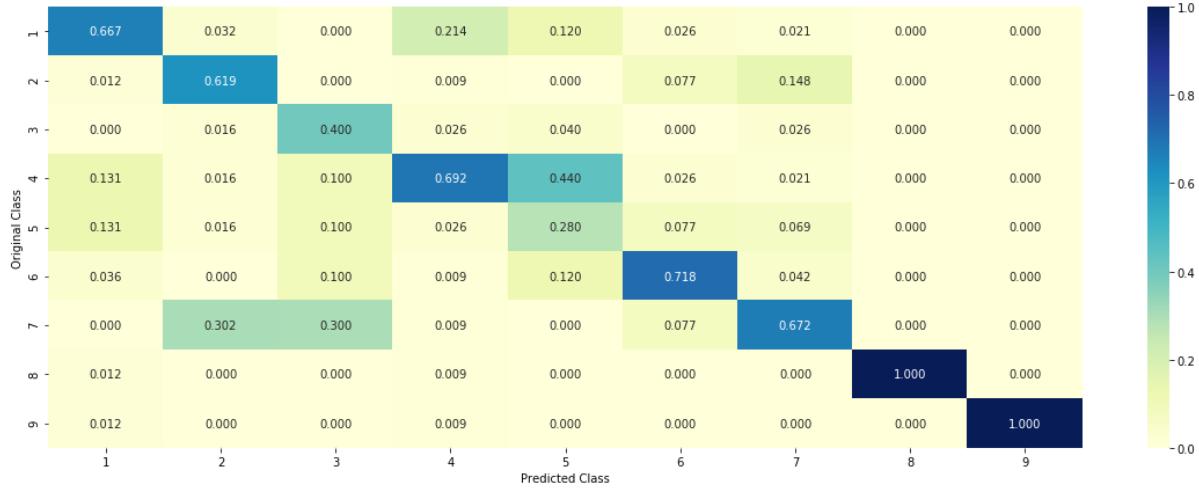
```
In [24]: # find more about KNeighborsClassifier() here http://scikit-Learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html
# -----
# default parameter
# KNeighborsClassifier(n_neighbors=5, weights='uniform', algorithm='auto', leaf_size=30, p=2,
# metric='minkowski', metric_params=None, n_jobs=1, **kwargs)

# methods of
# fit(X, y) : Fit the model using X as training data and y as target values
# predict(X):Predict the class labels for the provided data
# predict_proba(X):Return probability estimates for the test data X.
#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/k-nearest-neighbors-geometric-intuition-with-a-toy-example-1/
#-----
clf = KNeighborsClassifier(n_neighbors=alpha[best_alpha])
predict_and_plot_confusion_matrix(train_x_responseCoding, train_y, cv_x_responseCoding, cv_y, clf)
```

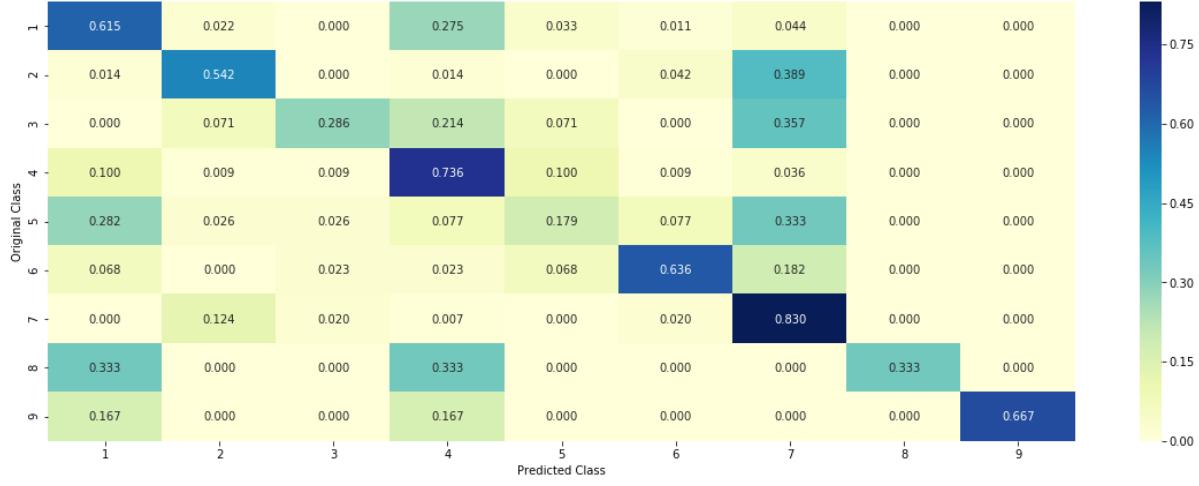
Log loss : 1.0109009426519564
 Number of mis-classified points : 0.34774436090225563
 ----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.2.3.Sample Query point -1

```
In [66]: clf = KNeighborsClassifier(n_neighbors=alpha[best_alpha])
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

test_point_index = 0
predicted_cls = sig_clf.predict(test_x_responseCoding[test_point_index].reshape(1,-1))
print("Predicted Class :", predicted_cls[0])
print("Actual Class :", test_y[test_point_index])
neighbors = clf.kneighbors(test_x_responseCoding[test_point_index].reshape(1,-1), alpha[best_alpha])
print("The ",alpha[best_alpha]," nearest neighbours of the test points belongs to classes",train_y[neighbors[1][0]])
print("Frequency of nearest points :",Counter(train_y[neighbors[1][0]]))
```

```
Predicted Class : 7
Actual Class : 7
The 11 nearest neighbours of the test points belongs to classes [7 7 7 7 7
7 7 7 7 7]
Frequency of nearest points : Counter({7: 11})
```

```
In [67]: neighbors
```

```
Out[67]: (array([[ 0.01036054,  0.01036054,  0.01036054,  0.01036054,  0.01191569,
       0.01510356,  0.0162892 ,  0.0162892 ,  0.0162892 ,  0.0162892 ,
       0.0162892 ]]),
 array([[ 420,   461,   902,  1979,  1254,   859,  1066,  1897,   366,  1761,   190]],

      dtype=int64))
```

```
In [68]: train_y[461]
```

```
Out[68]: 7
```

4.2.4. Sample Query Point-2

```
In [69]: clf = KNeighborsClassifier(n_neighbors=alpha[best_alpha])
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

test_point_index = 100

predicted_cls = sig_clf.predict(test_x_responseCoding[test_point_index].reshape(1,-1))
print("Predicted Class :", predicted_cls[0])
print("Actual Class :", test_y[test_point_index])
neighbors = clf.kneighbors(test_x_responseCoding[test_point_index].reshape(1,-1), alpha[best_alpha])
print("the k value for knn is",alpha[best_alpha],"and the nearest neighbours of the test points belongs to classes",train_y[neighbors[1][0]])
print("Frequency of nearest points :",Counter(train_y[neighbors[1][0]]))
```

```
Predicted Class : 1
Actual Class : 1
the k value for knn is 11 and the nearest neighbours of the test points belongs to classes [1 1 1 1 1 1 1 1 1 1]
Frequency of nearest points : Counter({1: 11})
```

4.3. Logistic Regression

4.3.1. With Class balancing

4.3.1.1. Hyper parameter tuning

```
In [70]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight])      Fit the calibrated model
# get_params([deep])      Get parameters for this estimator.
# predict(X)      Predict the target of new samples.
# predict_proba(X)    Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10 ** x for x in range(-6, 3)]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = SGDClassifier(class_weight='balanced', alpha=i, penalty='L2', loss='log', random_state=42)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
        # to avoid rounding error while multiplying probabilites we use log-probability estimates
    print("Log Loss :",log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
```

```
        ax.annotate((alpha[i],str(txt)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

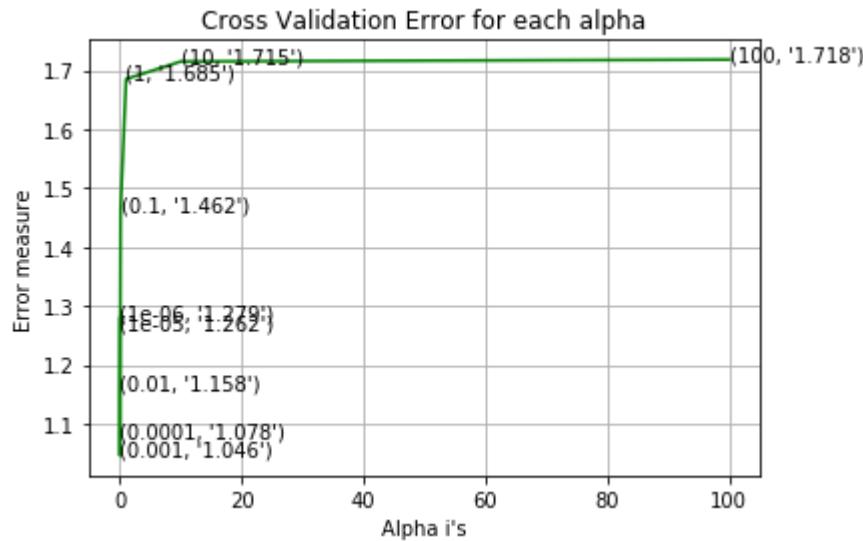
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty=
'12', loss='log', random_state=42)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss i
s:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation
log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss i
s:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 1e-06
Log Loss : 1.2789793164016956
for alpha = 1e-05
Log Loss : 1.2624669712593821
for alpha = 0.0001
Log Loss : 1.0779519350867943
for alpha = 0.001
Log Loss : 1.046058461000146
for alpha = 0.01
Log Loss : 1.1584225799872376
for alpha = 0.1
Log Loss : 1.4622561570695394
for alpha = 1
Log Loss : 1.6848912317575597
for alpha = 10
Log Loss : 1.7147107939022008
for alpha = 100
Log Loss : 1.7178785390961164

```



For values of best alpha = 0.001 The train log loss is: 0.5777638516853337
 For values of best alpha = 0.001 The cross validation log loss is: 1.046058461000146
 For values of best alpha = 0.001 The test log loss is: 1.0786856577858994

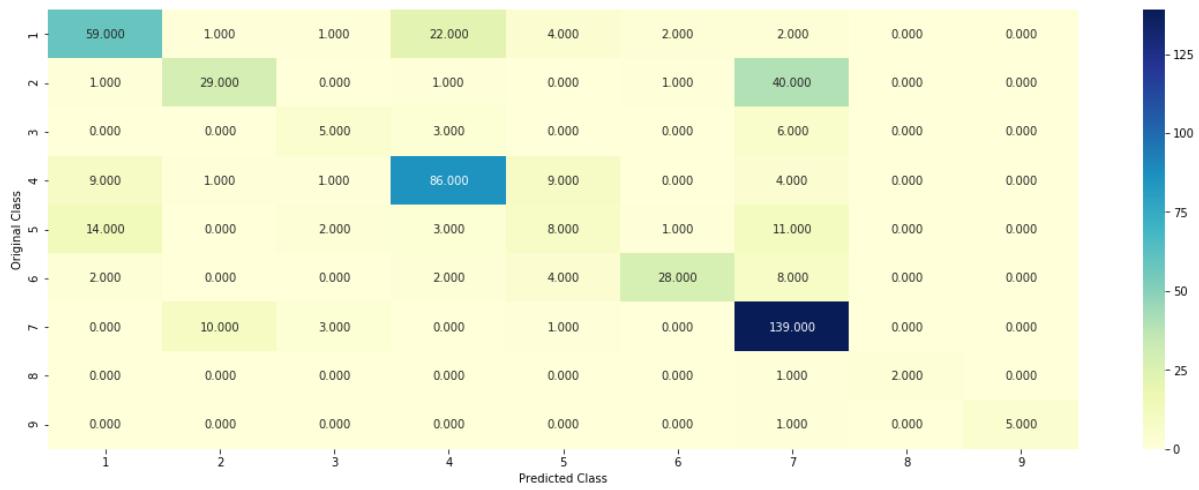
4.3.1.2. Testing the model with best hyper paramters

```
In [71]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

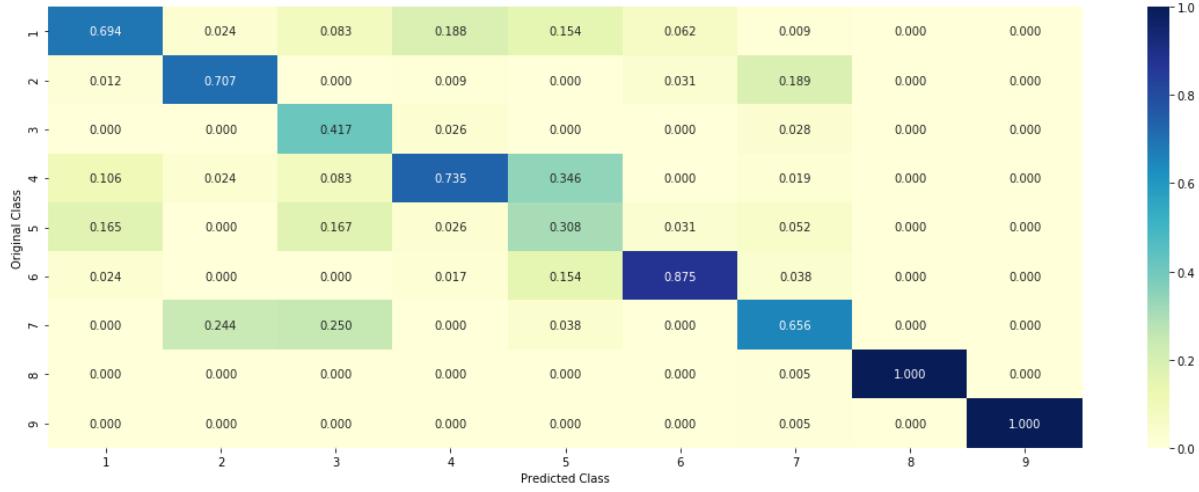
# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)    Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCoding, cv_y, clf)
```

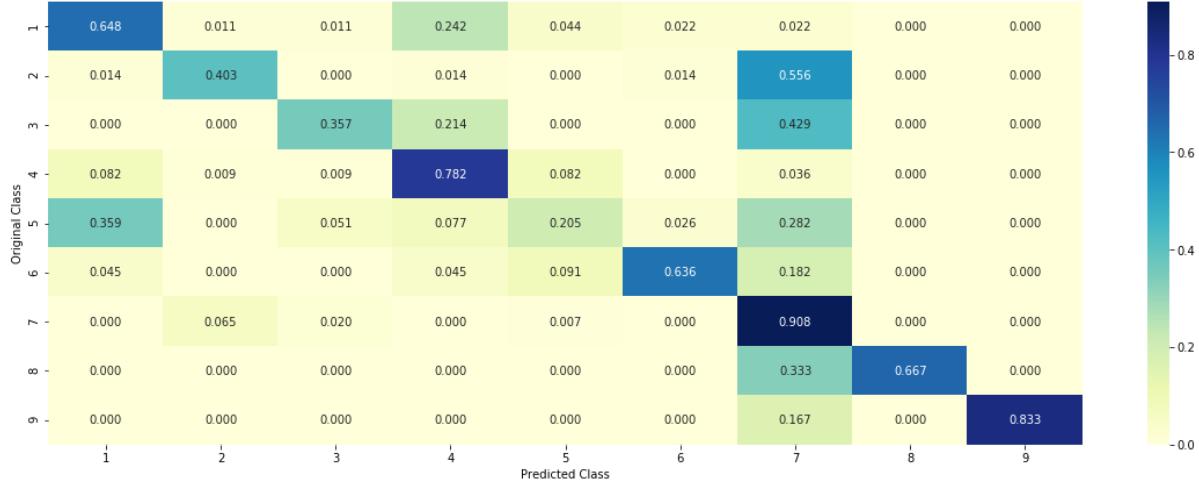
Log loss : 1.046058461000146
 Number of mis-classified points : 0.32142857142857145
 ----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.3.1.3. Feature Importance

```
In [0]: def get_imp_feature_names(text, indices, removed_ind = []):
    word_present = 0
    tabulte_list = []
    incresingorder_ind = 0
    for i in indices:
        if i < train_gene_feature_onehotCoding.shape[1]:
            tabulte_list.append([incresingorder_ind, "Gene", "Yes"])
        elif i < 18:
            tabulte_list.append([incresingorder_ind, "Variation", "Yes"])
        if ((i > 17) & (i not in removed_ind)):
            word = train_text_features[i]
            yes_no = True if word in text.split() else False
            if yes_no:
                word_present += 1
            tabulte_list.append([incresingorder_ind, train_text_features[i], yes_no])
        incresingorder_ind += 1
    print(word_present, "most important features are present in our query point")
    print("-"*50)
    print("The features that are most important of the ",predicted_cls[0]," class:")
    print (tabulate(tabulte_list, headers=["Index", 'Feature name', 'Present or Not']))
```

```
In [77]: (indices).shape
```

```
Out[77]: (1, 500)
```

4.3.1.3.1. Correctly Classified point

```
In [83]: # from tabulate import tabulate
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
clf.fit(train_x_onehotCoding,train_y)
test_point_index = 1
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class : ", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_onehotCoding[test_point_index]),4))
print("Actual Class : ", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)
```

```
Predicted Class : 4
Predicted Class Probabilities: [[0.0487 0.0648 0.0108 0.7572 0.0325 0.0201 0.
0539 0.0059 0.0062]]
Actual Class : 4
-----
238 Text feature [novel] present in test data point [True]
307 Text feature [instability] present in test data point [True]
370 Text feature [potential] present in test data point [True]
375 Text feature [inhibitor] present in test data point [True]
387 Text feature [suppressor] present in test data point [True]
433 Text feature [cycloheximide] present in test data point [True]
468 Text feature [behavioural] present in test data point [True]
Out of the top 500 features 7 are present in query point
```

4.3.1.3.2. Incorrectly Classified point

```
In [84]: test_point_index = 361
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_impfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

```
Predicted Class : 3
Predicted Class Probabilities: [[0.0181 0.0301 0.512  0.184  0.1832 0.0056 0.
0582 0.0051 0.0039]]
Actual Class : 7
-----
292 Text feature [2a] present in test data point [True]
328 Text feature [confirmed] present in test data point [True]
372 Text feature [2b] present in test data point [True]
377 Text feature [demonstrate] present in test data point [True]
378 Text feature [novel] present in test data point [True]
388 Text feature [yet] present in test data point [True]
441 Text feature [lines] present in test data point [True]
442 Text feature [described34] present in test data point [True]
454 Text feature [cell] present in test data point [True]
Out of the top 500 features 9 are present in query point
```

4.3.2. Without Class balancing

4.3.2.1. Hyper parameter tuning

```
In [85]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight])  Fit the calibrated model
# get_params([deep])  Get parameters for this estimator.
# predict(X)  Predict the target of new samples.
# predict_proba(X)  Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10 ** x for x in range(-6, 1)]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = SGDClassifier(alpha=i, penalty='l2', loss='log', random_state=42)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
    print("Log Loss :", log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array, c='g')
for i, txt in enumerate(np.round(cv_log_error_array, 3)):
    ax.annotate((alpha[i], str(txt)), (alpha[i], cv_log_error_array[i]))
plt.grid()
```

```
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

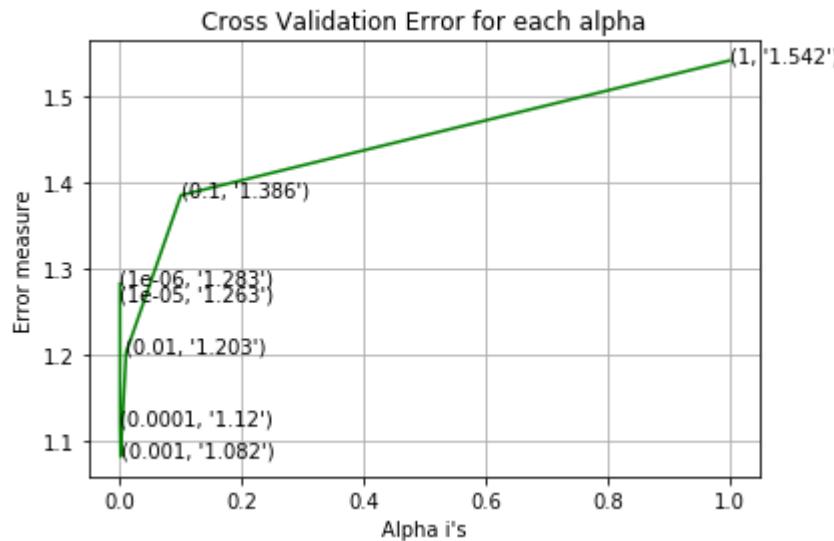
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_
state=42)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss i
s:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation
log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss i
s:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 1e-06
Log Loss : 1.2829556174199352
for alpha = 1e-05
Log Loss : 1.263125984921128
for alpha = 0.0001
Log Loss : 1.120029450631444
for alpha = 0.001
Log Loss : 1.082280308904442
for alpha = 0.01
Log Loss : 1.2029388623234842
for alpha = 0.1
Log Loss : 1.3855537525890227
for alpha = 1
Log Loss : 1.5416864384646043

```



For values of best alpha = 0.001 The train log loss is: 0.5795939891007535
 For values of best alpha = 0.001 The cross validation log loss is: 1.082280308904442
 For values of best alpha = 0.001 The test log loss is: 1.1180239393975904

4.3.2.2. Testing model with best hyper parameters

```
In [86]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)    Predict class labels for samples in X.

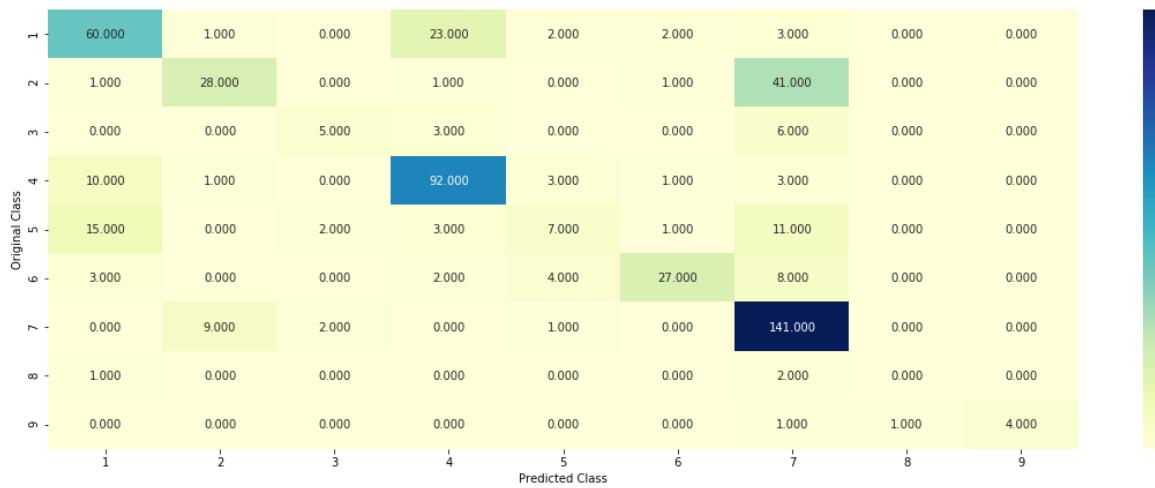
#-----
# video link:
#-----

clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCoding, cv_y, clf)
```

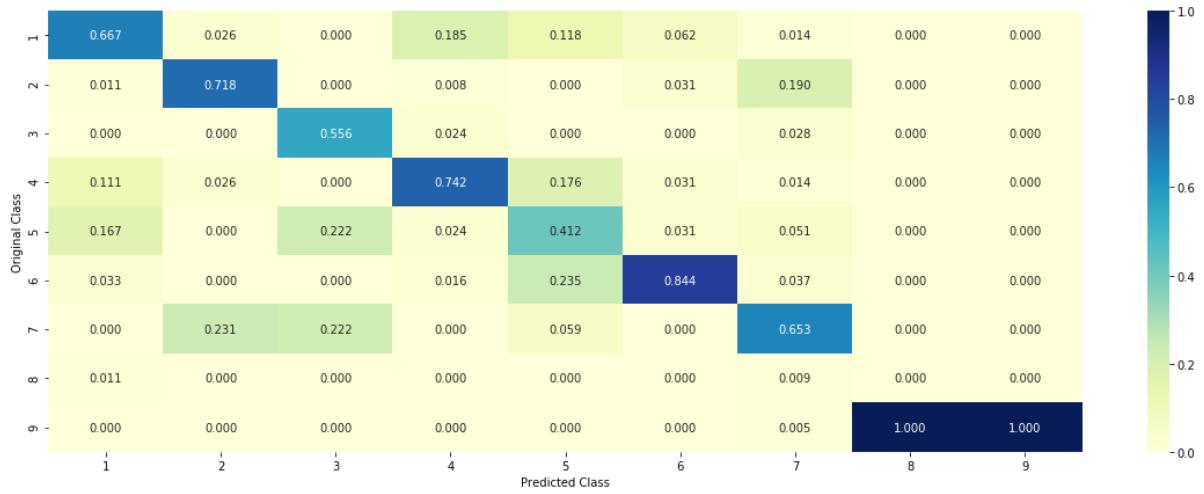
Log loss : 1.082280308904442

Number of mis-classified points : 0.3157894736842105

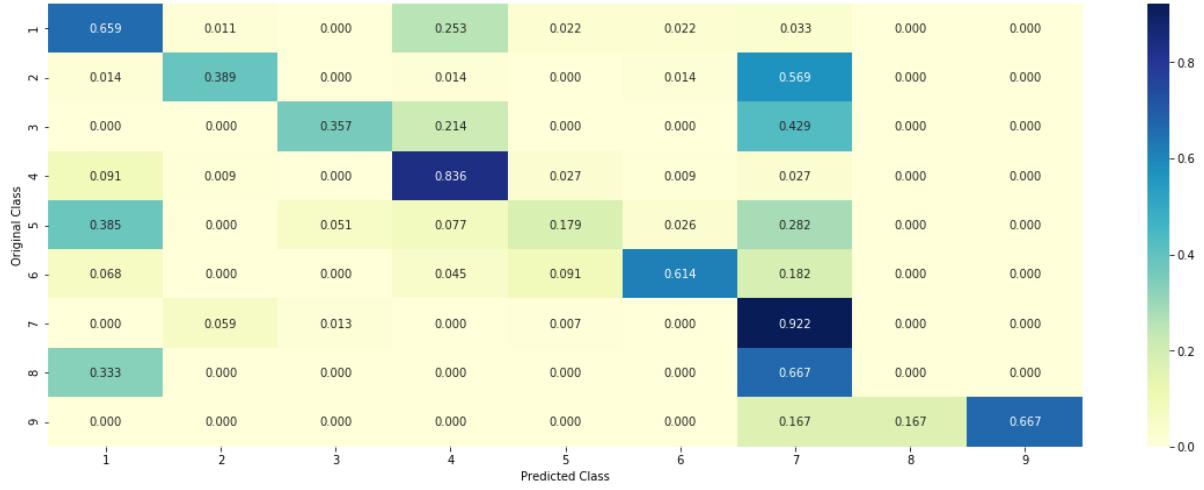
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.3.2.3. Feature Importance, Correctly Classified point

```
In [87]: clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_
state=42)
clf.fit(train_x_onehotCoding,train_y)
test_point_index = 1
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)

Predicted Class : 4
Predicted Class Probabilities: [[0.0476 0.0637 0.0155 0.7552 0.034 0.0206 0.
0535 0.0059 0.004 ]]
Actual Class : 4
-----
374 Text feature [instability] present in test data point [True]
432 Text feature [suppressor] present in test data point [True]
436 Text feature [novel] present in test data point [True]
484 Text feature [behavioural] present in test data point [True]
Out of the top 500 features 4 are present in query point
```

4.3.2.4. Feature Importance, Inorrectly Classified point

```
In [92]: test_point_index = 125
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)

Predicted Class : 4
Predicted Class Probabilities: [[0.073 0.0968 0.0191 0.4863 0.1897 0.0292 0.
0936 0.0072 0.0052]]
Actual Class : 5
-----
432 Text feature [suppressor] present in test data point [True]
Out of the top 500 features 1 are present in query point
```

4.4. Linear Support Vector Machines

4.4.1. Hyper parameter tuning

```
In [93]: # read more about support vector machines with Linear kernels here http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

# -----
# default parameters
# SVC(C=1.0, kernel='rbf', degree=3, gamma='auto', coef0=0.0, shrinking=True,
# probability=False, tol=0.001,
# cache_size=200, class_weight=None, verbose=False, max_iter=-1, decision_function_shape='ovr', random_state=None)

# Some of methods of SVM()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given training data.
# predict(X) Perform classification on samples in X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/mathematical-derivation-copy-8/
# -----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10 ** x for x in range(-5, 3)]
cv_log_error_array = []
for i in alpha:
    print("for C =", i)
    # clf = SVC(C=i,kernel='linear',probability=True, class_weight='balanced')
    clf = SGDClassifier(class_weight='balanced', alpha=i, penalty='l2', loss='hinge', random_state=42)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
    print("Log Loss :", log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array, c='g')
for i, txt in enumerate(np.round(cv_log_error_array, 3)):
    ax.annotate((alpha[i], str(txt)), (alpha[i], cv_log_error_array[i]))
```

```
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

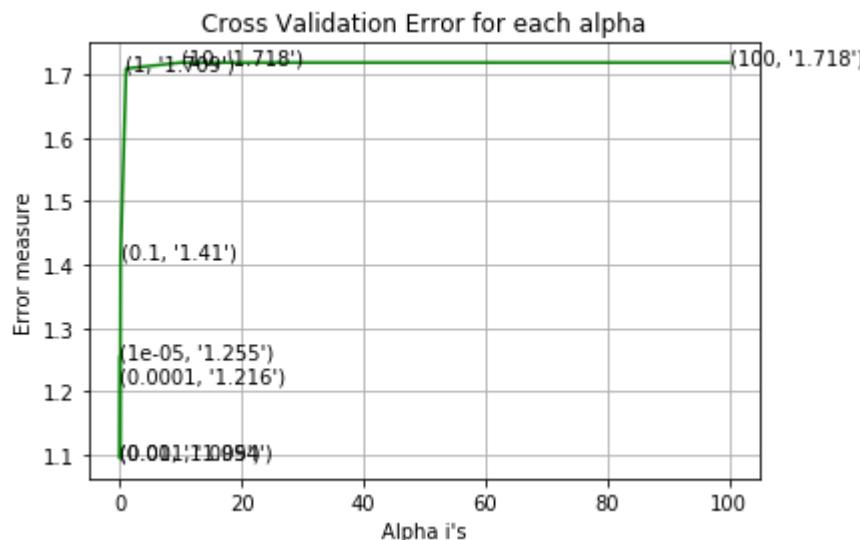
best_alpha = np.argmin(cv_log_error_array)
# clf = SVC(C=i,kernel='linear',probability=True, class_weight='balanced')
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='hinge', random_state=42)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for C = 1e-05
Log Loss : 1.254578750728679
for C = 0.0001
Log Loss : 1.2155614792360583
for C = 0.001
Log Loss : 1.0944931450460749
for C = 0.01
Log Loss : 1.095087277732763
for C = 0.1
Log Loss : 1.4095233415495647
for C = 1
Log Loss : 1.7090746999249902
for C = 10
Log Loss : 1.718440956285807
for C = 100
Log Loss : 1.7184419060141585

```



For values of best alpha = 0.001 The train log loss is: 0.5908638088671851
 For values of best alpha = 0.001 The cross validation log loss is: 1.0944931450460749
 For values of best alpha = 0.001 The test log loss is: 1.116440219341754

4.4.2. Testing model with best hyper parameters

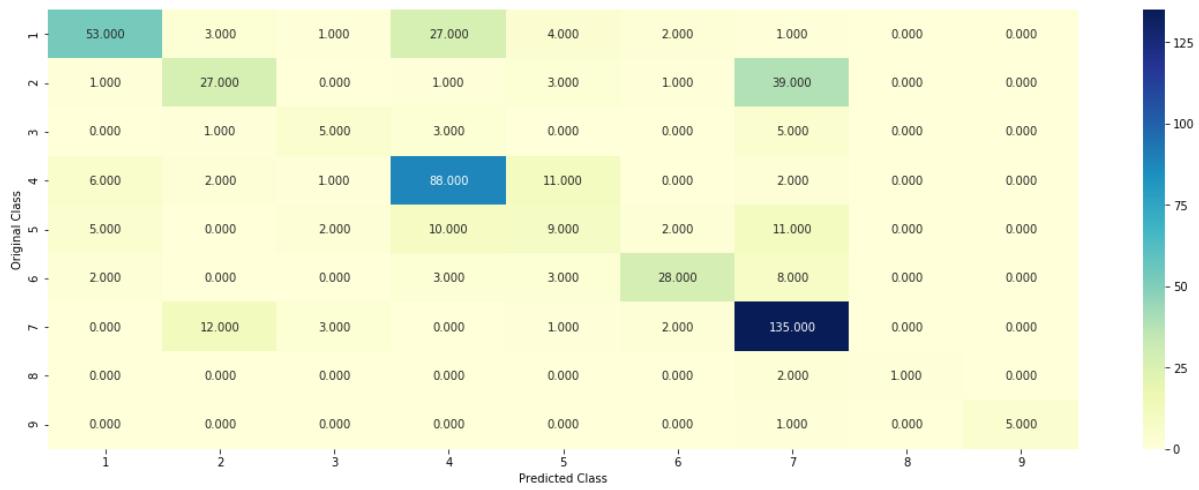
```
In [94]: # read more about support vector machines with Linear kernels here http://scikit-learn.org/stable/modules/generated/skLearn.svm.SVC.html

# -----
# default parameters
# SVC(C=1.0, kernel='rbf', degree=3, gamma='auto', coef0=0.0, shrinking=True,
# probability=False, tol=0.001,
# cache_size=200, class_weight=None, verbose=False, max_iter=-1, decision_function_shape='ovr', random_state=None)

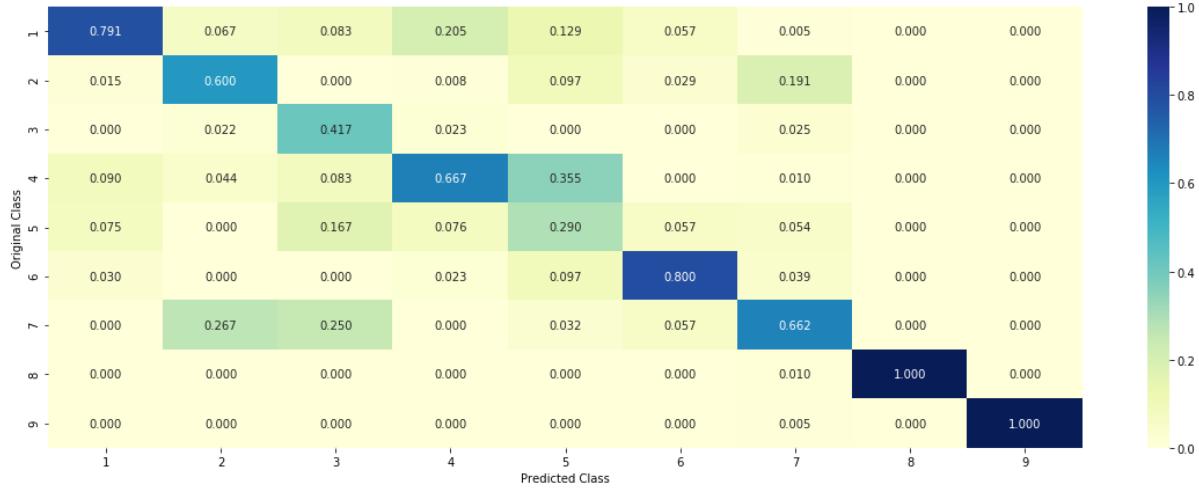
# Some of methods of SVM()
# fit(X, y, [sample_weight])      Fit the SVM model according to the given training data.
# predict(X)        Perform classification on samples in X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/mathematical-derivation-copy-8/
# -----


# clf = SVC(C=alpha[best_alpha],kernel='Linear',probability=True, class_weight='balanced')
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='hinge', random_state=42, class_weight='balanced')
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCoding, cv_y, clf)
```

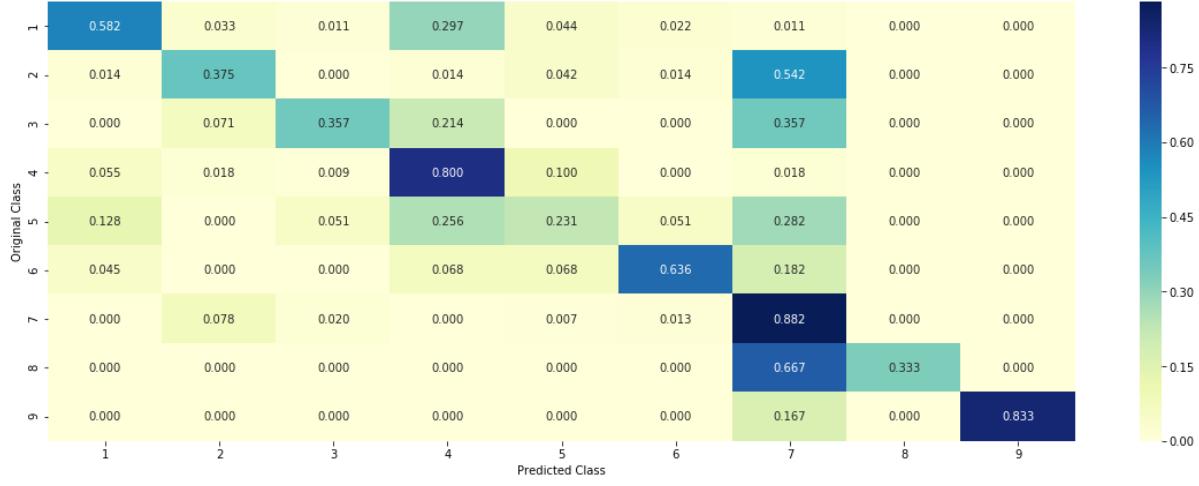
Log loss : 1.0944931450460749
 Number of mis-classified points : 0.34022556390977443
 ----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.3.3. Feature Importance

4.3.3.1. For Correctly classified point

```
In [96]: clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='hinge', random_state=42)
clf.fit(train_x_onehotCoding,train_y)
# test_point_index = 1
test_point_index = 100
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)

Predicted Class : 1
Predicted Class Probabilities: [[0.5774 0.0535 0.0045 0.0783 0.0136 0.0117 0.254 0.0027 0.0043]]
Actual Class : 1
-----
365 Text feature [jp] present in test data point [True]
499 Text feature [hec] present in test data point [True]
Out of the top 500 features 2 are present in query point
```

4.3.3.2. For Incorrectly classified point

```
In [99]: test_point_index = 200
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)

Predicted Class : 7
Predicted Class Probabilities: [[0.0187 0.1084 0.0037 0.0309 0.0087 0.0096 0.814 0.0026 0.0034]]
Actual Class : 2
-----
Out of the top 500 features 0 are present in query point
```

4.5 Random Forest Classifier

4.5.1. Hyper parameter tuning (With One hot Encoding)

In [100]:

```

# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', max_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given training data.
# predict(X) Perform classification on samples in X.
# predict_proba (X) Perform classification on samples in X.

# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/random-forest-and-their-construction-2/
# -----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [100,200,500,1000,2000]
max_depth = [5, 10]
cv_log_error_array = []
for i in alpha:
    for j in max_depth:
        print("for n_estimators =", i,"and max depth = ", j)
        clf = RandomForestClassifier(n_estimators=i, criterion='gini', max_depth=j, random_state=42, n_jobs=-1)
        clf.fit(train_x_onehotCoding, train_y)
        sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
        sig_clf.fit(train_x_onehotCoding, train_y)
        sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
        cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))

```

```
print("Log Loss :",log_loss(cv_y, sig_clf_probs))

'''fig, ax = plt.subplots()
features = np.dot(np.array(alpha)[:,None],np.array(max_depth)[None]).ravel()
ax.plot(features, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
    ax.annotate((alpha[int(i/2)],max_depth[int(i%2)],str(txt)), (features[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()
'''

best_alpha = np.argmin(cv_log_error_array)
clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/2)], criterion='gini', max_depth=max_depth[int(best_alpha%2)], random_state=42, n_jobs=-1)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best estimator = ', alpha[int(best_alpha/2)], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best estimator = ', alpha[int(best_alpha/2)], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best estimator = ', alpha[int(best_alpha/2)], "The test 1 log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```
for n_estimators = 100 and max depth = 5
Log Loss : 1.216052598778494
for n_estimators = 100 and max depth = 10
Log Loss : 1.1622743557027624
for n_estimators = 200 and max depth = 5
Log Loss : 1.2057197303916936
for n_estimators = 200 and max depth = 10
Log Loss : 1.1446914106199184
for n_estimators = 500 and max depth = 5
Log Loss : 1.2018139855328338
for n_estimators = 500 and max depth = 10
Log Loss : 1.1315819522573314
for n_estimators = 1000 and max depth = 5
Log Loss : 1.1999937512455041
for n_estimators = 1000 and max depth = 10
Log Loss : 1.1301463337721482
for n_estimators = 2000 and max depth = 5
Log Loss : 1.1986645140496466
for n_estimators = 2000 and max depth = 10
Log Loss : 1.1286939020185216
For values of best estimator = 2000 The train log loss is: 0.706001795234539
1
For values of best estimator = 2000 The cross validation log loss is: 1.1286
939020185216
For values of best estimator = 2000 The test log loss is: 1.1460809246060735
```

4.5.2. Testing model with best hyper parameters (One Hot Encoding)

In [101]:

```
# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', m
ax_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_l
eaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_s
tate=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given train
ing data.
# predict(X) Perform classification on samples in X.
# predict_proba (X) Perform classification on samples in X.

# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

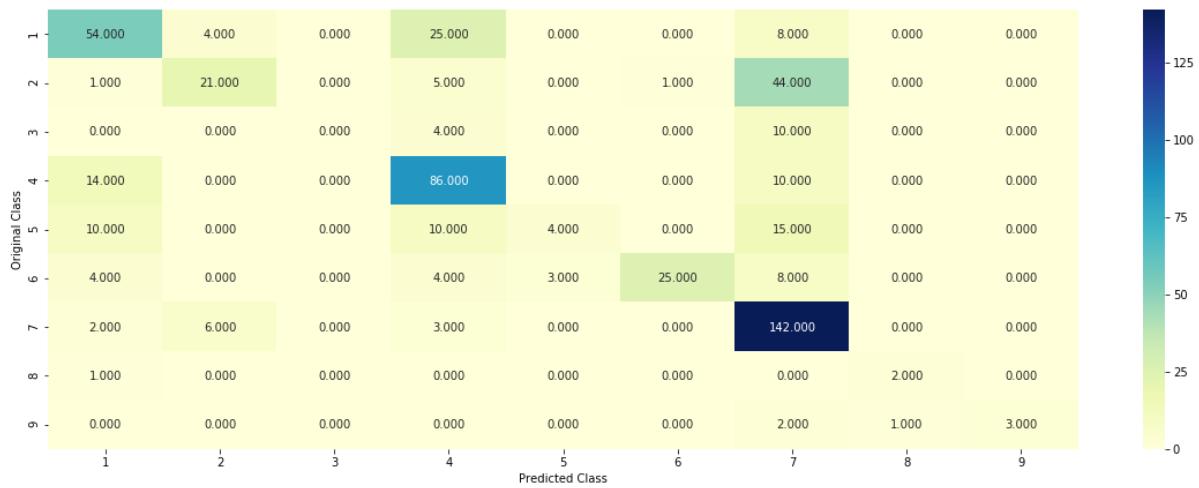
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/
Lessons/random-forest-and-their-construction-2/
# -----
```

```
clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/2)], criterion=
'gini', max_depth=max_depth[int(best_alpha%2)], random_state=42, n_jobs=-1)
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCod
ing, cv_y, clf)
```

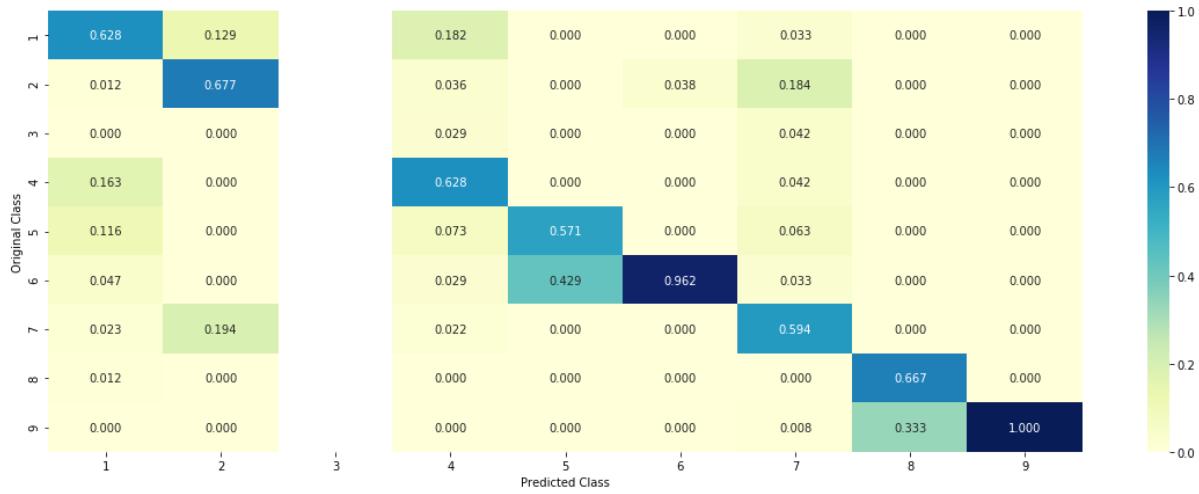
Log loss : 1.1286939020185216

Number of mis-classified points : 0.36654135338345867

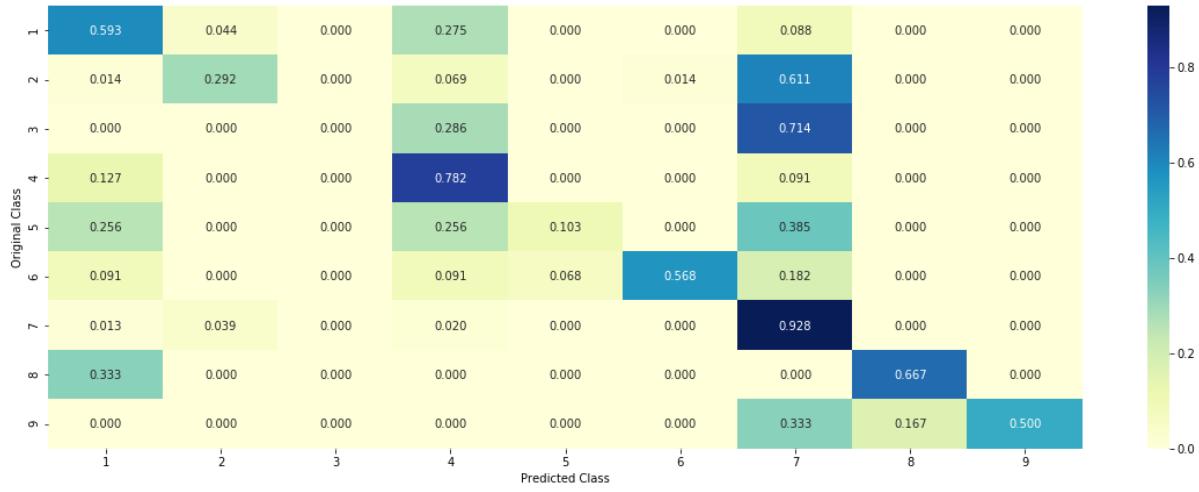
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.5.3. Feature Importance

4.5.3.1. Correctly Classified point

```
In [106]: # test_point_index = 10
clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/2)], criterion=
'gini', max_depth=max_depth[int(best_alpha%2)], random_state=42, n_jobs=-1)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

test_point_index = 1
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.feature_importances_)
print("-"*50)
get_impfeature_names(indices[:no_feature], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)
```

```
Predicted Class : 4
Predicted Class Probabilities: [[0.1002 0.0355 0.0175 0.711  0.0417 0.0338 0.
0503 0.0034 0.0066]]
Actual Class : 4
-----
0 Text feature [kinase] present in test data point [True]
1 Text feature [inhibitors] present in test data point [True]
6 Text feature [suppressor] present in test data point [True]
7 Text feature [inhibitor] present in test data point [True]
8 Text feature [phosphorylation] present in test data point [True]
9 Text feature [activated] present in test data point [True]
12 Text feature [missense] present in test data point [True]
15 Text feature [function] present in test data point [True]
17 Text feature [cells] present in test data point [True]
26 Text feature [akt] present in test data point [True]
29 Text feature [stability] present in test data point [True]
30 Text feature [loss] present in test data point [True]
32 Text feature [trials] present in test data point [True]
33 Text feature [inhibition] present in test data point [True]
34 Text feature [downstream] present in test data point [True]
37 Text feature [growth] present in test data point [True]
39 Text feature [functional] present in test data point [True]
41 Text feature [expressing] present in test data point [True]
43 Text feature [yeast] present in test data point [True]
44 Text feature [unstable] present in test data point [True]
47 Text feature [patients] present in test data point [True]
48 Text feature [treated] present in test data point [True]
56 Text feature [cell] present in test data point [True]
62 Text feature [truncating] present in test data point [True]
63 Text feature [protein] present in test data point [True]
72 Text feature [clinical] present in test data point [True]
74 Text feature [phosphatase] present in test data point [True]
86 Text feature [patient] present in test data point [True]
90 Text feature [pten] present in test data point [True]
92 Text feature [mutant] present in test data point [True]
98 Text feature [proteins] present in test data point [True]
99 Text feature [phosphoinositide] present in test data point [True]
Out of the top 100 features 32 are present in query point
```

4.5.3.2. Inorrectly Classified point

```
In [107]: test_point_index = 100
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.feature_importances_)
print("-"*50)
get_impfeature_names(indices[:no_feature], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)
```

Predicted Class : 7
Predicted Class Probabilities: [[0.2024 0.1222 0.0205 0.1033 0.0457 0.0403 0.
4476 0.0088 0.0093]]
Actuall Class : 1

0 Text feature [kinase] present in test data point [True]
1 Text feature [inhibitors] present in test data point [True]
2 Text feature [tyrosine] present in test data point [True]
3 Text feature [activating] present in test data point [True]
4 Text feature [activation] present in test data point [True]
5 Text feature [constitutive] present in test data point [True]
6 Text feature [suppressor] present in test data point [True]
7 Text feature [inhibitor] present in test data point [True]
8 Text feature [phosphorylation] present in test data point [True]
10 Text feature [oncogenic] present in test data point [True]
11 Text feature [treatment] present in test data point [True]
13 Text feature [transforming] present in test data point [True]
14 Text feature [signaling] present in test data point [True]
15 Text feature [function] present in test data point [True]
17 Text feature [cells] present in test data point [True]
18 Text feature [therapy] present in test data point [True]
20 Text feature [drug] present in test data point [True]
22 Text feature [nonsense] present in test data point [True]
23 Text feature [therapeutic] present in test data point [True]
24 Text feature [mek] present in test data point [True]
25 Text feature [erk] present in test data point [True]
26 Text feature [akt] present in test data point [True]
30 Text feature [loss] present in test data point [True]
33 Text feature [inhibition] present in test data point [True]
34 Text feature [downstream] present in test data point [True]
36 Text feature [deleterious] present in test data point [True]
37 Text feature [growth] present in test data point [True]
38 Text feature [resistance] present in test data point [True]
39 Text feature [functional] present in test data point [True]
41 Text feature [expressing] present in test data point [True]
43 Text feature [yeast] present in test data point [True]
45 Text feature [ic50] present in test data point [True]
47 Text feature [patients] present in test data point [True]
48 Text feature [treated] present in test data point [True]
49 Text feature [months] present in test data point [True]
56 Text feature [cell] present in test data point [True]
59 Text feature [survival] present in test data point [True]
60 Text feature [lines] present in test data point [True]
63 Text feature [protein] present in test data point [True]
67 Text feature [frameshift] present in test data point [True]
68 Text feature [factor] present in test data point [True]
71 Text feature [lung] present in test data point [True]
72 Text feature [clinical] present in test data point [True]
74 Text feature [phosphatase] present in test data point [True]
75 Text feature [proliferation] present in test data point [True]
78 Text feature [egfr] present in test data point [True]
81 Text feature [p53] present in test data point [True]
82 Text feature [oncogene] present in test data point [True]
83 Text feature [dna] present in test data point [True]
85 Text feature [response] present in test data point [True]
86 Text feature [patient] present in test data point [True]
88 Text feature [respond] present in test data point [True]

```
90 Text feature [pten] present in test data point [True]
92 Text feature [mutant] present in test data point [True]
93 Text feature [assays] present in test data point [True]
95 Text feature [phospho] present in test data point [True]
98 Text feature [proteins] present in test data point [True]
99 Text feature [phosphoinositide] present in test data point [True]
Out of the top 100 features 58 are present in query point
```

4.5.3. Hyper parameter tuning (With Response Coding)

In [108]:

```

# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', max_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given training data.
# predict(X) Perform classification on samples in X.
# predict_proba (X) Perform classification on samples in X.

# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/random-forest-and-their-construction-2/
# -----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10,50,100,200,500,1000]
max_depth = [2,3,5,10]
cv_log_error_array = []
for i in alpha:
    for j in max_depth:
        print("for n_estimators =", i, "and max depth = ", j)
        clf = RandomForestClassifier(n_estimators=i, criterion='gini', max_depth=j, random_state=42, n_jobs=-1)
        clf.fit(train_x_responseCoding, train_y)
        sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
        sig_clf.fit(train_x_responseCoding, train_y)
        sig_clf_probs = sig_clf.predict_proba(cv_x_responseCoding)
        cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))

```

```
        print("Log Loss :",log_loss(cv_y, sig_clf_probs))
    ...
    fig, ax = plt.subplots()
    features = np.dot(np.array(alpha)[:,None],np.array(max_depth)[None]).ravel()
    ax.plot(features, cv_log_error_array,c='g')
    for i, txt in enumerate(np.round(cv_log_error_array,3)):
        ax.annotate((alpha[int(i/4)],max_depth[int(i%4)],str(txt)), (features[i],cv_log_error_array[i]))
    plt.grid()
    plt.title("Cross Validation Error for each alpha")
    plt.xlabel("Alpha i's")
    plt.ylabel("Error measure")
    plt.show()
    '''

best_alpha = np.argmin(cv_log_error_array)
clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/4)], criterion='gini', max_depth=max_depth[int(best_alpha%4)], random_state=42, n_jobs=-1)
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_responseCoding)
print('For values of best alpha = ', alpha[int(best_alpha/4)], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_responseCoding)
print('For values of best alpha = ', alpha[int(best_alpha/4)], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_responseCoding)
print('For values of best alpha = ', alpha[int(best_alpha/4)], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for n_estimators = 10 and max depth = 2
Log Loss : 2.1179624203118155
for n_estimators = 10 and max depth = 3
Log Loss : 1.7853627943890886
for n_estimators = 10 and max depth = 5
Log Loss : 1.5718120625512921
for n_estimators = 10 and max depth = 10
Log Loss : 2.0120854977740414
for n_estimators = 50 and max depth = 2
Log Loss : 1.7627564426840363
for n_estimators = 50 and max depth = 3
Log Loss : 1.5131291375179743
for n_estimators = 50 and max depth = 5
Log Loss : 1.4365038442977867
for n_estimators = 50 and max depth = 10
Log Loss : 1.6931437101579332
for n_estimators = 100 and max depth = 2
Log Loss : 1.6489738779313445
for n_estimators = 100 and max depth = 3
Log Loss : 1.5536196603086667
for n_estimators = 100 and max depth = 5
Log Loss : 1.3745561420585717
for n_estimators = 100 and max depth = 10
Log Loss : 1.74696216225021
for n_estimators = 200 and max depth = 2
Log Loss : 1.7030387693322697
for n_estimators = 200 and max depth = 3
Log Loss : 1.5507318013034408
for n_estimators = 200 and max depth = 5
Log Loss : 1.4129391915618166
for n_estimators = 200 and max depth = 10
Log Loss : 1.7557119997300097
for n_estimators = 500 and max depth = 2
Log Loss : 1.7604841527597928
for n_estimators = 500 and max depth = 3
Log Loss : 1.6527300379441865
for n_estimators = 500 and max depth = 5
Log Loss : 1.4392492035116047
for n_estimators = 500 and max depth = 10
Log Loss : 1.7682531836094613
for n_estimators = 1000 and max depth = 2
Log Loss : 1.7335159959245123
for n_estimators = 1000 and max depth = 3
Log Loss : 1.6442215466730772
for n_estimators = 1000 and max depth = 5
Log Loss : 1.4330964130063824
for n_estimators = 1000 and max depth = 10
Log Loss : 1.7648085685886328
For values of best alpha = 100 The train log loss is: 0.055553324245587475
For values of best alpha = 100 The cross validation log loss is: 1.374556142
0585715
For values of best alpha = 100 The test log loss is: 1.4182353449496692

```

4.5.4. Testing model with best hyper parameters (Response Coding)

In [109]:

```
# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', max_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given training data.
# predict(X) Perform classification on samples in X.
# predict_proba (X) Perform classification on samples in X.

# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

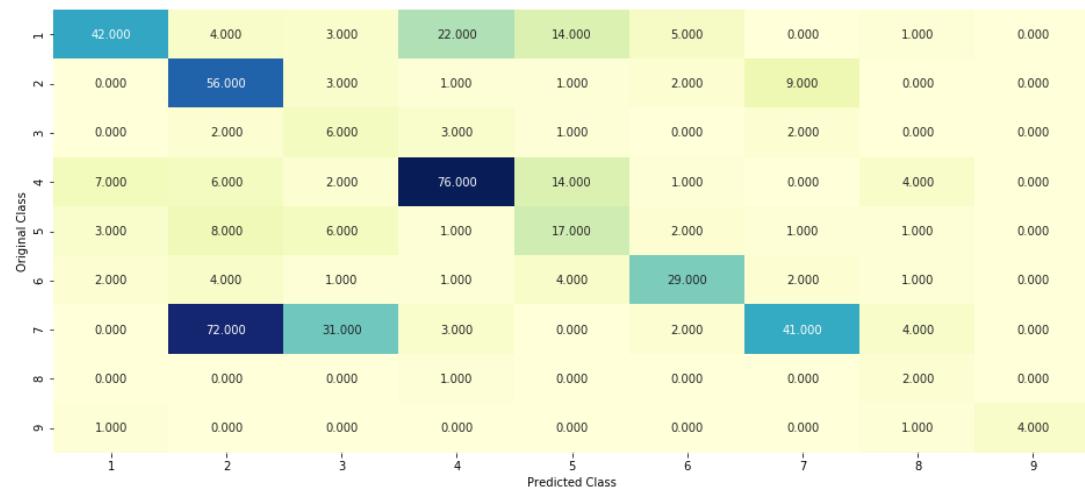
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/random-forest-and-their-construction-2/
# -----


clf = RandomForestClassifier(max_depth=max_depth[int(best_alpha*4)], n_estimators=alpha[int(best_alpha/4)], criterion='gini', max_features='auto', random_state=42)
predict_and_plot_confusion_matrix(train_x_responseCoding, train_y, cv_x_responseCoding, cv_y, clf)
```

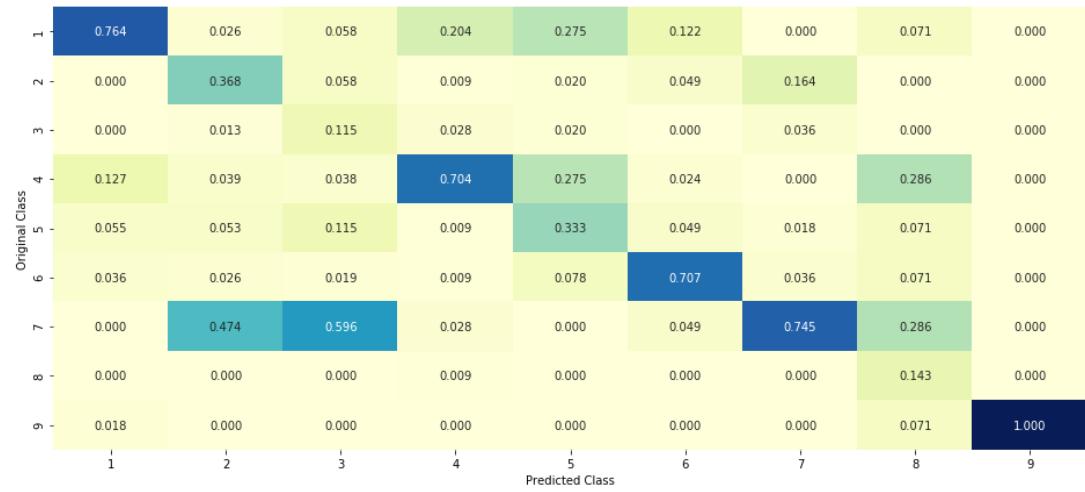
Log loss : 1.3745561420585715

Number of mis-classified points : 0.4868421052631579

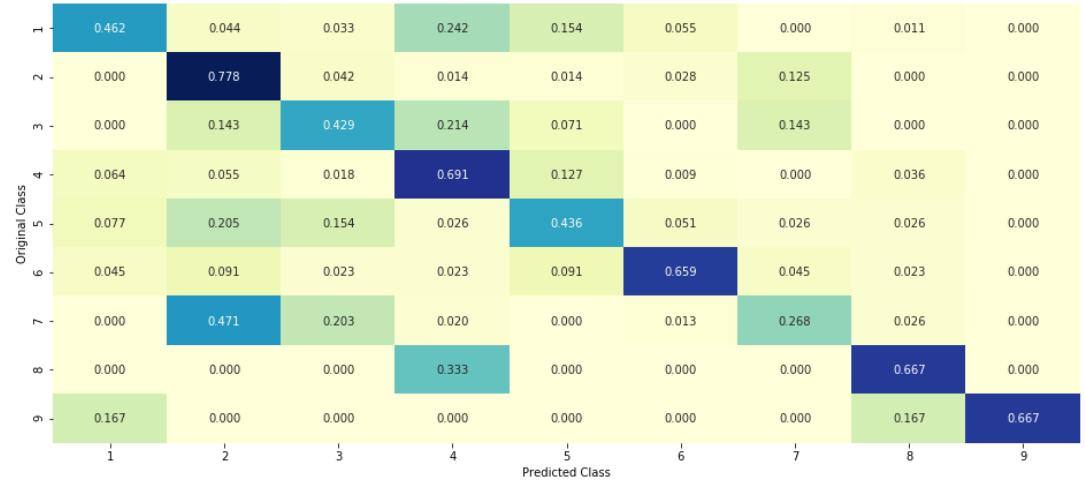
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.5.5. Feature Importance

4.5.5.1. Correctly Classified point

```
In [110]: clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/4)], criterion='gini', max_depth=max_depth[int(best_alpha%4)], random_state=42, n_jobs=-1)
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

test_point_index = 1
no_feature = 27
predicted_cls = sig_clf.predict(test_x_responseCoding[test_point_index].reshape(1,-1))
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_responseCoding[test_point_index].reshape(1,-1)),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.feature_importances_)
print("-"*50)
for i in indices:
    if i<9:
        print("Gene is important feature")
    elif i<18:
        print("Variation is important feature")
    else:
        print("Text is important feature")
```

```
Predicted Class : 4
Predicted Class Probabilities: [[0.0521 0.0206 0.2657 0.5224 0.0277 0.0383 0.
0101 0.0401 0.023 ]]
Actual Class : 4
-----
Variation is important feature
Variation is important feature
Variation is important feature
Variation is important feature
Gene is important feature
Variation is important feature
Variation is important feature
Text is important feature
Text is important feature
Text is important feature
Gene is important feature
Text is important feature
Gene is important feature
Gene is important feature
Text is important feature
Text is important feature
Variation is important feature
Gene is important feature
Gene is important feature
Variation is important feature
Text is important feature
Text is important feature
Gene is important feature
Text is important feature
Variation is important feature
Gene is important feature
Gene is important feature
```

4.5.5.2. Incorrectly Classified point

```
In [113]: test_point_index = 200
predicted_cls = sig_clf.predict(test_x_responseCoding[test_point_index].reshape(1,-1))
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
responseCoding[test_point_index].reshape(1,-1)),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.feature_importances_)
print("-"*50)
for i in indices:
    if i<9:
        print("Gene is important feature")
    elif i<18:
        print("Variation is important feature")
    else:
        print("Text is important feature")
```

```
Predicted Class : 3
Predicted Class Probabilities: [[0.0132 0.1537 0.4766 0.0188 0.0307 0.0357 0.
2172 0.0419 0.0123]]
Actual Class : 2
-----
Variation is important feature
Variation is important feature
Variation is important feature
Variation is important feature
Gene is important feature
Variation is important feature
Variation is important feature
Text is important feature
Text is important feature
Text is important feature
Gene is important feature
Text is important feature
Gene is important feature
Gene is important feature
Text is important feature
Text is important feature
Variation is important feature
Gene is important feature
Gene is important feature
Variation is important feature
Text is important feature
Text is important feature
Gene is important feature
Text is important feature
Variation is important feature
Gene is important feature
Gene is important feature
```

4.7 Stack the models

4.7.1 testing with hyper parameter tuning

```
In [114]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/geometric-intuition-1/
#-----


# read more about support vector machines with Linear kernels here http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html
# -----
# default parameters
# SVC(C=1.0, kernel='rbf', degree=3, gamma='auto', coef0=0.0, shrinking=True,
# probability=False, tol=0.001,
# cache_size=200, class_weight=None, verbose=False, max_iter=-1, decision_function_shape='ovr', random_state=None)

# Some of methods of SVM()
# fit(X, y, [sample_weight])   Fit the SVM model according to the given training data.
# predict(X)     Perform classification on samples in X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/mathematical-derivation-copy-8/
# -----


# read more about support vector machines with Linear kernels here http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html
# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', max_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight])   Fit the SVM model according to the given training data.
```

```

# predict(X)      Perform classification on samples in X.
# predict_proba (X)    Perform classification on samples in X.

# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/
Lessons/random-forest-and-their-construction-2/
# -----


clf1 = SGDClassifier(alpha=0.001, penalty='l2', loss='log', class_weight='balanced', random_state=0)
clf1.fit(train_x_onehotCoding, train_y)
sig_clf1 = CalibratedClassifierCV(clf1, method="sigmoid")

clf2 = SGDClassifier(alpha=1, penalty='l2', loss='hinge', class_weight='balanced', random_state=0)
clf2.fit(train_x_onehotCoding, train_y)
sig_clf2 = CalibratedClassifierCV(clf2, method="sigmoid")

clf3 = MultinomialNB(alpha=0.001)
clf3.fit(train_x_onehotCoding, train_y)
sig_clf3 = CalibratedClassifierCV(clf3, method="sigmoid")

sig_clf1.fit(train_x_onehotCoding, train_y)
print("Logistic Regression : Log Loss: %0.2f" % (log_loss(cv_y, sig_clf1.predict_proba(cv_x_onehotCoding))))
sig_clf2.fit(train_x_onehotCoding, train_y)
print("Support vector machines : Log Loss: %0.2f" % (log_loss(cv_y, sig_clf2.predict_proba(cv_x_onehotCoding))))
sig_clf3.fit(train_x_onehotCoding, train_y)
print("Naive Bayes : Log Loss: %0.2f" % (log_loss(cv_y, sig_clf3.predict_proba(cv_x_onehotCoding))))
print("-"*50)
alpha = [0.0001, 0.001, 0.01, 0.1, 1, 10]
best_alpha = 999
for i in alpha:
    lr = LogisticRegression(C=i)
    sclf = StackingClassifier(classifiers=[sig_clf1, sig_clf2, sig_clf3], meta_classifier=lr, use_probas=True)
    sclf.fit(train_x_onehotCoding, train_y)
    print("Stacking Classifier : for the value of alpha: %f Log Loss: %0.3f" %
(i, log_loss(cv_y, sclf.predict_proba(cv_x_onehotCoding))))
    log_error = log_loss(cv_y, sclf.predict_proba(cv_x_onehotCoding))
    if best_alpha > log_error:
        best_alpha = log_error

```

Logistic Regression : Log Loss: 1.05
Support vector machines : Log Loss: 1.71
Naive Bayes : Log Loss: 1.25

Stacking Classifier : for the value of alpha: 0.000100 Log Loss: 2.178
Stacking Classifier : for the value of alpha: 0.001000 Log Loss: 2.036
Stacking Classifier : for the value of alpha: 0.010000 Log Loss: 1.505
Stacking Classifier : for the value of alpha: 0.100000 Log Loss: 1.079
Stacking Classifier : for the value of alpha: 1.000000 Log Loss: 1.158
Stacking Classifier : for the value of alpha: 10.000000 Log Loss: 1.433

4.7.2 testing the model with the best hyper parameters

```
In [115]: lr = LogisticRegression(C=0.1)
sclf = StackingClassifier(classifiers=[sig_clf1, sig_clf2, sig_clf3], meta_classifier=lr, use_probas=True)
sclf.fit(train_x_onehotCoding, train_y)

log_error = log_loss(train_y, sclf.predict_proba(train_x_onehotCoding))
print("Log loss (train) on the stacking classifier :",log_error)

log_error = log_loss(cv_y, sclf.predict_proba(cv_x_onehotCoding))
print("Log loss (CV) on the stacking classifier :",log_error)

log_error = log_loss(test_y, sclf.predict_proba(test_x_onehotCoding))
print("Log loss (test) on the stacking classifier :",log_error)

print("Number of missclassified point :", np.count_nonzero((sclf.predict(test_x_onehotCoding)- test_y))/test_y.shape[0])
plot_confusion_matrix(test_y=test_y, predict_y=sclf.predict(test_x_onehotCoding))
```

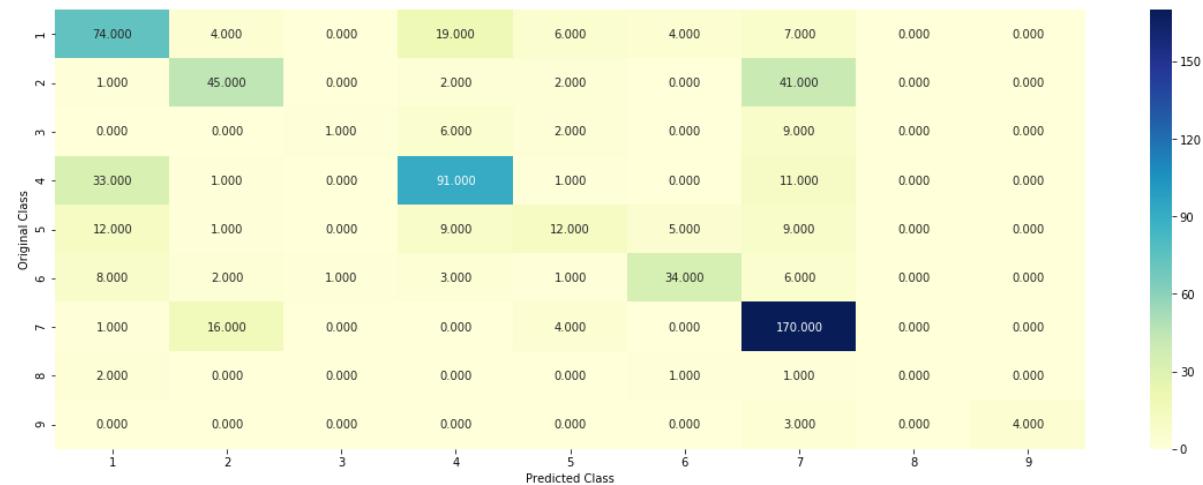
Log loss (train) on the stacking classifier : 0.6570870522901856

Log loss (CV) on the stacking classifier : 1.0790743062470094

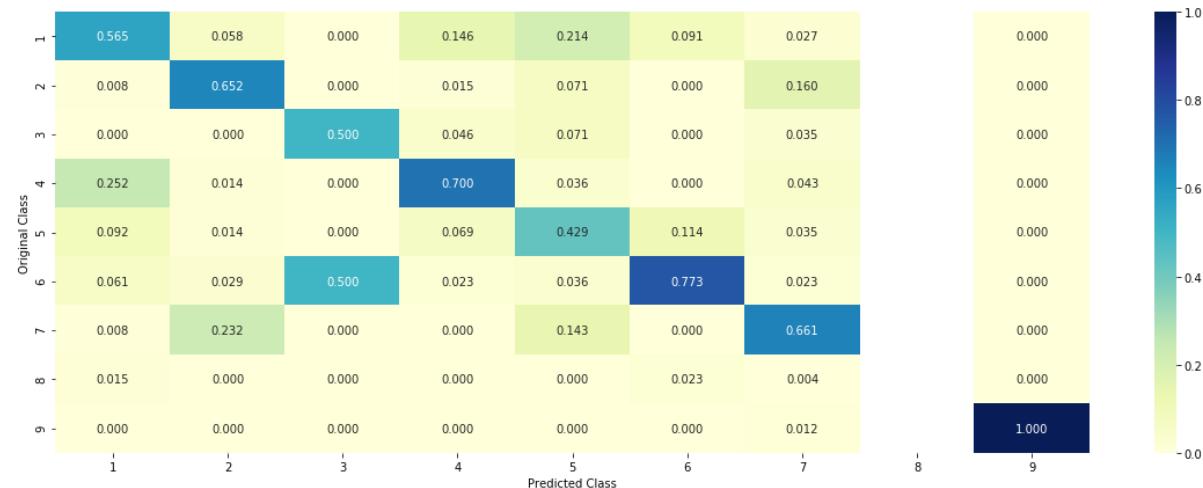
Log loss (test) on the stacking classifier : 1.108588807488054

Number of missclassified point : 0.3518796992481203

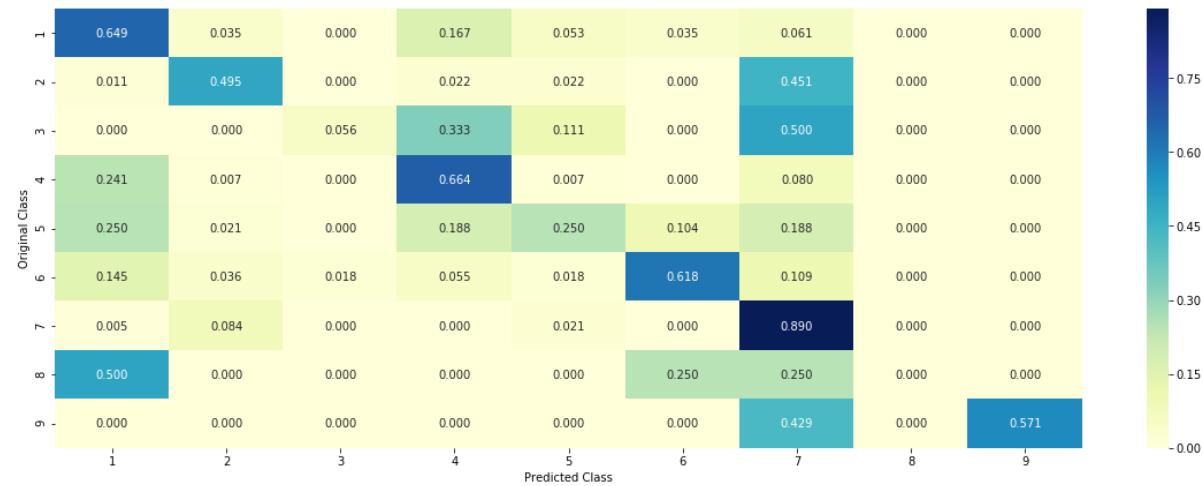
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.7.3 Maximum Voting classifier

```
In [116]: #Refer:http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.VotingClassifier.html
from sklearn.ensemble import VotingClassifier
vclf = VotingClassifier(estimators=[('lr', sig_clf1), ('svc', sig_clf2), ('rf', sig_clf3)], voting='soft')
vclf.fit(train_x_onehotCoding, train_y)
print("Log loss (train) on the VotingClassifier :", log_loss(train_y, vclf.predict_proba(train_x_onehotCoding)))
print("Log loss (CV) on the VotingClassifier :", log_loss(cv_y, vclf.predict_proba(cv_x_onehotCoding)))
print("Log loss (test) on the VotingClassifier :", log_loss(test_y, vclf.predict_proba(test_x_onehotCoding)))
print("Number of missclassified point :", np.count_nonzero((vclf.predict(test_x_onehotCoding)- test_y)/test_y.shape[0]))
plot_confusion_matrix(test_y=test_y, predict_y=vclf.predict(test_x_onehotCoding))
```

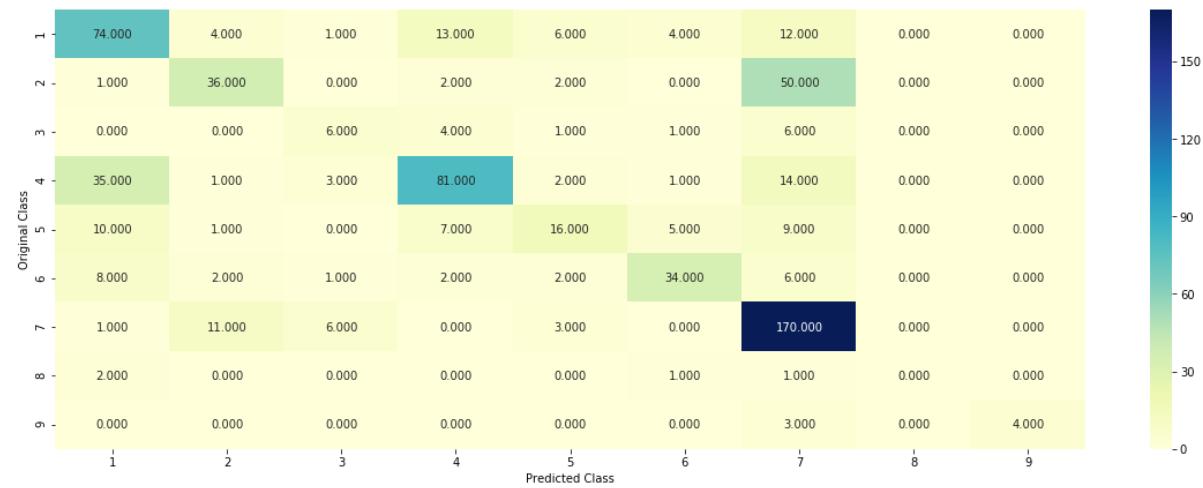
Log loss (train) on the VotingClassifier : 0.9088580015613595

Log loss (CV) on the VotingClassifier : 1.17890065358294

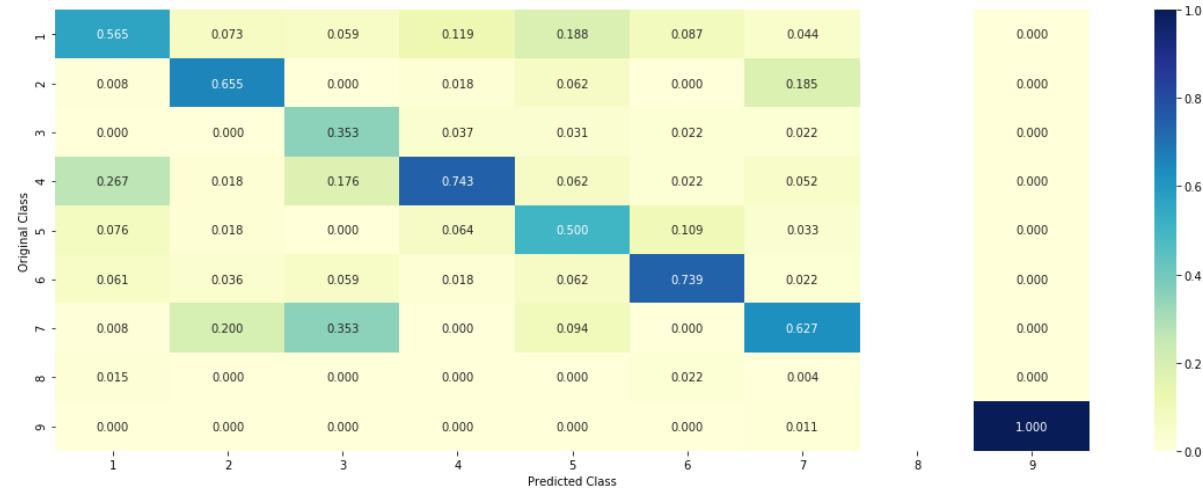
Log loss (test) on the VotingClassifier : 1.1993418040861672

Number of missclassified point : 0.3669172932330827

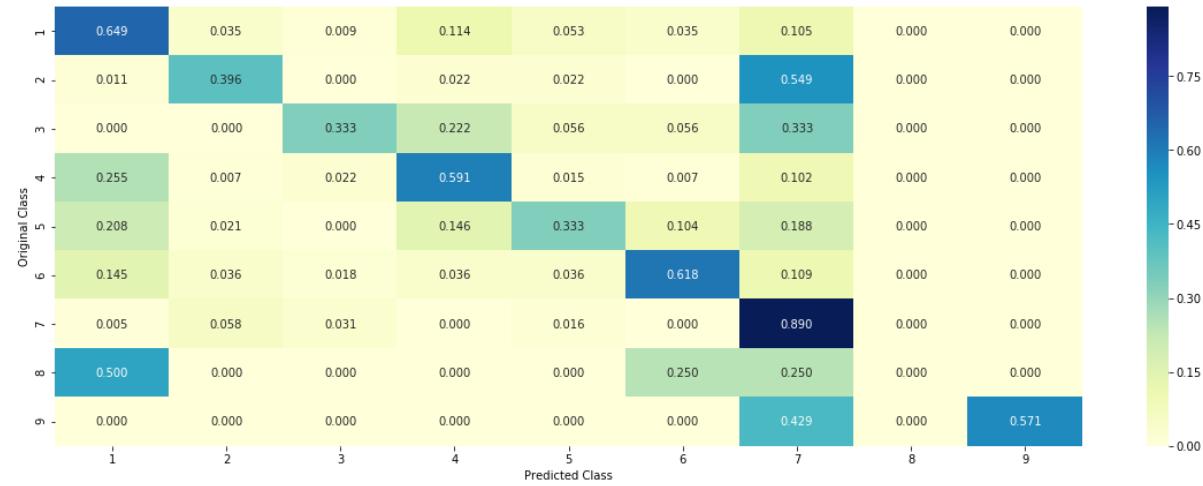
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



5. Assignments

1. Apply All the models with tf-idf features (Replace CountVectorizer with tfidfVectorizer and run the same cells)
2. Instead of using all the words in the dataset, use only the top 1000 words based of tf-idf values
3. Apply Logistic regression with CountVectorizer Features, including both unigrams and bigrams
4. Try any of the feature engineering techniques discussed in the course to reduce the CV and test log-loss to a value less than 1.0

1 & 2:

```
In [24]: # https://machinelearningmastery.com/how-to-fix-futurewarning-messages-in-scikit-learn/
from warnings import simplefilter
# ignore all future warnings
simplefilter(action='ignore', category=FutureWarning)

import pandas as pd
import matplotlib.pyplot as plt
import re
import time
import warnings
import numpy as np
from nltk.corpus import stopwords
from sklearn.decomposition import TruncatedSVD
from sklearn.preprocessing import normalize
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.manifold import TSNE
import seaborn as sns
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix
from sklearn.metrics.classification import accuracy_score, log_loss
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import SGDClassifier
from imblearn.over_sampling import SMOTE
from collections import Counter
from scipy.sparse import hstack
from sklearn.multiclass import OneVsRestClassifier
from sklearn.svm import SVC
from sklearn.model_selection import StratifiedKFold
from collections import Counter, defaultdict
from sklearn.calibration import CalibratedClassifierCV
from sklearn.naive_bayes import MultinomialNB
from sklearn.naive_bayes import GaussianNB
from sklearn.model_selection import train_test_split
from sklearn.model_selection import GridSearchCV
import math
from sklearn.metrics import normalized_mutual_info_score
from sklearn.ensemble import RandomForestClassifier
warnings.filterwarnings("ignore")

from mlxtend.classifier import StackingClassifier

from sklearn import model_selection
from sklearn.linear_model import LogisticRegression
```

3.1. Reading Data

3.1.1. Reading Gene and Variation Data

```
In [2]: data = pd.read_csv('training/training_variants')
print('Number of data points : ', data.shape[0])
print('Number of features : ', data.shape[1])
print('Features : ', data.columns.values)
data.head()
```

Number of data points : 3321
 Number of features : 4
 Features : ['ID' 'Gene' 'Variation' 'Class']

Out[2]:

	ID	Gene	Variation	Class
0	0	FAM58A	Truncating Mutations	1
1	1	CBL	W802*	2
2	2	CBL	Q249E	2
3	3	CBL	N454D	3
4	4	CBL	L399V	4

training/training_variants is a comma separated file containing the description of the genetic mutations used for training.

Fields are

- **ID** : the id of the row used to link the mutation to the clinical evidence
- **Gene** : the gene where this genetic mutation is located
- **Variation** : the aminoacid change for this mutations
- **Class** : 1-9 the class this genetic mutation has been classified on

3.1.2. Reading Text Data

```
In [3]: # note the separator in this file
data_text = pd.read_csv("training/training_text", sep="\|\|", engine="python", names=["ID", "TEXT"], skiprows=1)
print('Number of data points : ', data_text.shape[0])
print('Number of features : ', data_text.shape[1])
print('Features : ', data_text.columns.values)
data_text.head()
```

Number of data points : 3321
 Number of features : 2
 Features : ['ID' 'TEXT']

Out[3]:

	ID	TEXT
0	0	Cyclin-dependent kinases (CDKs) regulate a var...
1	1	Abstract Background Non-small cell lung canc...
2	2	Abstract Background Non-small cell lung canc...
3	3	Recent evidence has demonstrated that acquired...
4	4	Oncogenic mutations in the monomeric Casitas B...

3.1.3. Preprocessing of text

```
In [4]: # Loading stop words from nltk library
stop_words = set(stopwords.words('english'))

def nlp_preprocessing(total_text, index, column):
    if type(total_text) is not int:
        string = ""
        # replace every special char with space
        total_text = re.sub('[^a-zA-Z0-9\n]', ' ', total_text)
        # replace multiple spaces with single space
        total_text = re.sub('\s+', ' ', total_text)
        # converting all the chars into lower-case.
        total_text = total_text.lower()

        for word in total_text.split():
            # if the word is a not a stop word then retain that word from the data
            if not word in stop_words:
                string += word + " "

        data_text[column][index] = string
```

```
In [5]: #text processing stage.
start_time = time.clock()
for index, row in data_text.iterrows():
    if type(row['TEXT']) is str:
        nlp_preprocessing(row['TEXT'], index, 'TEXT')
    else:
        print("there is no text description for id:",index)
print('Time took for preprocessing the text :',time.clock() - start_time, "seconds")
```

there is no text description for id: 1109
 there is no text description for id: 1277
 there is no text description for id: 1407
 there is no text description for id: 1639
 there is no text description for id: 2755
 Time took for preprocessing the text : 160.110181092 seconds

```
In [6]: #merging both gene_variations and text data based on ID
result = pd.merge(data, data_text, on='ID', how='left')
result.head()
```

Out[6]:

	ID	Gene	Variation	Class	TEXT
0	0	FAM58A	Truncating Mutations	1	cyclin dependent kinases cdks regulate variety...
1	1	CBL	W802*	2	abstract background non small cell lung cancer...
2	2	CBL	Q249E	2	abstract background non small cell lung cancer...
3	3	CBL	N454D	3	recent evidence demonstrated acquired uniparen...
4	4	CBL	L399V	4	oncogenic mutations monomeric casitas b lineage...

```
In [7]: result[result.isnull().any(axis=1)]
```

Out[7]:

	ID	Gene	Variation	Class	TEXT
1109	1109	FANCA	S1088F	1	NaN
1277	1277	ARID5B	Truncating Mutations	1	NaN
1407	1407	FGFR3	K508M	6	NaN
1639	1639	FLT1	Amplification	6	NaN
2755	2755	BRAF	G596C	7	NaN

```
In [8]: result.loc[result['TEXT'].isnull(),'TEXT'] = result['Gene'] + ' '+result['Variation']
```

```
In [9]: result[result['ID']==1109]
```

Out[9]:

	ID	Gene	Variation	Class	TEXT
1109	1109	FANCA	S1088F	1	FANCA S1088F

3.1.4. Test, Train and Cross Validation Split

3.1.4.1. Splitting data into train, test and cross validation (64:20:16)

```
In [10]: y_true = result['Class'].values
result.Gene      = result.Gene.str.replace('\s+', '_')
result.Variation = result.Variation.str.replace('\s+', '_')

# split the data into test and train by maintaining same distribution of output variable 'y_true' [stratify=y_true]
X_train, test_df, y_train, y_test = train_test_split(result, y_true, stratify=y_true, test_size=0.2)
# split the train data into train and cross validation by maintaining same distribution of output variable 'y_train' [stratify=y_train]
train_df, cv_df, y_train, y_cv = train_test_split(X_train, y_train, stratify=y_train, test_size=0.2)
```

```
In [11]: y_train
```

```
Out[11]: array([1, 4, 4, ..., 1, 1, 7], dtype=int64)
```

We split the data into train, test and cross validation data sets, preserving the ratio of class distribution in the original data set

```
In [12]: print('Number of data points in train data:', train_df.shape[0])
print('Number of data points in test data:', test_df.shape[0])
print('Number of data points in cross validation data:', cv_df.shape[0])
```

```
Number of data points in train data: 2124
Number of data points in test data: 665
Number of data points in cross validation data: 532
```

3.1.4.2. Distribution of y_i's in Train, Test and Cross Validation datasets

```
In [13]: pd.__version__
```

```
Out[13]: '0.24.2'
```

```
In [ ]: # X_train = X_train.drop(['Class'],axis=1)
# test_df = test_df.drop(['Class'],axis=1)
# cv_df = cv_df.drop(['Class'],axis=1)
```

3.2 Prediction using a 'Random' Model

In a 'Random' Model, we generate the NINE class probabilites randomly such that they sum to 1.

```
In [15]: # This function plots the confusion matrices given y_i, y_i_hat.

def plot_confusion_matrix(test_y, predict_y):
    C = confusion_matrix(test_y, predict_y)
    # C = 9,9 matrix, each cell (i,j) represents number of points of class i a
    re predicted class j

    A =(((C.T)/(C.sum(axis=1))).T)
    #divid each element of the confusion matrix with the sum of elements in th
    at column

    # C = [[1, 2],
    #      [3, 4]]
    # C.T = [[1, 3],
    #          [2, 4]]
    # C.sum(axis = 1) axis=0 corresonds to columns and axis=1 corresponds to
    rows in two dimensional array
    # C.sum(axis =1) = [[3, 7]]
    # ((C.T)/(C.sum(axis=1))) = [[1/3, 3/7]
    #                               [2/3, 4/7]]

    # ((C.T)/(C.sum(axis=1))).T = [[1/3, 2/3]
    #                               [3/7, 4/7]]
    # sum of row elements = 1

    B =(C/C.sum(axis=0))
    #divid each element of the confusion matrix with the sum of elements in th
    at row

    # C = [[1, 2],
    #      [3, 4]]
    # C.sum(axis = 0) axis=0 corresonds to columns and axis=1 corresponds to
    rows in two dimensional array
    # C.sum(axis =0) = [[4, 6]]
    # (C/C.sum(axis=0)) = [[1/4, 2/6],
    #                      [3/4, 4/6]]

    labels = [1,2,3,4,5,6,7,8,9]
    # representing A in heatmap format
    print("-"*20, "Confusion matrix", "*"-20)
    plt.figure(figsize=(20,7))
    sns.heatmap(C, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y
    ticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.show()

    print("-"*20, "Precision matrix (Column Sum=1)", "*"-20)
    plt.figure(figsize=(20,7))
    sns.heatmap(B, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y
    ticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.show()

    # representing B in heatmap format
    print("-"*20, "Recall matrix (Row sum=1)", "*"-20)
    plt.figure(figsize=(20,7))
```

```
    sns.heatmap(A, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y  
    ticklabels=labels)  
    plt.xlabel('Predicted Class')  
    plt.ylabel('Original Class')  
    plt.show()
```

```
In [16]: # we need to generate 9 numbers and the sum of numbers should be 1
# one solution is to generate 9 numbers and divide each of the numbers by their sum
# ref: https://stackoverflow.com/a/18662466/4084039
test_data_len = test_df.shape[0]
cv_data_len = cv_df.shape[0]

# we create a output array that has exactly same size as the CV data
cv_predicted_y = np.zeros((cv_data_len,9))
for i in range(cv_data_len):
    rand_probs = np.random.rand(1,9)
    cv_predicted_y[i] = ((rand_probs/sum(sum(rand_probs)))[0])
print("Log loss on Cross Validation Data using Random Model",log_loss(y_cv,cv_predicted_y, eps=1e-15))

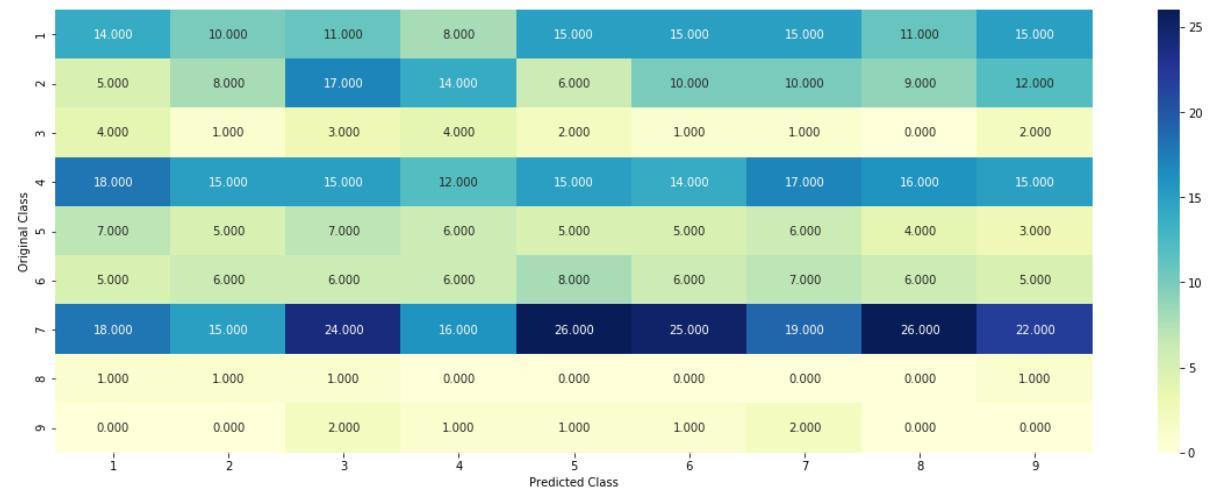
# Test-Set error.
#we create a output array that has exactly same as the test data
test_predicted_y = np.zeros((test_data_len,9))
for i in range(test_data_len):
    rand_probs = np.random.rand(1,9)
    test_predicted_y[i] = ((rand_probs/sum(sum(rand_probs)))[0])
print("Log loss on Test Data using Random Model",log_loss(y_test,test_predicted_y, eps=1e-15))

predicted_y =np.argmax(test_predicted_y, axis=1)
plot_confusion_matrix(y_test, predicted_y+1)
```

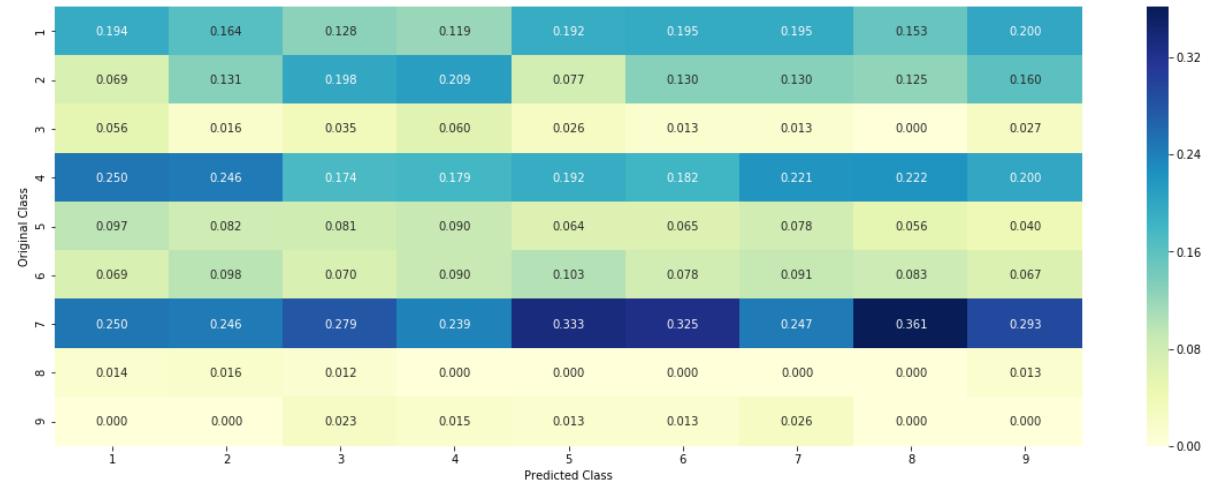
Log loss on Cross Validation Data using Random Model 2.5146188634667106

Log loss on Test Data using Random Model 2.4835270143310155

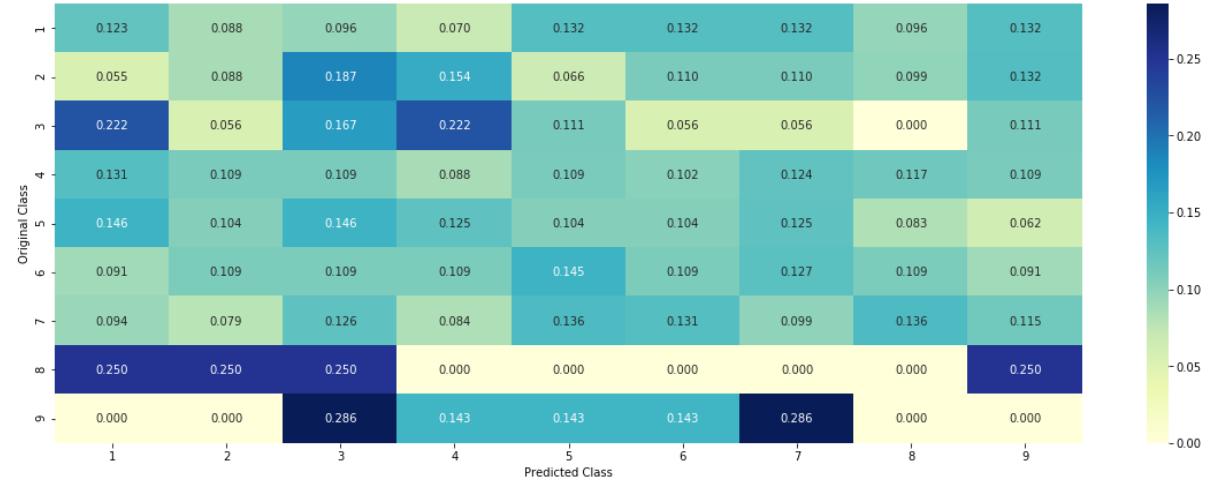
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



3.3 Univariate Analysis


```
In [17]: # code for response coding with Laplace smoothing.
# alpha : used for Laplace smoothing
# feature: ['gene', 'variation']
# df: ['train_df', 'test_df', 'cv_df']
# algorithm
# -----
# Consider all unique values and the number of occurrences of given feature in
# train data dataframe
# build a vector (1*9) , the first element = (number of times it occurred in cl
ass1 + 10*alpha / number of time it occurred in total data+90*alpha)
# gv_dict is like a look up table, for every gene it store a (1*9) representat
ion of it
# for a value of feature in df:
# if it is in train data:
# we add the vector that was stored in 'gv_dict' Look up table to 'gv_fea'
# if it is not there is train:
# we add [1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9] to 'gv_fea'
# return 'gv_fea'
# -----
# get_gv_fea_dict: Get Gene variation Feature Dict
def get_gv_fea_dict(alpha, feature, df):
    # value_count: it contains a dict like
    # print(train_df['Gene'].value_counts())
    # output:
    #     {BRCA1      174
    #      TP53      106
    #      EGFR      86
    #      BRCA2      75
    #      PTEN      69
    #      KIT       61
    #      BRAF      60
    #      ERBB2      47
    #      PDGFRA     46
    #      ...}
    # print(train_df['Variation'].value_counts())
    # output:
    # {
    # Truncating_Mutations      63
    # Deletion                  43
    # Amplification             43
    # Fusions                   22
    # Overexpression            3
    # E17K                      3
    # Q61L                      3
    # S222D                     2
    # P130S                     2
    # ...
    # }
    value_count = train_df[feature].value_counts()

    # gv_dict : Gene Variation Dict, which contains the probability array for
    # each gene/variation
    gv_dict = dict()

    # denominator will contain the number of time that particular feature occu
```

```

red in whole data
for i, denominator in value_count.items():
    # vec will contain ( $p(y_i=1/G_i)$ ) probability of gene/variation belongs
    to particular class
    # vec is 9 dimensional vector
    vec = []
    for k in range(1,10):
        # print(train_df.loc[(train_df['Class']==1) & (train_df['Gene']=
        = 'BRCA1')])]

        #           ID   Gene          Variation Class
        # 2470  2470  BRCA1      S1715C      1
        # 2486  2486  BRCA1      S1841R      1
        # 2614  2614  BRCA1      M1R         1
        # 2432  2432  BRCA1      L1657P      1
        # 2567  2567  BRCA1      T1685A      1
        # 2583  2583  BRCA1      E1660G      1
        # 2634  2634  BRCA1      W1718L      1
        # cls_cnt.shape[0] will return the number of rows

    cls_cnt = train_df.loc[(train_df['Class']==k) & (train_df[feature]==i)]

    # cls_cnt.shape[0](numerator) will contain the number of time that
    # particular feature occurred in whole data
    vec.append((cls_cnt.shape[0] + alpha*10)/ (denominator + 90*alpha))
)

# we are adding the gene/variation to the dict as key and vec as value
gv_dict[i]=vec
return gv_dict

# Get Gene variation feature
def get_gv_feature(alpha, feature, df):
    # print(gv_dict)
    #     {'BRCA1': [0.200757575757575, 0.037878787878788, 0.068181818181818177, 0.13636363636363635, 0.25, 0.1931818181818181, 0.03787878787878788, 0.037878787878788, 0.037878787878788], 'TP53': [0.32142857142857145, 0.061224489795918366, 0.061224489795918366, 0.27040816326530615, 0.061224489795918366, 0.066326530612244902, 0.051020408163265307, 0.051020408163265307, 0.056122448979591837], 'EGFR': [0.0568181818181816, 0.21590909090909091, 0.0625, 0.06818181818177, 0.06818181818177, 0.0625, 0.34659090909090912, 0.0625, 0.0568181818181816], 'BRCA2': [0.1333333333333333, 0.060606060606060608, 0.060606060606060608, 0.0787878787878782, 0.1393939393939394, 0.34545454545454546, 0.060606060606060608, 0.060606060606060608, 0.060606060606060608], 'PTEN': [0.06918238937106917, 0.062893081761006289, 0.069182389937106917, 0.46540880503144655, 0.075471698113207544, 0.062893081761006289, 0.069182389937106917, 0.062893081761006289, 0.062893081761006289], 'KIT': [0.066225165562913912, 0.25165562913907286, 0.072847682119205295, 0.072847682119205295, 0.066225165562913912, 0.066225165562913912, 0.066225165562913912, 0.27152317880794702, 0.066225165562913912, 0.066225165562913912], 'BRAF': [0.066666666666666666, 0.1799999999999999, 0.073333333333333334, 0.07333333333333334, 0.09333333333333338, 0.08000000000000002, 0.2999999999999999, 0.066666666666666666, 0.066666666666666666], '#     ...'
#     }

```

```

gv_dict = get_gv_fea_dict(alpha, feature, df)
# value_count is similar in get_gv_fea_dict
value_count = train_df[feature].value_counts()

# gv_fea: Gene_variation feature, it will contain the feature for each feature value in the data
gv_fea = []
# for every feature values in the given data frame we will check if it is there in the train data then we will add the feature to gv_fea
# if not we will add [1/9,1/9,1/9,1/9,1/9,1/9,1/9,1/9] to gv_fea
for index, row in df.iterrows():
    if row[feature] in dict(value_count).keys():
        gv_fea.append(gv_dict[row[feature]])
    else:
        gv_fea.append([1/9,1/9,1/9,1/9,1/9,1/9,1/9,1/9])
#
return gv_fea

```

when we calculate the probability of a feature belongs to any particular class, we apply laplace smoothing

- $(\text{numerator} + 10^{10} \alpha) / (\text{denominator} + 90^{10} \alpha)$

Q3. How to featurize this Gene feature ?

Ans.there are two ways we can featurize this variable check out this video:

<https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/handling-categorical-and-numerical-features/>

1. One hot Encoding
2. Response coding

We will choose the appropriate featurization based on the ML model we use. For this problem of multi-class classification with categorical features, one-hot encoding is better for Logistic regression while response coding is better for Random Forests.

```

In [22]: #response-coding of the Gene feature
# alpha is used for Laplace smoothing
alpha = 1
# train gene feature
train_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", train_df))
# test gene feature
test_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", test_df))
# cross validation gene feature
cv_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", cv_df))

```

In [23]: `print("train_gene_feature_responseCoding is converted feature using response coding method. The shape of gene feature:", train_gene_feature_responseCoding.shape)`

train_gene_feature_responseCoding is converted feature using response coding method. The shape of gene feature: (2124, 9)

In [25]: `# one-hot encoding of Gene feature.
gene_vectorizer = CountVectorizer()
train_gene_feature_onehotCoding = gene_vectorizer.fit_transform(train_df['Gene'])
test_gene_feature_onehotCoding = gene_vectorizer.transform(test_df['Gene'])
cv_gene_feature_onehotCoding = gene_vectorizer.transform(cv_df['Gene'])`

In [33]: `print("train_gene_feature_onehotCoding is converted feature using one-hot encoding method. The shape of gene feature:", train_gene_feature_onehotCoding.shape)`

train_gene_feature_onehotCoding is converted feature using one-hot encoding method. The shape of gene feature: (2124, 233)

Q9. How to featurize this Variation feature ?

Ans. There are two ways we can featurize this variable check out this video:

<https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/handling-categorical-and-numerical-features/>

1. One hot Encoding
2. Response coding

We will be using both these methods to featurize the Variation Feature

In [27]: `# alpha is used for Laplace smoothing
alpha = 1
train gene feature
train_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", train_df))
test gene feature
test_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", test_df))
cross validation gene feature
cv_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", cv_df))`

In [35]: `print("train_variation_feature_responseCoding is a converted feature using the response coding method. The shape of Variation feature:", train_variation_feature_responseCoding.shape)`

train_variation_feature_responseCoding is a converted feature using the response coding method. The shape of Variation feature: (2124, 9)

```
In [29]: # one-hot encoding of variation feature.
variation_vectorizer = CountVectorizer()
train_variation_feature_onehotCoding = variation_vectorizer.fit_transform(train_df['Variation'])
test_variation_feature_onehotCoding = variation_vectorizer.transform(test_df['Variation'])
cv_variation_feature_onehotCoding = variation_vectorizer.transform(cv_df['Variation'])
```

```
In [34]: print("train_variation_feature_onehotEncoded is converted feature using the one-hot encoding method. The shape of Variation feature:", train_variation_feature_onehotCoding.shape)
```

train_variation_feature_onehotEncoded is converted feature using the one-hot encoding method. The shape of Variation feature: (2124, 1955)

3.2.3 Univariate Analysis on Text Feature

1. How many unique words are present in train data?
2. How are word frequencies distributed?
3. How to featurize text field?
4. Is the text feature useful in predicting y_i?
5. Is the text feature stable across train, test and CV datasets?

```
In [44]: # building a CountVectorizer with all the words that occurred minimum 3 times in train data
text_vectorizer = TfidfVectorizer(lowercase=False,min_df=10)

train_text_feature_tfidf_Coding = text_vectorizer.fit_transform(train_df['TEXT'])
test_text_feature_tfidf_Coding = text_vectorizer.transform(test_df['TEXT'])
cv_text_feature_tfidf_Coding = text_vectorizer.transform(cv_df['TEXT'])
```

```
In [45]: print(train_text_feature_tfidf_Coding.shape)
print(test_text_feature_tfidf_Coding.shape)
print(cv_text_feature_tfidf_Coding.shape)
```

(2124, 25766)
(665, 25766)
(532, 25766)

```
In [47]: from sklearn.feature_selection import SelectKBest
from sklearn.feature_selection import chi2

chi = SelectKBest(chi2,k=1000).fit(train_text_feature_tfidf_Coding, y_train)

train_text_feature_tfidf_Coding = chi.transform(train_text_feature_tfidf_Coding)
test_text_feature_tfidf_Coding = chi.transform(test_text_feature_tfidf_Coding)
cv_text_feature_tfidf_Coding = chi.transform(cv_text_feature_tfidf_Coding)
```

```
In [48]: print(train_text_feature_tfidf_Coding.shape)
print(test_text_feature_tfidf_Coding.shape)
print(cv_text_feature_tfidf_Coding.shape)
```

```
(2124, 1000)
(665, 1000)
(532, 1000)
```

```
In [ ]: #####
```

```
In [ ]: #####
```

```
In [55]: # # don't forget to normalize every feature
# train_text_feature_onehotCoding = normalize(train_text_feature_onehotCoding,
axis=0)

# test_text_feature_onehotCoding = normalize(test_text_feature_onehotCoding, a
xis=0)

# cv_text_feature_onehotCoding = normalize(cv_text_feature_onehotCoding, axis=
0)
```

```
In [1]: import dill
# dill.dump_session('notebook_env_assinm.db')
dill.load_session('notebook_env_assinm.db')
```

```
In [2]: # Train a Logistic regression+Calibration model using text features which are
# on-hot encoded
alpha = [10 ** x for x in range(-5, 1)]

# read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generate
# nered/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_i
# ntercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_
# rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochast
# ic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link:
#-----


cv_log_error_array=[]
for i in alpha:
    clf = SGDClassifier(alpha=i, penalty='l2', loss='log', random_state=42)
    clf.fit(train_text_feature_tfidf_Coding, y_train)

    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_text_feature_tfidf_Coding, y_train)
    predict_y = sig_clf.predict_proba(cv_text_feature_tfidf_Coding)
    cv_log_error_array.append(log_loss(y_cv, predict_y, labels=clf.classes_, e
ps=1e-15))
    print('For values of alpha = ', i, "The log loss is:",log_loss(y_cv, predi
ct_y, labels=clf.classes_, eps=1e-15))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
    ax.annotate((alpha[i],np.round(txt,3)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_
state=42)
clf.fit(train_text_feature_tfidf_Coding, y_train)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_text_feature_tfidf_Coding, y_train)

predict_y = sig_clf.predict_proba(train_text_feature_tfidf_Coding)
```

```
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_text_feature_tfidf_Coding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_text_feature_tfidf_Coding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```
C:\Users\LENOVO\Anaconda3\lib\site-packages\sklearn\model_selection\_split.py:1978: FutureWarning: The default value of cv will change from 3 to 5 in version 0.22. Specify it explicitly to silence this warning.
    warnings.warn(CV_WARNING, FutureWarning)

For values of alpha =  1e-05 The log loss is: 1.1146966037389887

C:\Users\LENOVO\Anaconda3\lib\site-packages\sklearn\model_selection\_split.py:1978: FutureWarning: The default value of cv will change from 3 to 5 in version 0.22. Specify it explicitly to silence this warning.
    warnings.warn(CV_WARNING, FutureWarning)

For values of alpha =  0.0001 The log loss is: 1.0955479347989876

C:\Users\LENOVO\Anaconda3\lib\site-packages\sklearn\model_selection\_split.py:1978: FutureWarning: The default value of cv will change from 3 to 5 in version 0.22. Specify it explicitly to silence this warning.
    warnings.warn(CV_WARNING, FutureWarning)

For values of alpha =  0.001 The log loss is: 1.1909093488192404

C:\Users\LENOVO\Anaconda3\lib\site-packages\sklearn\model_selection\_split.py:1978: FutureWarning: The default value of cv will change from 3 to 5 in version 0.22. Specify it explicitly to silence this warning.
    warnings.warn(CV_WARNING, FutureWarning)

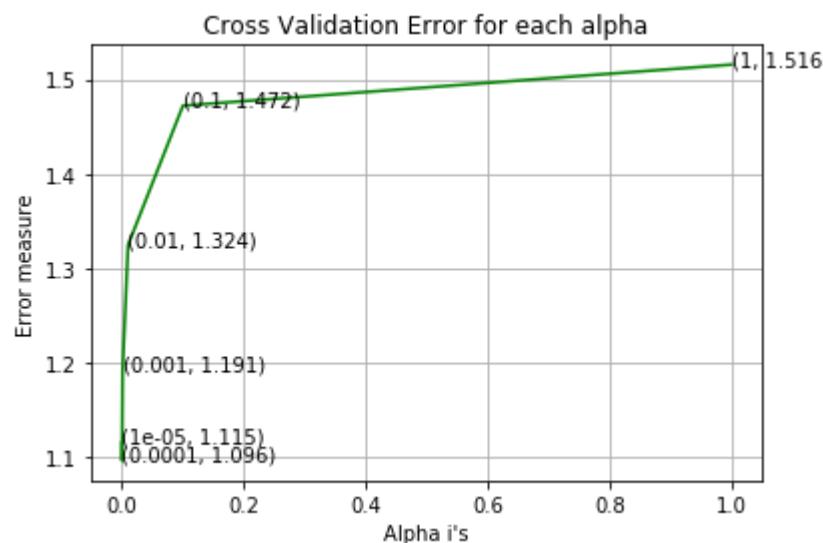
For values of alpha =  0.01 The log loss is: 1.324294615882801

C:\Users\LENOVO\Anaconda3\lib\site-packages\sklearn\model_selection\_split.py:1978: FutureWarning: The default value of cv will change from 3 to 5 in version 0.22. Specify it explicitly to silence this warning.
    warnings.warn(CV_WARNING, FutureWarning)

For values of alpha =  0.1 The log loss is: 1.4721101201548596

C:\Users\LENOVO\Anaconda3\lib\site-packages\sklearn\model_selection\_split.py:1978: FutureWarning: The default value of cv will change from 3 to 5 in version 0.22. Specify it explicitly to silence this warning.
    warnings.warn(CV_WARNING, FutureWarning)

For values of alpha =  1 The log loss is: 1.5156974852366925
```



```
C:\Users\LENOVO\Anaconda3\lib\site-packages\sklearn\model_selection\_split.py:1978: FutureWarning: The default value of cv will change from 3 to 5 in version 0.22. Specify it explicitly to silence this warning.
    warnings.warn(CV_WARNING, FutureWarning)

For values of best alpha =  0.0001 The train log loss is: 0.8572427222921676
For values of best alpha =  0.0001 The cross validation log loss is: 1.0955479347989876
For values of best alpha =  0.0001 The test log loss is: 1.0328635462164693
```

Q. Is the Text feature stable across all the data sets (Test, Train, Cross validation)?

Ans. Yes, it seems like!

4. Machine Learning Models

In [3]: #Data preparation for ML models.

```
#Misc. functionns for ML models

def predict_and_plot_confusion_matrix(train_x, train_y,test_x, test_y, clf):
    clf.fit(train_x, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x, train_y)
    pred_y = sig_clf.predict(test_x)

    # for calculating Log_Loss we will provide the array of probabilities belongs to each class
    print("Log loss :",log_loss(test_y, sig_clf.predict_proba(test_x)))
    # calculating the number of data points that are misclassified
    print("Number of mis-classified points :", np.count_nonzero((pred_y- test_y))/test_y.shape[0])
    plot_confusion_matrix(test_y, pred_y)
```

In [4]: def report_log_loss(train_x, train_y, test_x, test_y, clf):

```
clf.fit(train_x, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x, train_y)
sig_clf_probs = sig_clf.predict_proba(test_x)
return log_loss(test_y, sig_clf_probs, eps=1e-15)
```

```
In [5]: # this function will be used just for naive bayes
# for the given indices, we will print the name of the features
# and we will check whether the feature present in the test point text or not
def get_imptfeature_names(indices, text, gene, var, no_features):
    gene_count_vec = TfidfVectorizer(lowercase=False,min_df=10)
    var_count_vec = TfidfVectorizer(lowercase=False,min_df=10)
    text_count_vec = TfidfVectorizer(lowercase=False,min_df=10)

    gene_vec = gene_count_vec.fit(train_df['Gene'])
    var_vec = var_count_vec.fit(train_df['Variation'])
    text_vec = text_count_vec.fit(train_df['TEXT'])

    fea1_len = len(gene_vec.vocabulary_.keys())
    fea2_len = len(var_count_vec.vocabulary_.keys())

    word_present = 0
    for i,v in enumerate(indices):
        if (v < fea1_len):
            word = gene_vec.get_feature_names()[v]
            yes_no = True if word == gene else False
            if yes_no:
                word_present += 1
                print(i, "Gene feature [{}] present in test data point [{}].format(word,yes_no))
            elif (v < fea1_len+fea2_len):
                word = var_vec.get_feature_names()[v-(fea1_len)]
                yes_no = True if word == var else False
                if yes_no:
                    word_present += 1
                    print(i, "variation feature [{}] present in test data point [{}].format(word,yes_no))
            else:
                word = text_vec.get_feature_names()[v-(fea1_len+fea2_len)]
                yes_no = True if word in text.split() else False
                if yes_no:
                    word_present += 1
                    print(i, "Text feature [{}] present in test data point [{}].format(word,yes_no))

    print("Out of the top ",no_features," features ", word_present, "are present in query point")
```

Stacking the three types of features

In [6]: # merging gene, variance and text features

```
# building train, test and cross validation data sets
# a = [[1, 2],
#       [3, 4]]
# b = [[4, 5],
#       [6, 7]]
# hstack(a, b) = [[1, 2, 4, 5],
#                  [3, 4, 6, 7]]

train_gene_var_onehotCoding = hstack((train_gene_feature_onehotCoding,train_variation_feature_onehotCoding))
test_gene_var_onehotCoding = hstack((test_gene_feature_onehotCoding,test_variation_feature_onehotCoding))
cv_gene_var_onehotCoding = hstack((cv_gene_feature_onehotCoding,cv_variation_feature_onehotCoding))

train_x_onehotCoding = hstack((train_gene_var_onehotCoding, train_text_feature_tfidf_Coding)).tocsr()

test_x_onehotCoding = hstack((test_gene_var_onehotCoding, test_text_feature_tfidf_Coding)).tocsr()

cv_x_onehotCoding = hstack((cv_gene_var_onehotCoding, cv_text_feature_tfidf_Coding)).tocsr()

train_gene_var_responseCoding = np.hstack((train_gene_feature_responseCoding,train_variation_feature_responseCoding))
test_gene_var_responseCoding = np.hstack((test_gene_feature_responseCoding,test_variation_feature_responseCoding))
cv_gene_var_responseCoding = np.hstack((cv_gene_feature_responseCoding,cv_variation_feature_responseCoding))

train_x_responseCoding = hstack((train_gene_var_responseCoding, train_text_feature_tfidf_Coding)).tocsr()
test_x_responseCoding = hstack((test_gene_var_responseCoding, test_text_feature_tfidf_Coding)).tocsr()
cv_x_responseCoding = hstack((cv_gene_var_responseCoding, cv_text_feature_tfidf_Coding)).tocsr()

train_y = np.array(list(train_df['Class']))
test_y = np.array(list(test_df['Class']))
cv_y = np.array(list(cv_df['Class']))
```

```
In [7]: print("One hot encoding features :")
print("(number of data points * number of features) in train data = ", train_x_onehotCoding.shape)
print("(number of data points * number of features) in test data = ", test_x_onehotCoding.shape)
print("(number of data points * number of features) in cross validation data
      =", cv_x_onehotCoding.shape)
```

One hot encoding features :
 (number of data points * number of features) in train data = (2124, 3188)
 (number of data points * number of features) in test data = (665, 3188)
 (number of data points * number of features) in cross validation data = (532, 3188)

```
In [8]: print(" Response encoding features :")
print("(number of data points * number of features) in train data = ", train_x_responseCoding.shape)
print("(number of data points * number of features) in test data = ", test_x_responseCoding.shape)
print("(number of data points * number of features) in cross validation data
      =", cv_x_responseCoding.shape)
```

Response encoding features :
 (number of data points * number of features) in train data = (2124, 1018)
 (number of data points * number of features) in test data = (665, 1018)
 (number of data points * number of features) in cross validation data = (532, 1018)

```
In [9]: # https://machineLearningmastery.com/how-to-fix-futurewarning-messages-in-scikit-learn/
from warnings import simplefilter
# ignore all future warnings
simplefilter(action='ignore', category=FutureWarning)
```

4.1. Base Line Model

4.1.1. Naive Bayes

4.1.1.1. Hyper parameter tuning

```
In [10]: # find more about Multinomial Naive base function here http://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html
# -----
# default paramters
# sklearn.naive_bayes.MultinomialNB(alpha=1.0, fit_prior=True, class_prior=None)

# some of methods of MultinomialNB()
# fit(X, y[, sample_weight]) Fit Naive Bayes classifier according to X, y
# predict(X) Perform classification on an array of test vectors X.
# predict_log_proba(X) Return log-probability estimates for the test vector X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/naive-bayes-algorithm-1/
# -----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/naive-bayes-algorithm-1/
# -----


alpha = [0.00001, 0.0001, 0.001, 0.1, 1, 10, 100, 1000]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = MultinomialNB(alpha=i)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
        # to avoid rounding error while multiplying probabilites we use log-probability estimates
    print("Log Loss :", log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(np.log10(alpha), cv_log_error_array, c='g')
for i, txt in enumerate(np.round(cv_log_error_array, 3)):
    ax.annotate((alpha[i], str(txt)), (np.log10(alpha[i]), cv_log_error_array[i]))
```

```
plt.grid()
plt.xticks(np.log10(alpha))
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

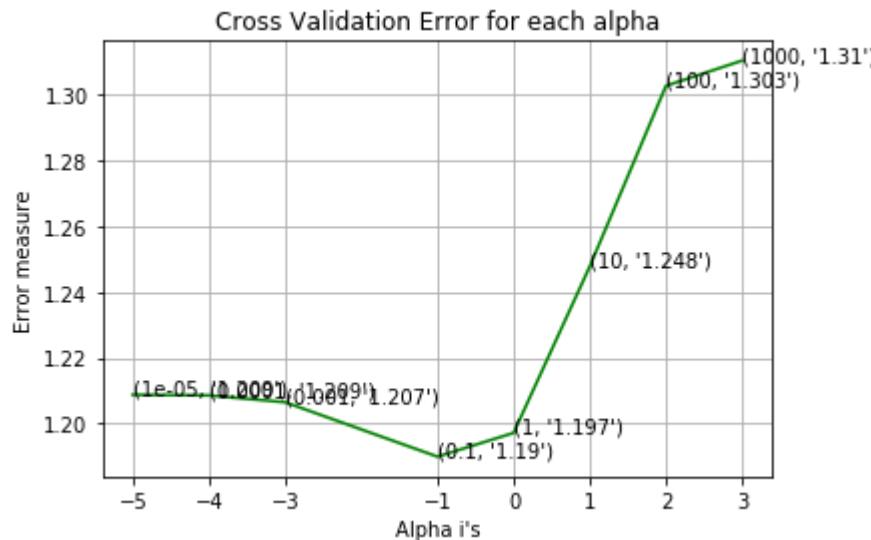
best_alpha = np.argmin(cv_log_error_array)
clf = MultinomialNB(alpha=alpha[best_alpha])
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 1e-05
Log Loss : 1.2087583883449589
for alpha = 0.0001
Log Loss : 1.2085791735509288
for alpha = 0.001
Log Loss : 1.2065022603444817
for alpha = 0.1
Log Loss : 1.1899587367794495
for alpha = 1
Log Loss : 1.1971761332205408
for alpha = 10
Log Loss : 1.2477813584455477
for alpha = 100
Log Loss : 1.3026677985165849
for alpha = 1000
Log Loss : 1.310334579455075

```



For values of best alpha = 0.1 The train log loss is: 0.6781816986617645
 For values of best alpha = 0.1 The cross validation log loss is: 1.1899587367794495
 For values of best alpha = 0.1 The test log loss is: 1.1175779226165785

4.1.1.2. Testing the model with best hyper parameters

```
In [11]: # find more about Multinomial Naive base function here http://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html
# -----
# default paramters
# sklearn.naive_bayes.MultinomialNB(alpha=1.0, fit_prior=True, class_prior=None)

# some of methods of MultinomialNB()
# fit(X, y[, sample_weight]) Fit Naive Bayes classifier according to X, y
# predict(X) Perform classification on an array of test vectors X.
# predict_log_proba(X) Return log-probability estimates for the test vector X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/naive-bayes-algorithm-1/
# -----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
# -----

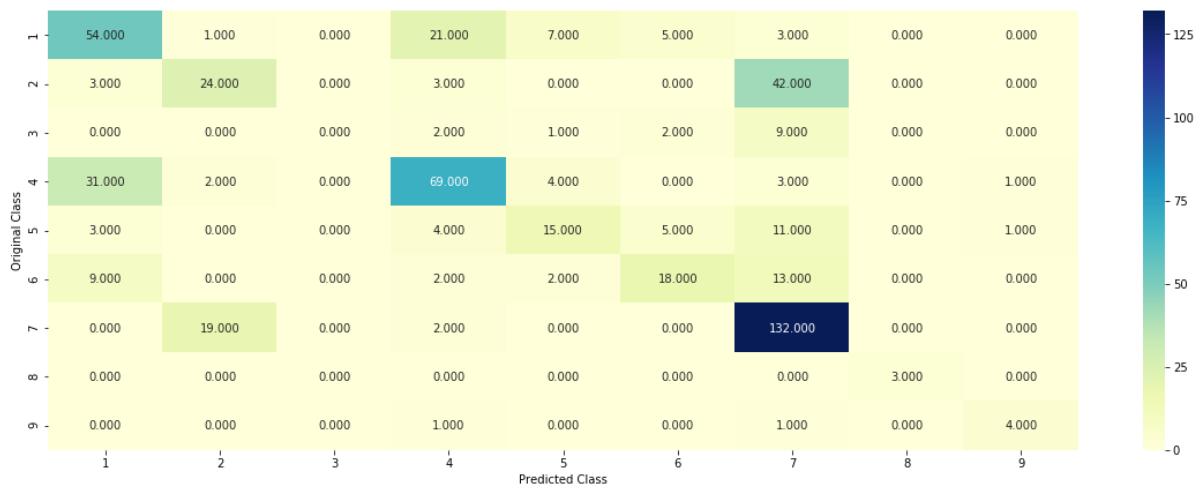

clf = MultinomialNB(alpha=alpha[best_alpha])
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)
sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
# to avoid rounding error while multiplying probabilit
y estimates
print("Log Loss :",log_loss(cv_y, sig_clf_probs))
print("Number of missclassified point :", np.count_nonzero((sig_clf.predict(cv_x_onehotCoding)- cv_y))/cv_y.shape[0])
plot_confusion_matrix(cv_y, sig_clf.predict(cv_x_onehotCoding.toarray()))
```

Log Loss : 1.1899587367794495

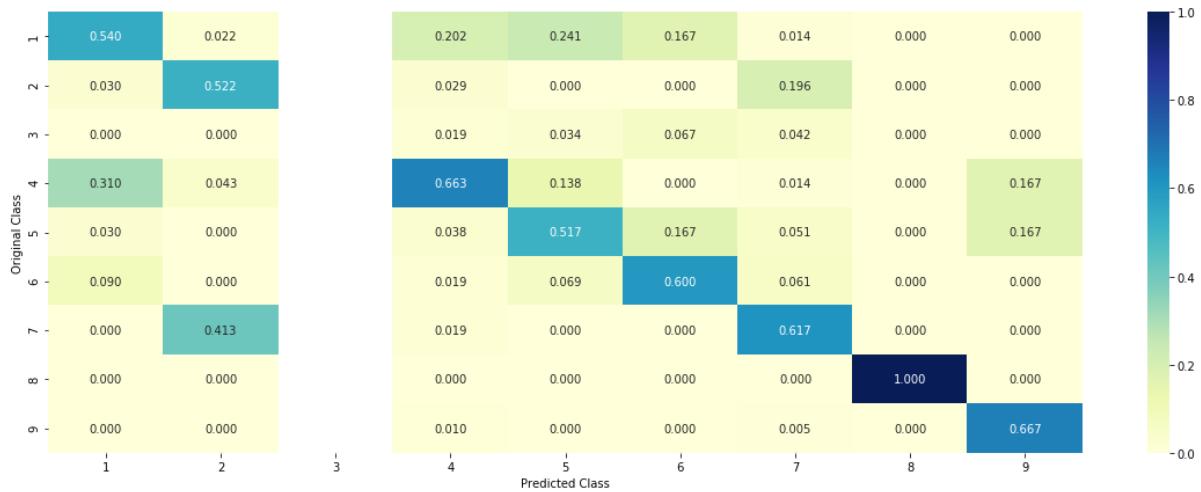
Number of missclassified point : 0.40037593984962405

----- Confusion matrix -----

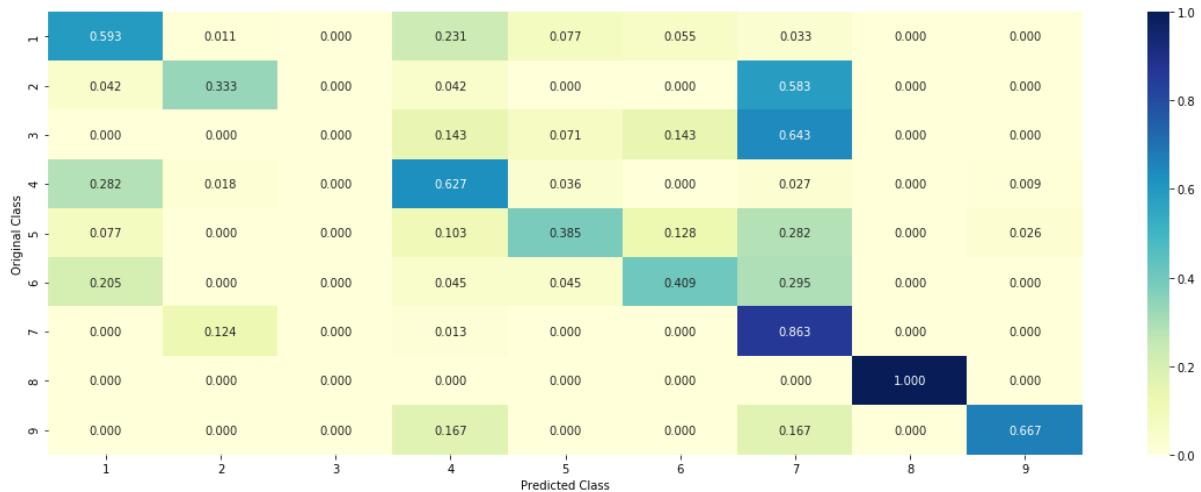
C:\Users\LENOVO\Anaconda3\lib\site-packages\ipykernel_launcher.py:22: RuntimeWarning: invalid value encountered in true_divide



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.1.1.3. Feature Importance, Correctly classified point

```
In [12]: test_point_index = 50
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.coef_)[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)

Predicted Class : 7
Predicted Class Probabilities: [[0.0515 0.0428 0.0182 0.0673 0.0386 0.0295 0.
7449 0.0039 0.0032]]
Actual Class : 7
-----
13 Text feature [021] present in test data point [True]
20 Text feature [012] present in test data point [True]
43 Text feature [activated] present in test data point [True]
52 Text feature [accumulation] present in test data point [True]
74 Text feature [acquisition] present in test data point [True]
78 Text feature [activating] present in test data point [True]
83 Text feature [af9] present in test data point [True]
Out of the top 100 features 7 are present in query point
```

4.1.1.4. Feature Importance, Incorrectly classified point

```
In [13]: test_point_index = 59
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.coef_)[predicted_cls-1][:,:no_feature]
print("-"*50)
get_impfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

```
Predicted Class : 1
Predicted Class Probabilities: [[0.4816 0.047  0.021  0.2973 0.0448 0.0346 0.
0657 0.0044 0.0037]]
Actual Class : 4
-----
16 Text feature [advantage] present in test data point [True]
25 Text feature [024] present in test data point [True]
82 Text feature [activate] present in test data point [True]
Out of the top 100 features 3 are present in query point
```

4.2. K Nearest Neighbour Classification

4.2.1. Hyper parameter tuning

```
In [14]: # find more about KNeighborsClassifier() here http://scikit-Learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html
# -----
# default parameter
# KNeighborsClassifier(n_neighbors=5, weights='uniform', algorithm='auto', leaf_size=30, p=2,
# metric='minkowski', metric_params=None, n_jobs=1, **kwargs)

# methods of
# fit(X, y) : Fit the model using X as training data and y as target values
# predict(X):Predict the class labels for the provided data
# predict_proba(X):Return probability estimates for the test data X.
#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/k-nearest-neighbors-geometric-intuition-with-a-toy-example-1/
#-----

# find more about CalibratedClassifierCV here at http://scikit-Learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
#-----
# video link:
#-----
```

```
alpha = [5, 11, 15, 21, 31, 41, 51, 99]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = KNeighborsClassifier(n_neighbors=i)
    clf.fit(train_x_responseCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_responseCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_responseCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
    # to avoid rounding error while multiplying probabilites we use log-probability estimates
    print("Log Loss :",log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
    ax.annotate((alpha[i],str(txt)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
```

```
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

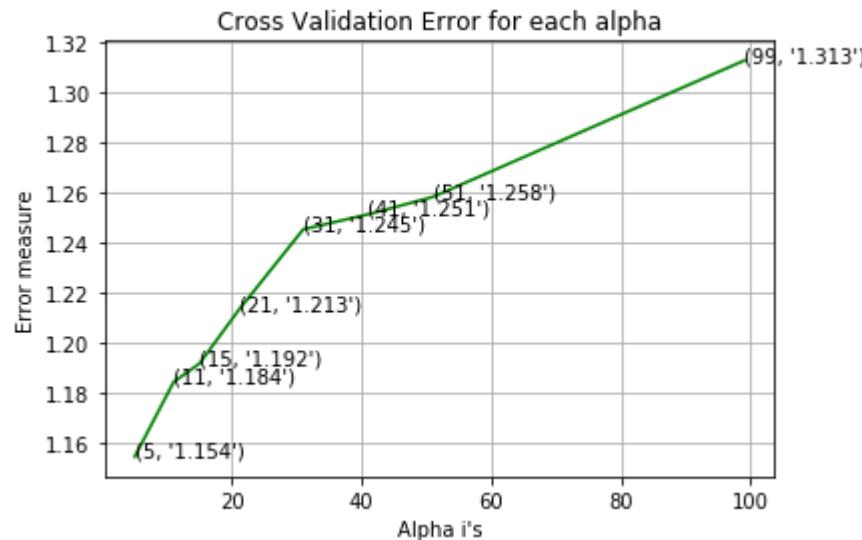
best_alpha = np.argmin(cv_log_error_array)
clf = KNeighborsClassifier(n_neighbors=alpha[best_alpha])
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_responseCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_responseCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_responseCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 5
Log Loss : 1.154457918476012
for alpha = 11
Log Loss : 1.1843099450827188
for alpha = 15
Log Loss : 1.1917225487968377
for alpha = 21
Log Loss : 1.2129716886466784
for alpha = 31
Log Loss : 1.2453906774329975
for alpha = 41
Log Loss : 1.251361808884378
for alpha = 51
Log Loss : 1.2583366961424436
for alpha = 99
Log Loss : 1.3128478899090807

```



For values of best alpha = 5 The train log loss is: 0.8257296138338655
 For values of best alpha = 5 The cross validation log loss is: 1.154457918476012
 For values of best alpha = 5 The test log loss is: 1.0796826950505716

4.2.2. Testing the model with best hyper parameters

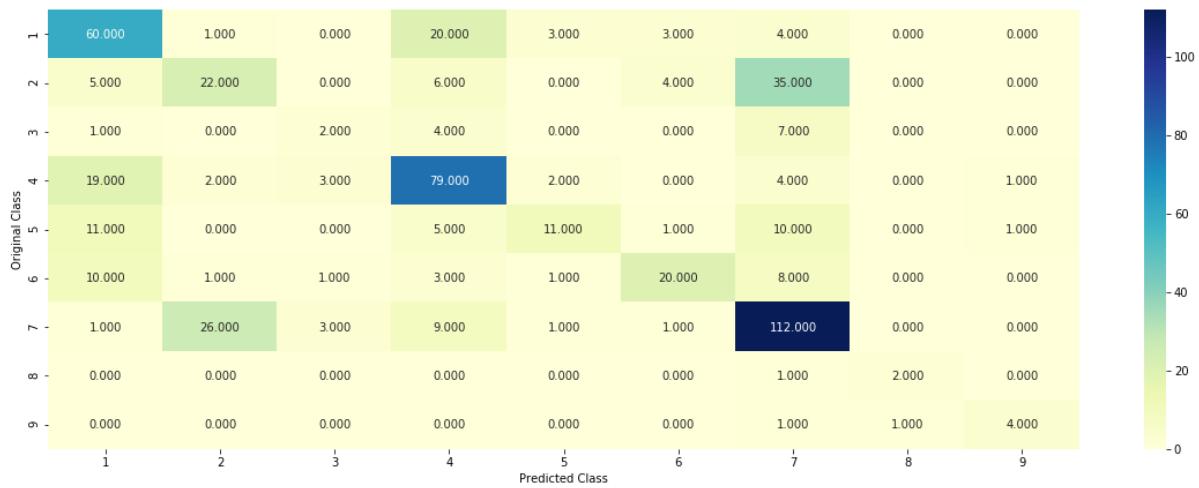
```
In [15]: # find more about KNeighborsClassifier() here http://scikit-Learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html
# -----
# default parameter
# KNeighborsClassifier(n_neighbors=5, weights='uniform', algorithm='auto', leaf_size=30, p=2,
# metric='minkowski', metric_params=None, n_jobs=1, **kwargs)

# methods of
# fit(X, y) : Fit the model using X as training data and y as target values
# predict(X):Predict the class labels for the provided data
# predict_proba(X):Return probability estimates for the test data X.
#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/k-nearest-neighbors-geometric-intuition-with-a-toy-example-1/
#-----
clf = KNeighborsClassifier(n_neighbors=alpha[best_alpha])
predict_and_plot_confusion_matrix(train_x_responseCoding, train_y, cv_x_responseCoding, cv_y, clf)
```

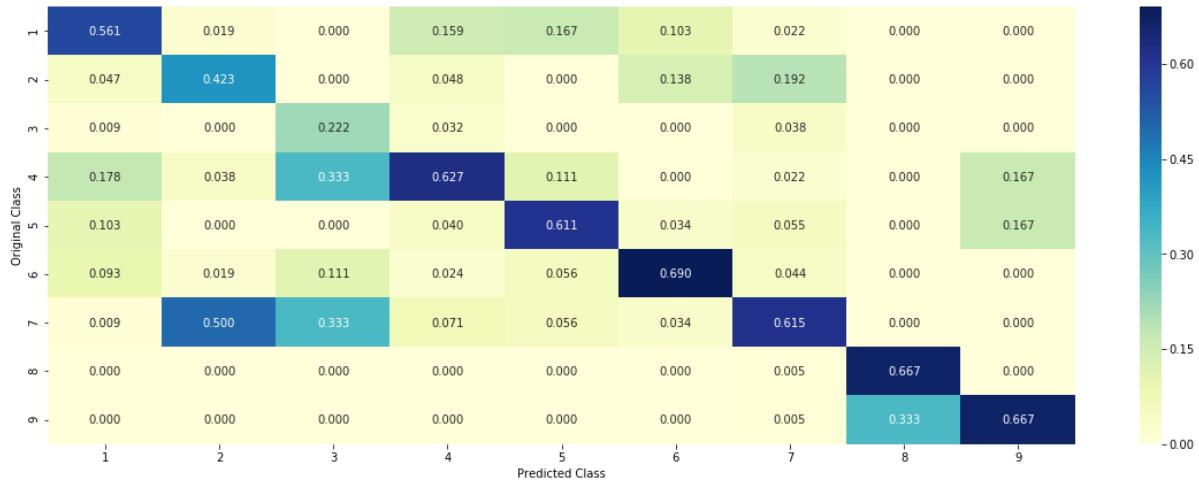
Log loss : 1.154457918476012

Number of mis-classified points : 0.41353383458646614

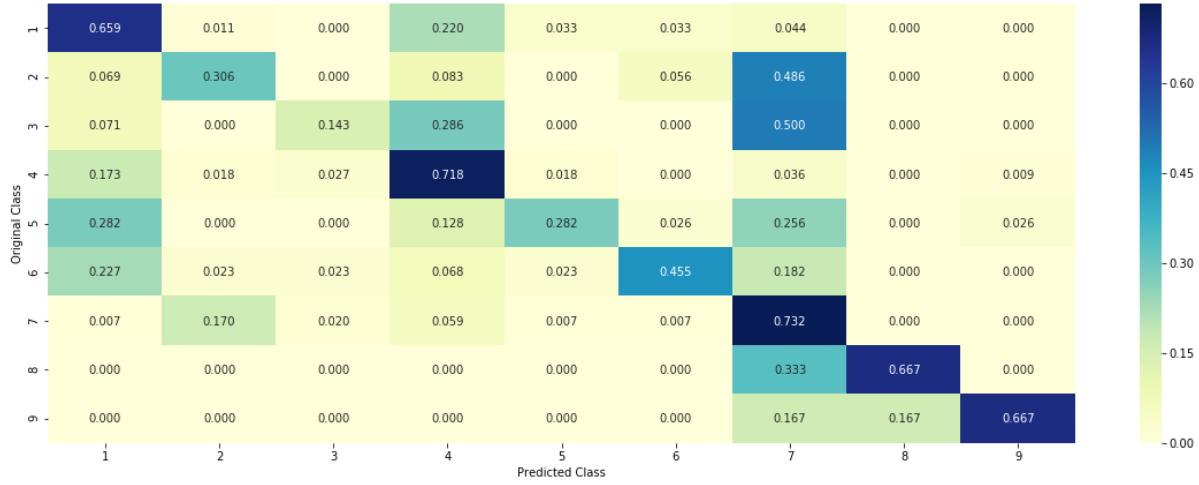
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.2.3.Sample Query point -1

```
In [16]: clf = KNeighborsClassifier(n_neighbors=alpha[best_alpha])
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

test_point_index = 0
predicted_cls = sig_clf.predict(test_x_responseCoding[test_point_index].reshape(1,-1))
print("Predicted Class :", predicted_cls[0])
print("Actual Class :", test_y[test_point_index])
neighbors = clf.kneighbors(test_x_responseCoding[test_point_index].reshape(1,-1), alpha[best_alpha])
print("The ",alpha[best_alpha]," nearest neighbours of the test points belongs to classes",train_y[neighbors[1][0]])
print("Frequency of nearest points :",Counter(train_y[neighbors[1][0]]))
```

Predicted Class : 4
 Actual Class : 4
 The 5 nearest neighbours of the test points belongs to classes [4 4 4 4 4]
 Frequency of nearest points : Counter({4: 5})

```
In [17]: neighbors
```

```
Out[17]: (array([[0.01036054, 0.11967768, 0.17971449, 0.26519249, 0.72975908]]),
 array([[1651, 331, 1044, 1589, 1797]], dtype=int64))
```

```
In [18]: train_y[461]
```

```
Out[18]: 4
```

4.2.4. Sample Query Point-2

```
In [19]: clf = KNeighborsClassifier(n_neighbors=alpha[best_alpha])
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

test_point_index = 59

predicted_cls = sig_clf.predict(test_x_responseCoding[test_point_index].reshape(1,-1))
print("Predicted Class :", predicted_cls[0])
print("Actual Class :", test_y[test_point_index])
neighbors = clf.kneighbors(test_x_responseCoding[test_point_index].reshape(1,-1), alpha[best_alpha])
print("the k value for knn is",alpha[best_alpha],"and the nearest neighbours of the test points belongs to classes",train_y[neighbors[1][0]])
print("Frequency of nearest points :",Counter(train_y[neighbors[1][0]]))
```

```
Predicted Class : 1
Actual Class : 4
the k value for knn is 5 and the nearest neighbours of the test points belongs to classes [4 1 1 1 4]
Frequency of nearest points : Counter({1: 3, 4: 2})
```

4.3. Logistic Regression

4.3.1. With Class balancing

4.3.1.1. Hyper parameter tuning

```
In [20]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight])      Fit the calibrated model
# get_params([deep])      Get parameters for this estimator.
# predict(X)      Predict the target of new samples.
# predict_proba(X)    Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10 ** x for x in range(-6, 3)]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = SGDClassifier(class_weight='balanced', alpha=i, penalty='L2', loss='log', random_state=42)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
        # to avoid rounding error while multiplying probabilites we use log-probability estimates
    print("Log Loss :",log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
```

```
        ax.annotate((alpha[i],str(txt)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

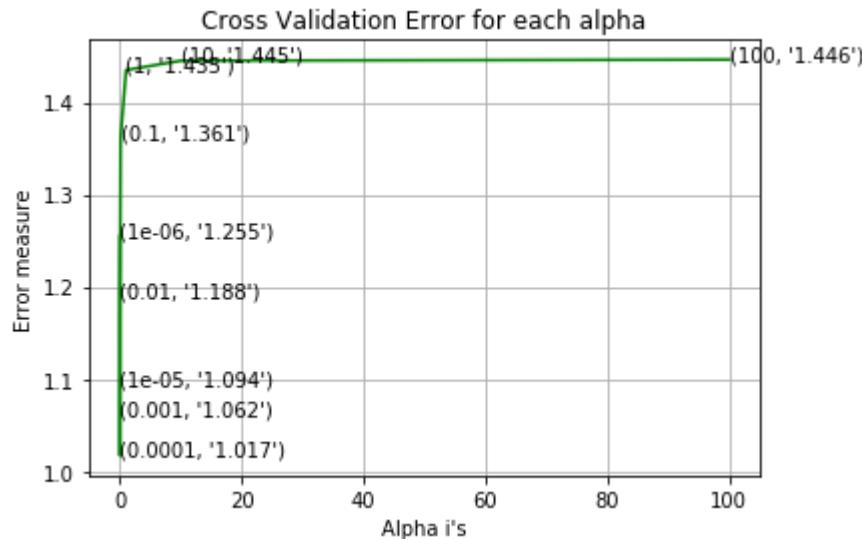
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty=
'12', loss='log', random_state=42)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss i
s:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation
log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss i
s:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 1e-06
Log Loss : 1.2548949178601423
for alpha = 1e-05
Log Loss : 1.0941780576381863
for alpha = 0.0001
Log Loss : 1.0171201996195784
for alpha = 0.001
Log Loss : 1.0615872355587639
for alpha = 0.01
Log Loss : 1.188332168904587
for alpha = 0.1
Log Loss : 1.3612462702412047
for alpha = 1
Log Loss : 1.4348393373182824
for alpha = 10
Log Loss : 1.4449339770970664
for alpha = 100
Log Loss : 1.4461522237437494

```



For values of best alpha = 0.0001 The train log loss is: 0.4435971612892727
 For values of best alpha = 0.0001 The cross validation log loss is: 1.0171201996195784
 For values of best alpha = 0.0001 The test log loss is: 0.9339972774275132

4.3.1.2. Testing the model with best hyper paramters

```
In [21]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)    Predict class labels for samples in X.

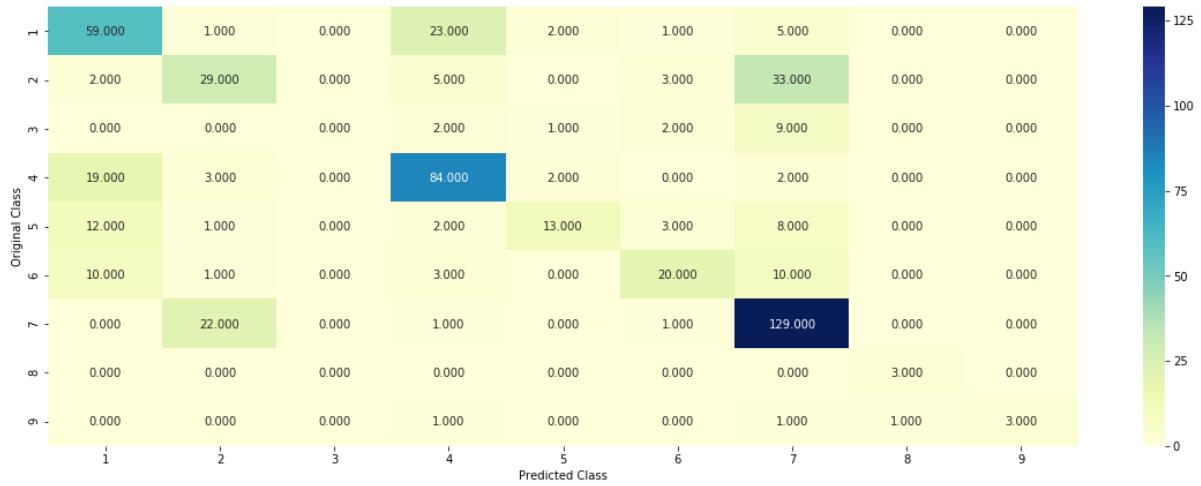
#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCoding, cv_y, clf)
```

Log loss : 1.0171201996195784

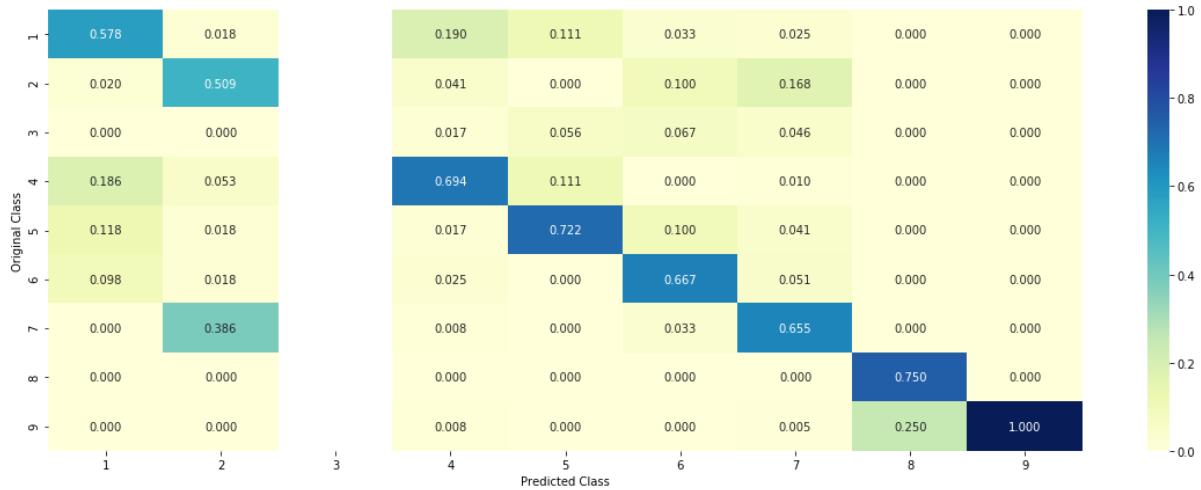
Number of mis-classified points : 0.3609022556390977

----- Confusion matrix -----

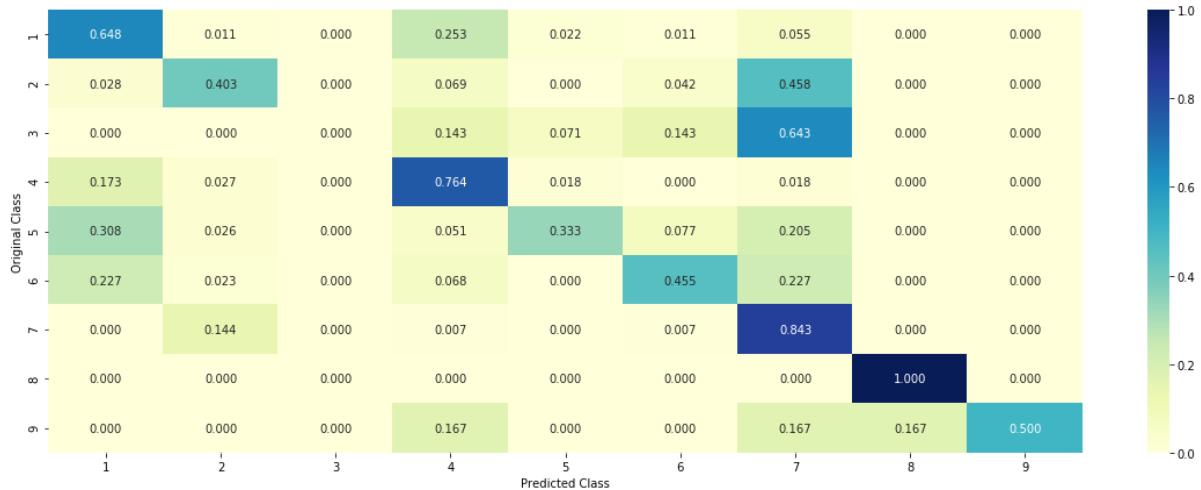
C:\Users\LENOVO\Anaconda3\lib\site-packages\ipykernel_launcher.py:22: RuntimeWarning: invalid value encountered in true_divide



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.3.1.3. Feature Importance

```
In [22]: def get_imp_feature_names(text, indices, removed_ind = []):
    word_present = 0
    tabulte_list = []
    incresingorder_ind = 0
    for i in indices:
        if i < train_gene_feature_onehotCoding.shape[1]:
            tabulte_list.append([incresingorder_ind, "Gene", "Yes"])
        elif i < 18:
            tabulte_list.append([incresingorder_ind, "Variation", "Yes"])
        if ((i > 17) & (i not in removed_ind)):
            word = train_text_features[i]
            yes_no = True if word in text.split() else False
            if yes_no:
                word_present += 1
            tabulte_list.append([incresingorder_ind, train_text_features[i], yes_no])
            incresingorder_ind += 1
    print(word_present, "most important features are present in our query point")
    print("-"*50)
    print("The features that are most importent of the ",predicted_cls[0]," class:")
    print (tabulate(tabulte_list, headers=["Index", 'Feature name', 'Present or Not']))
```

In [23]: (indices).shape

Out[23]: (1, 100)

4.3.1.3.1. Correctly Classified point

```
In [26]: # from tabulate import tabulate
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
clf.fit(train_x_onehotCoding,train_y)
test_point_index = 0
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class : ", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_onehotCoding[test_point_index]),4))
print("Actual Class : ", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imffeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)
```

```
Predicted Class : 4
Predicted Class Probabilities: [[0.0931 0.0399 0.0157 0.7287 0.0299 0.0254 0.
0521 0.0069 0.0082]]
Actual Class : 4
-----
10 Text feature [21] present in test data point [True]
72 Text feature [3d] present in test data point [True]
188 Text feature [agents] present in test data point [True]
196 Text feature [13] present in test data point [True]
215 Text feature [15] present in test data point [True]
227 Text feature [150] present in test data point [True]
277 Text feature [12] present in test data point [True]
330 Text feature [45] present in test data point [True]
344 Text feature [1640] present in test data point [True]
385 Text feature [3b] present in test data point [True]
404 Text feature [acid] present in test data point [True]
429 Text feature [4a] present in test data point [True]
486 Text feature [600] present in test data point [True]
Out of the top 500 features 13 are present in query point
```

4.3.1.3.2. Incorrectly Classified point

```
In [27]: test_point_index = 361
no_feature = 59
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_impfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

```
Predicted Class : 5
Predicted Class Probabilities: [[0.1264 0.0035 0.0246 0.2956 0.4619 0.082 0.
0023 0.0025 0.0011]]
Actual Class : 4
-----
0 Text feature [2013] present in test data point [True]
18 Text feature [1863] present in test data point [True]
27 Text feature [560] present in test data point [True]
30 Text feature [54] present in test data point [True]
39 Text feature [350] present in test data point [True]
54 Text feature [5411t] present in test data point [True]
Out of the top 59 features 6 are present in query point
```

4.3.2. Without Class balancing

4.3.2.1. Hyper parameter tuning

```
In [28]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight])  Fit the calibrated model
# get_params([deep])  Get parameters for this estimator.
# predict(X)  Predict the target of new samples.
# predict_proba(X)  Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10 ** x for x in range(-6, 1)]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = SGDClassifier(alpha=i, penalty='l2', loss='log', random_state=42)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
    print("Log Loss :", log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array, c='g')
for i, txt in enumerate(np.round(cv_log_error_array, 3)):
    ax.annotate((alpha[i], str(txt)), (alpha[i], cv_log_error_array[i]))
plt.grid()
```

```
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

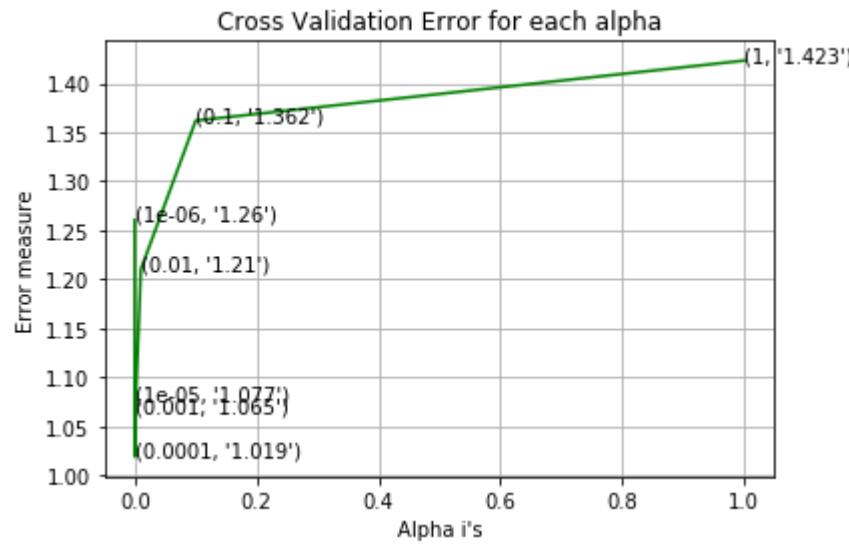
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_
state=42)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss i
s:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation
log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss i
s:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 1e-06
Log Loss : 1.2602895685857307
for alpha = 1e-05
Log Loss : 1.0773855849840543
for alpha = 0.0001
Log Loss : 1.018905024009811
for alpha = 0.001
Log Loss : 1.0650038196269849
for alpha = 0.01
Log Loss : 1.2099656135627903
for alpha = 0.1
Log Loss : 1.3618312105792558
for alpha = 1
Log Loss : 1.4230490362508987

```



For values of best alpha = 0.0001 The train log loss is: 0.4434317778813003
 For values of best alpha = 0.0001 The cross validation log loss is: 1.018905024009811
 For values of best alpha = 0.0001 The test log loss is: 0.9354344951768131

4.3.2.2. Testing model with best hyper parameters

```
In [29]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)    Predict class labels for samples in X.

#-----
# video link:
#-----

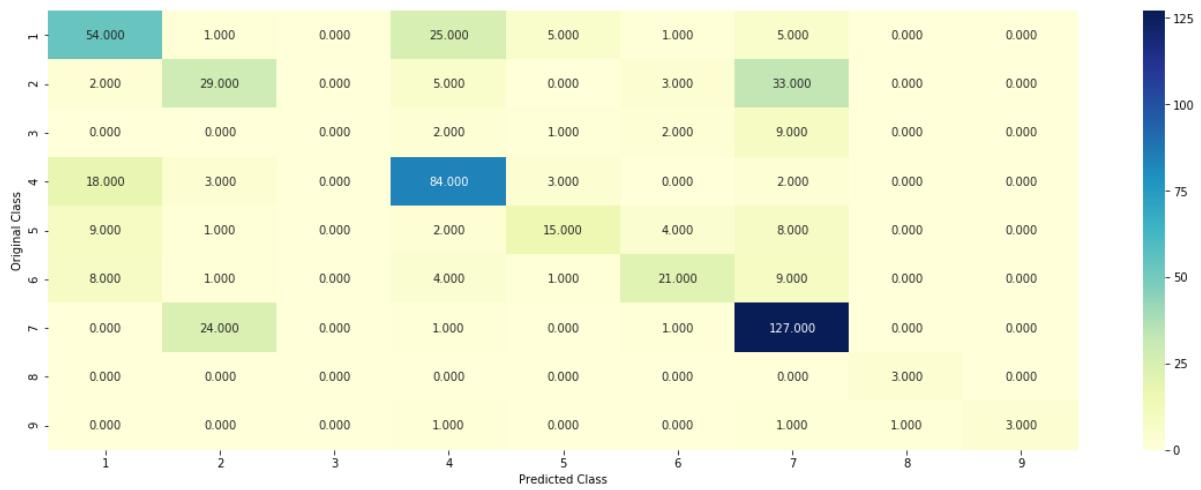
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCoding, cv_y, clf)
```

Log loss : 1.018905024009811

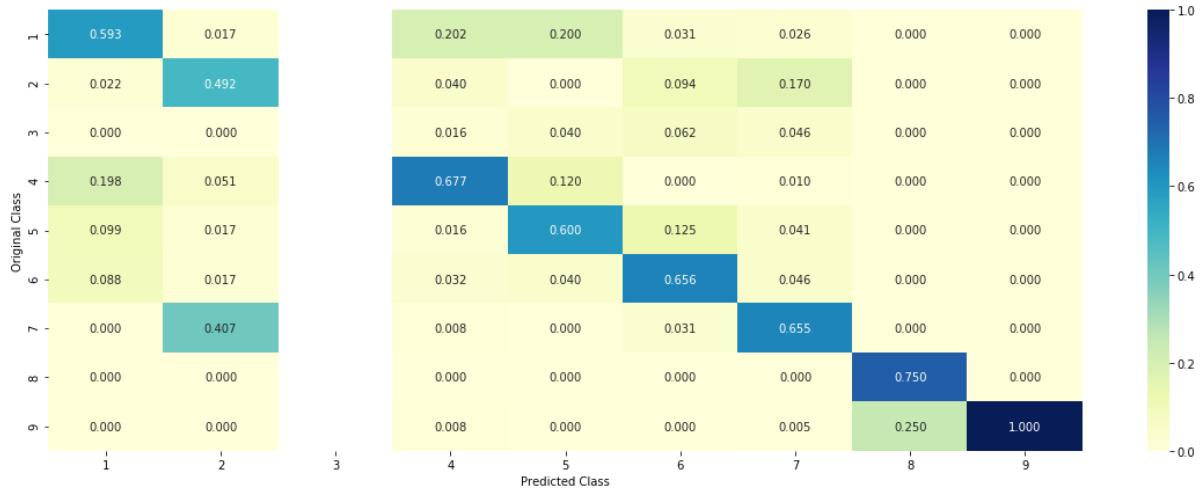
Number of mis-classified points : 0.3684210526315789

----- Confusion matrix -----

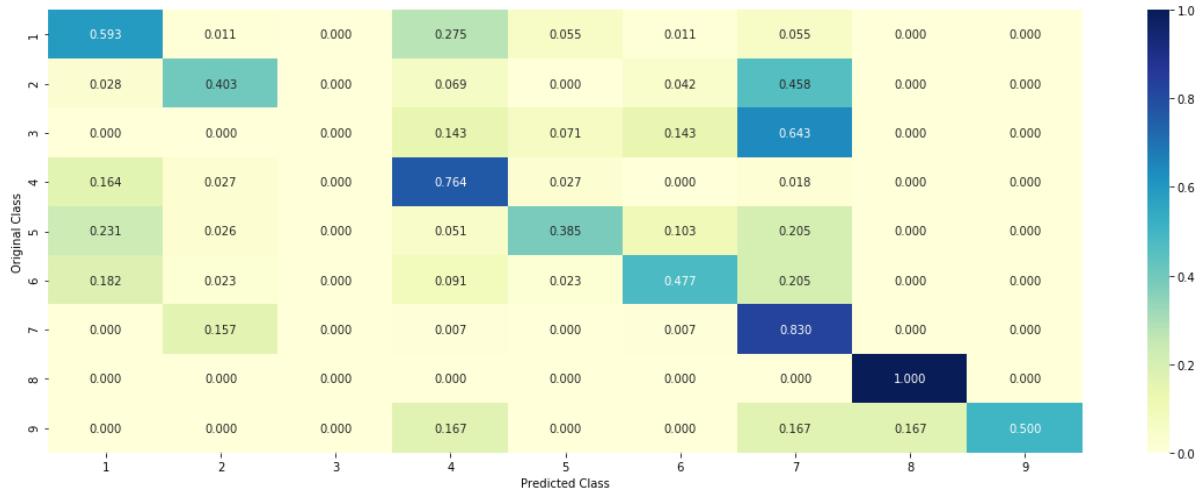
C:\Users\LENOVO\Anaconda3\lib\site-packages\ipykernel_launcher.py:22: RuntimeWarning: invalid value encountered in true_divide



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.3.2.3. Feature Importance, Correctly Classified point

```
In [36]: clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_
state=42)
clf.fit(train_x_onehotCoding,train_y)
test_point_index = 50
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

Predicted Class : 7
Predicted Class Probabilities: [[0.0146 0.0674 0.0041 0.0548 0.0164 0.0205 0.
8178 0.0025 0.002]]
Actual Class : 7

2 Text feature [1970s] present in test data point [True]
7 Text feature [189] present in test data point [True]
16 Text feature [6b] present in test data point [True]
41 Text feature [122] present in test data point [True]
48 Text feature [65] present in test data point [True]
56 Text feature [110] present in test data point [True]
67 Text feature [021] present in test data point [True]
80 Text feature [161] present in test data point [True]
84 Text feature [207] present in test data point [True]
89 Text feature [140] present in test data point [True]
102 Text feature [116] present in test data point [True]
142 Text feature [33] present in test data point [True]
162 Text feature [204] present in test data point [True]
187 Text feature [66] present in test data point [True]
195 Text feature [2a] present in test data point [True]
203 Text feature [141] present in test data point [True]
204 Text feature [152] present in test data point [True]
216 Text feature [124] present in test data point [True]
223 Text feature [400] present in test data point [True]
232 Text feature [250] present in test data point [True]
236 Text feature [100] present in test data point [True]
246 Text feature [012] present in test data point [True]
259 Text feature [adaptive] present in test data point [True]
269 Text feature [188] present in test data point [True]
282 Text feature [26] present in test data point [True]
295 Text feature [3e] present in test data point [True]
296 Text feature [93] present in test data point [True]
306 Text feature [17] present in test data point [True]
315 Text feature [29] present in test data point [True]
329 Text feature [159] present in test data point [True]
335 Text feature [14] present in test data point [True]
347 Text feature [589] present in test data point [True]
356 Text feature [210] present in test data point [True]
361 Text feature [106] present in test data point [True]
379 Text feature [56] present in test data point [True]
385 Text feature [64] present in test data point [True]
394 Text feature [53] present in test data point [True]
404 Text feature [1985] present in test data point [True]
417 Text feature [212] present in test data point [True]
423 Text feature [190] present in test data point [True]
424 Text feature [70] present in test data point [True]
430 Text feature [accomplished] present in test data point [True]
440 Text feature [142] present in test data point [True]
441 Text feature [25] present in test data point [True]
448 Text feature [18] present in test data point [True]
454 Text feature [213] present in test data point [True]
475 Text feature [163] present in test data point [True]
476 Text feature [137] present in test data point [True]
482 Text feature [203] present in test data point [True]
493 Text feature [42] present in test data point [True]
Out of the top 500 features 50 are present in query point

4.3.2.4. Feature Importance, Inorrectly Classified point

```
In [37]: test_point_index = 125
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d_
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)

Predicted Class : 4
Predicted Class Probabilities: [[0.3829 0.025  0.0135 0.4403 0.0249 0.0236 0.
0793 0.0053 0.0053]]
Actual Class : 1
-----
58 Text feature [3d] present in test data point [True]
88 Text feature [1997] present in test data point [True]
150 Text feature [2007] present in test data point [True]
157 Text feature [advantage] present in test data point [True]
189 Text feature [150] present in test data point [True]
192 Text feature [72] present in test data point [True]
203 Text feature [2010] present in test data point [True]
228 Text feature [12] present in test data point [True]
323 Text feature [6d] present in test data point [True]
329 Text feature [04] present in test data point [True]
338 Text feature [2013] present in test data point [True]
383 Text feature [3b] present in test data point [True]
408 Text feature [4a] present in test data point [True]
426 Text feature [2011] present in test data point [True]
428 Text feature [2009] present in test data point [True]
438 Text feature [actions] present in test data point [True]
448 Text feature [09] present in test data point [True]
449 Text feature [49] present in test data point [True]
450 Gene feature [PIK3R1] present in test data point [True]
Out of the top 500 features 19 are present in query point
```

4.4. Linear Support Vector Machines

4.4.1. Hyper parameter tuning

```
In [38]: # read more about support vector machines with Linear kernels here http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

# -----
# default parameters
# SVC(C=1.0, kernel='rbf', degree=3, gamma='auto', coef0=0.0, shrinking=True,
# probability=False, tol=0.001,
# cache_size=200, class_weight=None, verbose=False, max_iter=-1, decision_function_shape='ovr', random_state=None)

# Some of methods of SVM()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given training data.
# predict(X) Perform classification on samples in X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/mathematical-derivation-copy-8/
# -----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10 ** x for x in range(-5, 3)]
cv_log_error_array = []
for i in alpha:
    print("for C =", i)
    # clf = SVC(C=i,kernel='linear',probability=True, class_weight='balanced')
    clf = SGDClassifier(class_weight='balanced', alpha=i, penalty='l2', loss='hinge', random_state=42)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
    print("Log Loss :", log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array, c='g')
for i, txt in enumerate(np.round(cv_log_error_array, 3)):
    ax.annotate((alpha[i], str(txt)), (alpha[i], cv_log_error_array[i]))
```

```
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

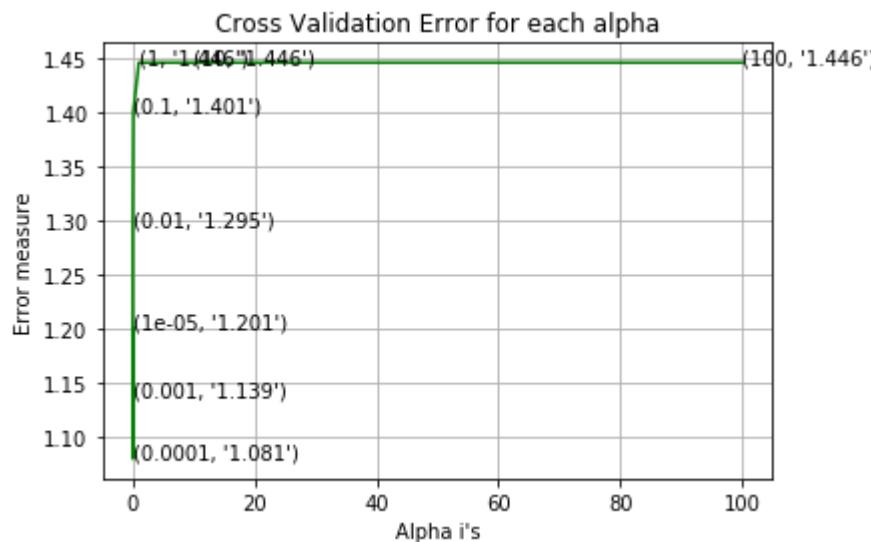
best_alpha = np.argmin(cv_log_error_array)
# clf = SVC(C=i,kernel='linear',probability=True, class_weight='balanced')
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='hinge', random_state=42)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for C = 1e-05
Log Loss : 1.2012276765711212
for C = 0.0001
Log Loss : 1.0805984169146248
for C = 0.001
Log Loss : 1.1387963409712925
for C = 0.01
Log Loss : 1.2950987178358815
for C = 0.1
Log Loss : 1.40116700878263
for C = 1
Log Loss : 1.446487790877563
for C = 10
Log Loss : 1.4464884509630414
for C = 100
Log Loss : 1.4464884906240076

```



For values of best alpha = 0.0001 The train log loss is: 0.4092955544188989
 For values of best alpha = 0.0001 The cross validation log loss is: 1.0805984169146248
 For values of best alpha = 0.0001 The test log loss is: 0.9934440770177451

4.4.2. Testing model with best hyper parameters

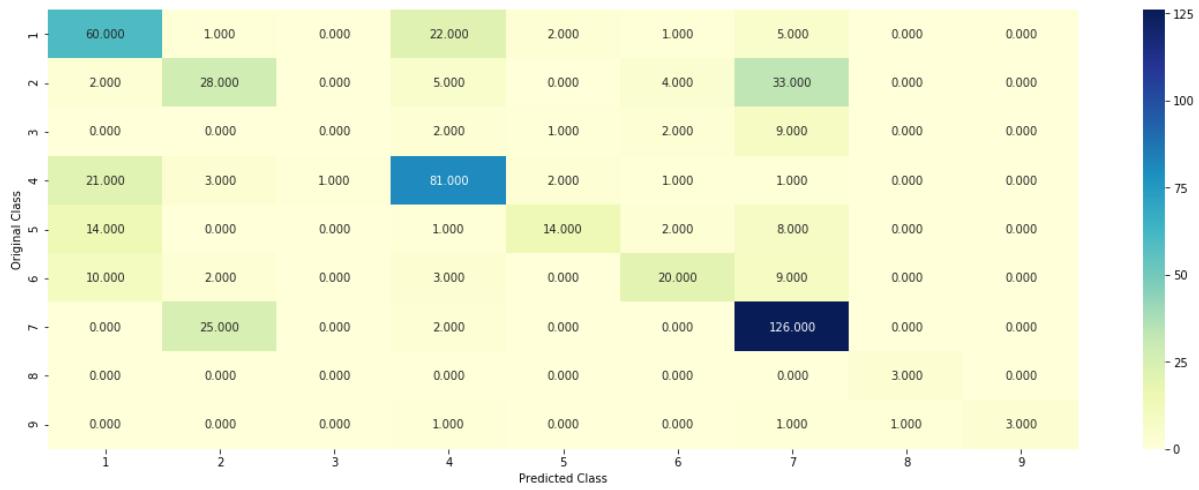
```
In [39]: # read more about support vector machines with Linear kernels here http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

# -----
# default parameters
# SVC(C=1.0, kernel='rbf', degree=3, gamma='auto', coef0=0.0, shrinking=True,
# probability=False, tol=0.001,
# cache_size=200, class_weight=None, verbose=False, max_iter=-1, decision_function_shape='ovr', random_state=None)

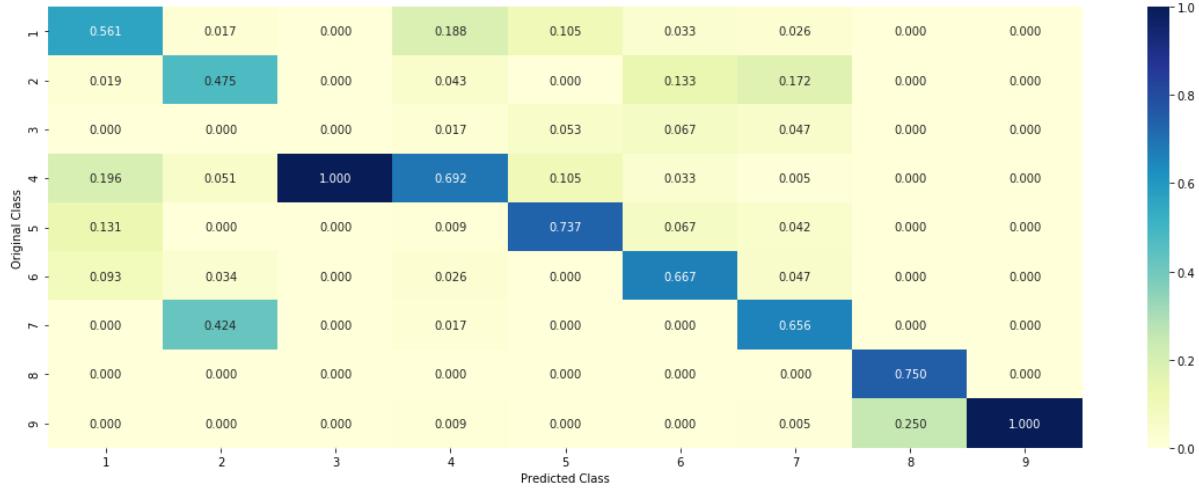
# Some of methods of SVM()
# fit(X, y, [sample_weight])      Fit the SVM model according to the given training data.
# predict(X)        Perform classification on samples in X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/mathematical-derivation-copy-8/
# -----


# clf = SVC(C=alpha[best_alpha],kernel='linear',probability=True, class_weight='balanced')
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='hinge', random_state=42, class_weight='balanced')
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCoding, cv_y, clf)
```

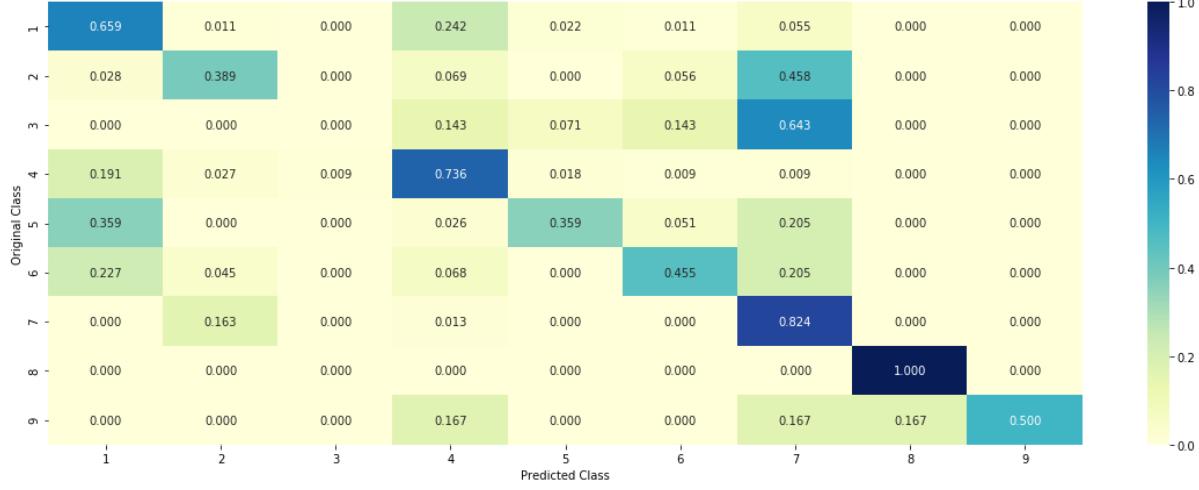
Log loss : 1.0805984169146248
 Number of mis-classified points : 0.37030075187969924
 ----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.3.3. Feature Importance

4.3.3.1. For Correctly classified point

```
In [40]: clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='hinge', random_state=42)
clf.fit(train_x_onehotCoding,train_y)
# test_point_index = 1
test_point_index = 100
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_impfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)
```

Predicted Class : 7
Predicted Class Probabilities: [[0.0301 0.3197 0.0212 0.0608 0.0744 0.0272 0.
4596 0.0034 0.0037]]
Actual Class : 7

1 Text feature [3t3] present in test data point [True]
13 Text feature [1st] present in test data point [True]
17 Text feature [6b] present in test data point [True]
24 Text feature [122] present in test data point [True]
58 Text feature [65] present in test data point [True]
59 Text feature [140] present in test data point [True]
64 Text feature [33] present in test data point [True]
66 Text feature [3f] present in test data point [True]
81 Text feature [26] present in test data point [True]
83 Text feature [295] present in test data point [True]
93 Text feature [598] present in test data point [True]
114 Text feature [adjuvant] present in test data point [True]
122 Text feature [17] present in test data point [True]
133 Text feature [107] present in test data point [True]
154 Text feature [336] present in test data point [True]
168 Text feature [116] present in test data point [True]
171 Text feature [378] present in test data point [True]
174 Text feature [1040] present in test data point [True]
185 Text feature [2a] present in test data point [True]
188 Text feature [0013] present in test data point [True]
194 Text feature [400] present in test data point [True]
215 Text feature [657] present in test data point [True]
233 Text feature [100] present in test data point [True]
240 Text feature [596kb] present in test data point [True]
265 Text feature [64] present in test data point [True]
269 Text feature [530] present in test data point [True]
286 Text feature [3rd] present in test data point [True]
296 Text feature [599] present in test data point [True]
297 Text feature [14] present in test data point [True]
307 Text feature [25] present in test data point [True]
365 Text feature [212] present in test data point [True]
375 Text feature [2001] present in test data point [True]
394 Text feature [29] present in test data point [True]
418 Text feature [0001] present in test data point [True]
423 Text feature [53] present in test data point [True]
442 Text feature [advance] present in test data point [True]
443 Text feature [93] present in test data point [True]
453 Text feature [activated] present in test data point [True]
455 Text feature [97] present in test data point [True]
462 Gene feature [BRAF] present in test data point [True]
473 Text feature [173] present in test data point [True]
475 Text feature [675] present in test data point [True]
479 Text feature [1993] present in test data point [True]
499 Text feature [200] present in test data point [True]

Out of the top 500 features 44 are present in query point

4.3.3.2. For Incorrectly classified point

```
In [43]: test_point_index = 550
no_feature = 500
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_impfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

Predicted Class : 5
Predicted Class Probabilities: [[0.2047 0.0677 0.0219 0.1329 0.4257 0.0914 0.
0419 0.0067 0.0071]]
Actual Class : 4

14 Text feature [2013] present in test data point [True]
27 Text feature [36] present in test data point [True]
37 Text feature [186] present in test data point [True]
38 Text feature [54] present in test data point [True]
43 Text feature [350] present in test data point [True]
47 Text feature [47] present in test data point [True]
67 Text feature [239] present in test data point [True]
71 Text feature [43] present in test data point [True]
78 Text feature [57] present in test data point [True]
90 Text feature [128] present in test data point [True]
99 Text feature [310] present in test data point [True]
119 Text feature [62] present in test data point [True]
142 Text feature [183] present in test data point [True]
159 Text feature [276] present in test data point [True]
162 Text feature [353] present in test data point [True]
169 Text feature [61] present in test data point [True]
170 Text feature [51] present in test data point [True]
186 Text feature [360] present in test data point [True]
205 Text feature [000] present in test data point [True]
209 Text feature [214] present in test data point [True]
221 Text feature [afforded] present in test data point [True]
239 Text feature [abbreviations] present in test data point [True]
245 Text feature [1b] present in test data point [True]
260 Text feature [326] present in test data point [True]
271 Text feature [685] present in test data point [True]
274 Text feature [401] present in test data point [True]
277 Text feature [169] present in test data point [True]
284 Text feature [198] present in test data point [True]
288 Text feature [250] present in test data point [True]
293 Gene feature [MLH1] present in test data point [True]
295 Text feature [33] present in test data point [True]
306 Text feature [22] present in test data point [True]
310 Text feature [116] present in test data point [True]
328 Text feature [270] present in test data point [True]
343 Text feature [300] present in test data point [True]
356 Text feature [accurate] present in test data point [True]
371 Text feature [935] present in test data point [True]
372 Text feature [accompanying] present in test data point [True]
382 Text feature [aag] present in test data point [True]
396 Text feature [act] present in test data point [True]
406 Text feature [account] present in test data point [True]
411 Text feature [accessible] present in test data point [True]
424 Text feature [97] present in test data point [True]
464 Text feature [05] present in test data point [True]
470 Text feature [5a] present in test data point [True]
471 Text feature [292] present in test data point [True]
473 Text feature [34] present in test data point [True]
Out of the top 500 features 47 are present in query point

4.5 Random Forest Classifier

4.5.1. Hyper parameter tuning (With One hot Encoding)

In [44]:

```

# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', max_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given training data.
# predict(X) Perform classification on samples in X.
# predict_proba (X) Perform classification on samples in X.

# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/random-forest-and-their-construction-2/
# -----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [100,200,500,1000,2000]
max_depth = [5, 10]
cv_log_error_array = []
for i in alpha:
    for j in max_depth:
        print("for n_estimators =", i,"and max depth = ", j)
        clf = RandomForestClassifier(n_estimators=i, criterion='gini', max_depth=j, random_state=42, n_jobs=-1)
        clf.fit(train_x_onehotCoding, train_y)
        sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
        sig_clf.fit(train_x_onehotCoding, train_y)
        sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
        cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))

```

```

        print("Log Loss :",log_loss(cv_y, sig_clf_probs))

'''fig, ax = plt.subplots()
features = np.dot(np.array(alpha)[:,None],np.array(max_depth)[None]).ravel()
ax.plot(features, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
    ax.annotate((alpha[int(i/2)],max_depth[int(i%2)],str(txt)), (features[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()
'''

best_alpha = np.argmin(cv_log_error_array)
clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/2)], criterion='gini', max_depth=max_depth[int(best_alpha%2)], random_state=42, n_jobs=-1)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best estimator = ', alpha[int(best_alpha/2)], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best estimator = ', alpha[int(best_alpha/2)], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best estimator = ', alpha[int(best_alpha/2)], "The test 1 log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))

for n_estimators = 100 and max depth = 5
Log Loss : 1.183955668547268
for n_estimators = 100 and max depth = 10
Log Loss : 1.0956785400729443
for n_estimators = 200 and max depth = 5
Log Loss : 1.1779254916800874
for n_estimators = 200 and max depth = 10
Log Loss : 1.0891468268869304
for n_estimators = 500 and max depth = 5
Log Loss : 1.17519335748105
for n_estimators = 500 and max depth = 10
Log Loss : 1.0865307201703858
for n_estimators = 1000 and max depth = 5
Log Loss : 1.174150871181727
for n_estimators = 1000 and max depth = 10
Log Loss : 1.0852399201037555
for n_estimators = 2000 and max depth = 5
Log Loss : 1.1732043003547163
for n_estimators = 2000 and max depth = 10
Log Loss : 1.0854244958625219
For values of best estimator = 1000 The train log loss is: 0.689501577266708
4
For values of best estimator = 1000 The cross validation log loss is: 1.0852399201037555
For values of best estimator = 1000 The test log loss is: 1.0109120692997888

```

4.5.2. Testing model with best hyper parameters (One Hot Encoding)

In [45]:

```
# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', m
ax_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_l
eaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_s
tate=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given train
ing data.
# predict(X) Perform classification on samples in X.
# predict_proba (X) Perform classification on samples in X.

# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

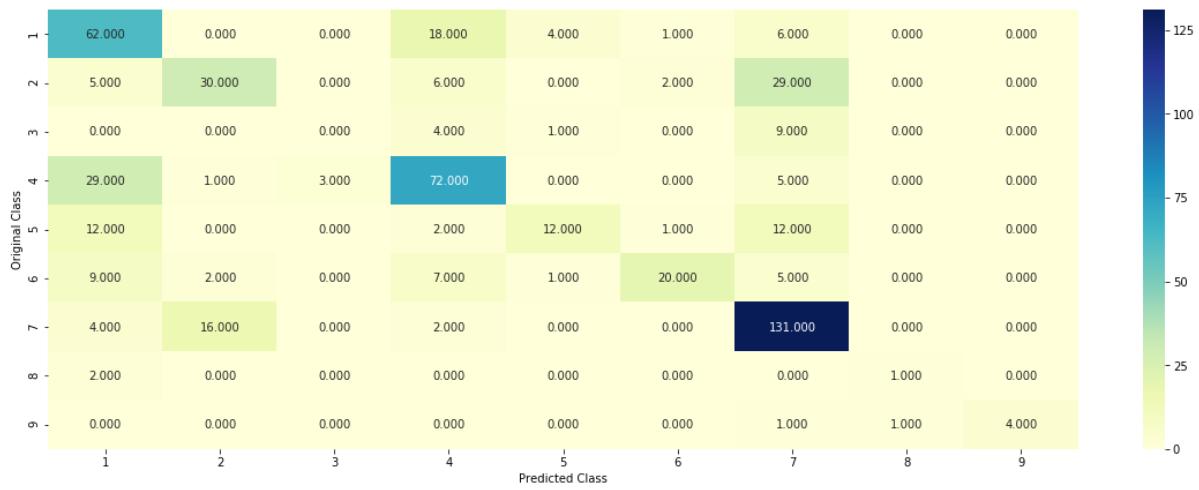
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/
Lessons/random-forest-and-their-construction-2/
# -----
```

```
clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/2)], criterion=
'gini', max_depth=max_depth[int(best_alpha%2)], random_state=42, n_jobs=-1)
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCod
ing, cv_y, clf)
```

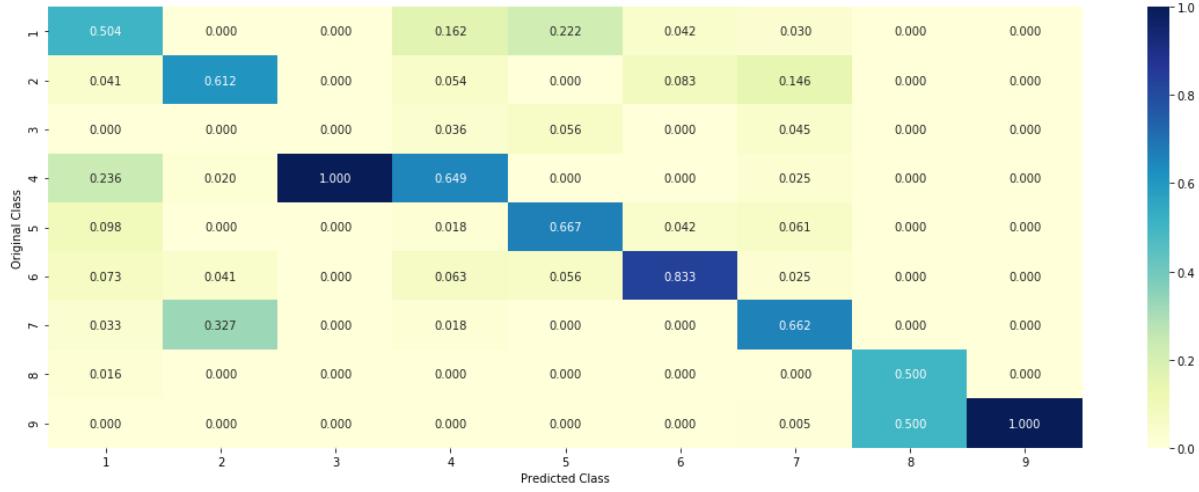
Log loss : 1.0852399201037555

Number of mis-classified points : 0.37593984962406013

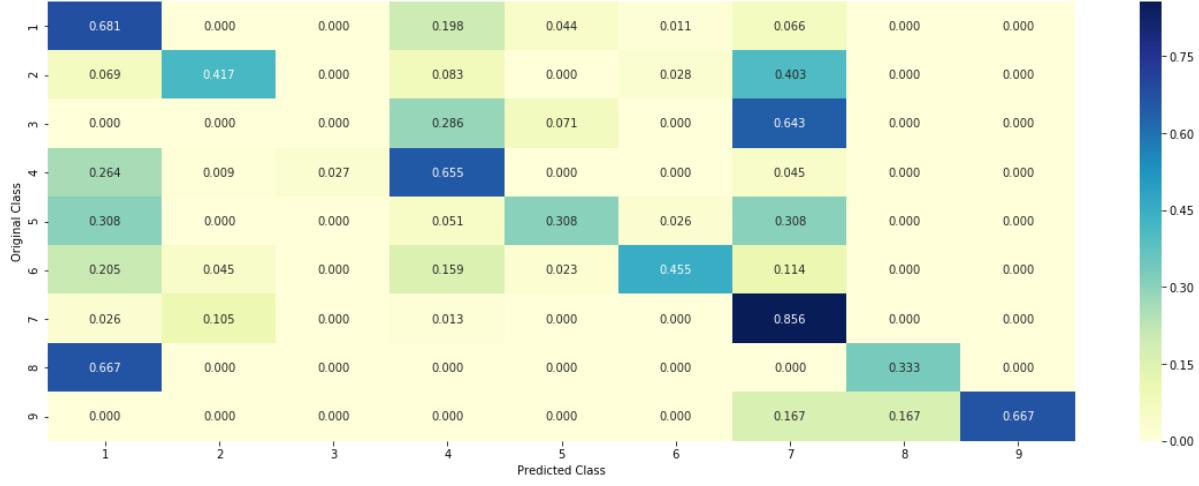
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.5.3. Feature Importance

4.5.3.1. Correctly Classified point

```
In [47]: # test_point_index = 10
clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/2)], criterion=
'gini', max_depth=max_depth[int(best_alpha%2)], random_state=42, n_jobs=-1)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

test_point_index = 56
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.feature_importances_)
print("-"*50)
get_imptfeature_names(indices[:no_feature], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_poin
t_index], no_feature)

Predicted Class : 9
Predicted Class Probabilities: [[0.2033 0.043  0.0158 0.2096 0.0462 0.0414 0.
0528 0.0189 0.369 ]]
Actual Class : 9
-----
4 Text feature [activated] present in test data point [True]
39 Text feature [acquisition] present in test data point [True]
75 Text feature [046] present in test data point [True]
83 Text feature [agency] present in test data point [True]
84 Text feature [ablate] present in test data point [True]
Out of the top 100 features 5 are present in query point
```

4.5.3.2. Inorrectly Classified point

```
In [49]: test_point_index = 550
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actuall Class :", test_y[test_point_index])
indices = np.argsort(-clf.feature_importances_)
print("-"*50)
get_imptfeature_names(indices[:no_feature], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)
```

```
Predicted Class : 1
Predicted Class Probabilities: [[0.3587 0.0261 0.021  0.278  0.1788 0.0991 0.
0271 0.0045 0.0067]]
Actuall Class : 4
-----
11 Text feature [age] present in test data point [True]
35 Text feature [actions] present in test data point [True]
39 Text feature [acquisition] present in test data point [True]
67 Text feature [accompanying] present in test data point [True]
69 Text feature [aberrant] present in test data point [True]
78 Text feature [97] present in test data point [True]
81 Text feature [account] present in test data point [True]
83 Text feature [agency] present in test data point [True]
95 Text feature [98] present in test data point [True]
Out of the top 100 features 9 are present in query point
```

4.5.3. Hyper parameter tuning (With Response Coding)

In [50]:

```

# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', max_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given training data.
# predict(X) Perform classification on samples in X.
# predict_proba (X) Perform classification on samples in X.

# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/random-forest-and-their-construction-2/
# -----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight]) Fit the calibrated model
# get_params([deep]) Get parameters for this estimator.
# predict(X) Predict the target of new samples.
# predict_proba(X) Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10,50,100,200,500,1000]
max_depth = [2,3,5,10]
cv_log_error_array = []
for i in alpha:
    for j in max_depth:
        print("for n_estimators =", i, "and max depth = ", j)
        clf = RandomForestClassifier(n_estimators=i, criterion='gini', max_depth=j, random_state=42, n_jobs=-1)
        clf.fit(train_x_responseCoding, train_y)
        sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
        sig_clf.fit(train_x_responseCoding, train_y)
        sig_clf_probs = sig_clf.predict_proba(cv_x_responseCoding)
        cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))

```

```
        print("Log Loss :",log_loss(cv_y, sig_clf_probs))
    ...
    fig, ax = plt.subplots()
    features = np.dot(np.array(alpha)[:,None],np.array(max_depth)[None]).ravel()
    ax.plot(features, cv_log_error_array,c='g')
    for i, txt in enumerate(np.round(cv_log_error_array,3)):
        ax.annotate((alpha[int(i/4)],max_depth[int(i%4)],str(txt)), (features[i],cv_log_error_array[i]))
    plt.grid()
    plt.title("Cross Validation Error for each alpha")
    plt.xlabel("Alpha i's")
    plt.ylabel("Error measure")
    plt.show()
    '''

best_alpha = np.argmin(cv_log_error_array)
clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/4)], criterion='gini', max_depth=max_depth[int(best_alpha%4)], random_state=42, n_jobs=-1)
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_responseCoding)
print('For values of best alpha = ', alpha[int(best_alpha/4)], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_responseCoding)
print('For values of best alpha = ', alpha[int(best_alpha/4)], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_responseCoding)
print('For values of best alpha = ', alpha[int(best_alpha/4)], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for n_estimators = 10 and max depth = 2
Log Loss : 1.4319772741372518
for n_estimators = 10 and max depth = 3
Log Loss : 1.305960912886437
for n_estimators = 10 and max depth = 5
Log Loss : 1.2175544791726336
for n_estimators = 10 and max depth = 10
Log Loss : 1.0372689526448795
for n_estimators = 50 and max depth = 2
Log Loss : 1.4210311426546973
for n_estimators = 50 and max depth = 3
Log Loss : 1.2595976046338504
for n_estimators = 50 and max depth = 5
Log Loss : 1.1249319691975996
for n_estimators = 50 and max depth = 10
Log Loss : 1.0017724528298393
for n_estimators = 100 and max depth = 2
Log Loss : 1.416472900105551
for n_estimators = 100 and max depth = 3
Log Loss : 1.2842131472849436
for n_estimators = 100 and max depth = 5
Log Loss : 1.1201450100715424
for n_estimators = 100 and max depth = 10
Log Loss : 0.9978374676858422
for n_estimators = 200 and max depth = 2
Log Loss : 1.386870268670287
for n_estimators = 200 and max depth = 3
Log Loss : 1.257988790445603
for n_estimators = 200 and max depth = 5
Log Loss : 1.104477756933601
for n_estimators = 200 and max depth = 10
Log Loss : 1.0036204997004872
for n_estimators = 500 and max depth = 2
Log Loss : 1.3212417846800704
for n_estimators = 500 and max depth = 3
Log Loss : 1.2196362275219985
for n_estimators = 500 and max depth = 5
Log Loss : 1.1102245390629306
for n_estimators = 500 and max depth = 10
Log Loss : 0.9995917087952124
for n_estimators = 1000 and max depth = 2
Log Loss : 1.3102714680360001
for n_estimators = 1000 and max depth = 3
Log Loss : 1.2155196291402768
for n_estimators = 1000 and max depth = 5
Log Loss : 1.1188168592170342
for n_estimators = 1000 and max depth = 10
Log Loss : 1.0014672313442274
For values of best alpha = 100 The train log loss is: 0.10819655263337148
For values of best alpha = 100 The cross validation log loss is: 0.997837467
6858422
For values of best alpha = 100 The test log loss is: 0.8924571744597392

```

4.5.4. Testing model with best hyper parameters (Response Coding)

In [51]:

```
# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', max_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight]) Fit the SVM model according to the given training data.
# predict(X) Perform classification on samples in X.
# predict_proba (X) Perform classification on samples in X.

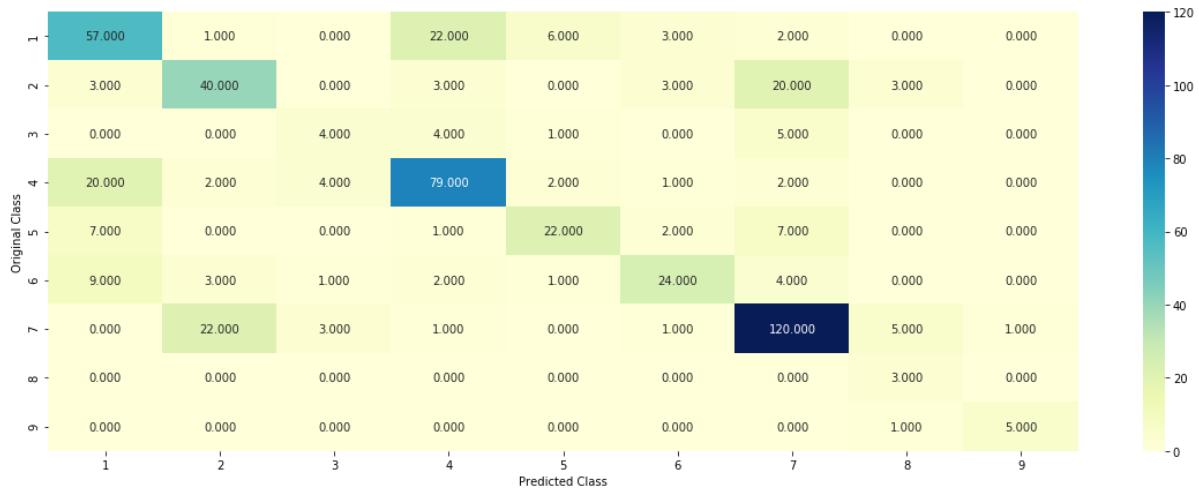
# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/random-forest-and-their-construction-2/
# -----

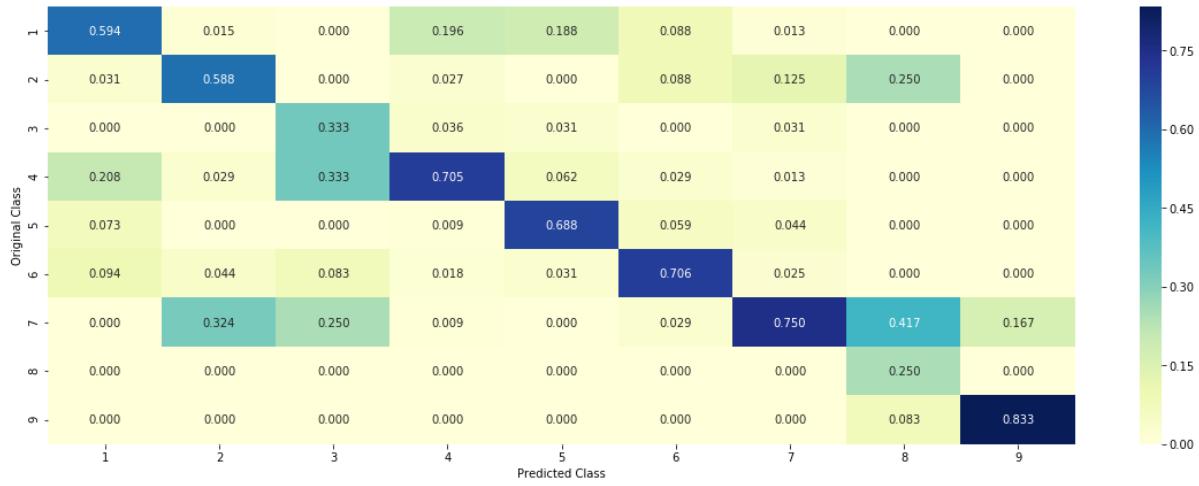


clf = RandomForestClassifier(max_depth=max_depth[int(best_alpha*4)], n_estimators=alpha[int(best_alpha/4)], criterion='gini', max_features='auto', random_state=42)
predict_and_plot_confusion_matrix(train_x_responseCoding, train_y, cv_x_responseCoding, cv_y, clf)
```

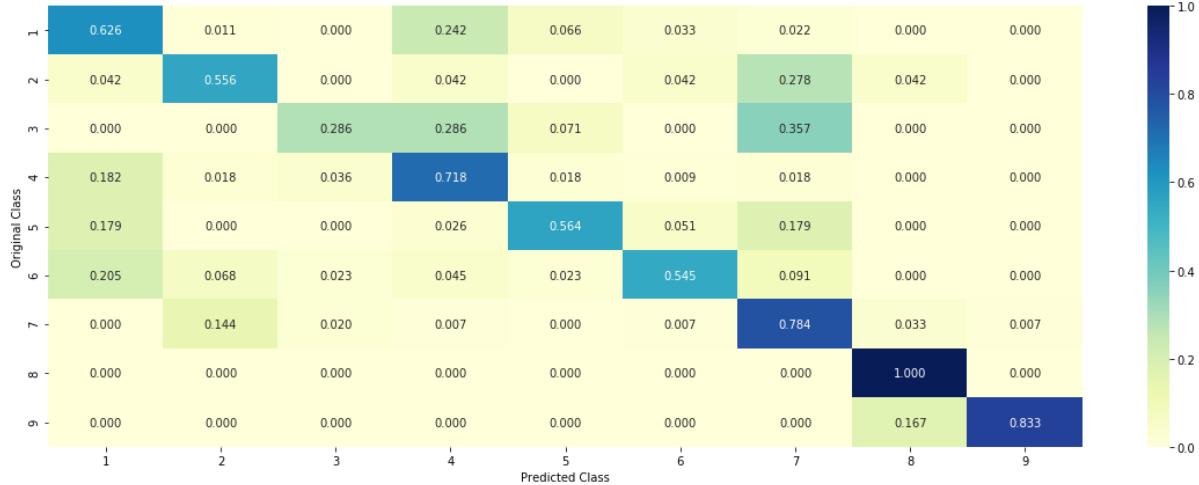
Log loss : 0.9978374676858422
 Number of mis-classified points : 0.33458646616541354
 ----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.5.5. Feature Importance

4.5.5.1. Correctly Classified point

```
In [53]: clf = RandomForestClassifier(n_estimators=alpha[int(best_alpha/4)], criterion='gini', max_depth=max_depth[int(best_alpha%4)], random_state=42, n_jobs=-1)
clf.fit(train_x_responseCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_responseCoding, train_y)

test_point_index = 50
no_feature = 27
predicted_cls = sig_clf.predict(test_x_responseCoding[test_point_index].reshape(1,-1))
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_responseCoding[test_point_index].reshape(1,-1)),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.feature_importances_)
print("-"*50)
for i in indices:
    if i<9:
        print("Gene is important feature")
    elif i<18:
        print("Variation is important feature")
    else:
        print("Text is important feature")
```


4.5.5.2. Incorrectly Classified point

```
In [59]: test_point_index = 550
predicted_cls = sig_clf.predict(test_x_responseCoding[test_point_index].reshape(1,-1))
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_responseCoding[test_point_index].reshape(1,-1)),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-clf.feature_importances_)
print("-"*50)
for i in indices:
    if i<9:
        print("Gene is important feature")
    elif i<18:
        print("Variation is important feature")
    else:
        print("Text is important feature")
```


4.7 Stack the models

4.7.1 testing with hyper parameter tuning

```
In [60]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/geometric-intuition-1/
#-----


# read more about support vector machines with Linear kernels here http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html
# -----
# default parameters
# SVC(C=1.0, kernel='rbf', degree=3, gamma='auto', coef0=0.0, shrinking=True,
# probability=False, tol=0.001,
# cache_size=200, class_weight=None, verbose=False, max_iter=-1, decision_function_shape='ovr', random_state=None)

# Some of methods of SVM()
# fit(X, y, [sample_weight])   Fit the SVM model according to the given training data.
# predict(X)     Perform classification on samples in X.
# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/Lessons/mathematical-derivation-copy-8/
# -----


# read more about support vector machines with Linear kernels here http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html
# -----
# default parameters
# sklearn.ensemble.RandomForestClassifier(n_estimators=10, criterion='gini', max_depth=None, min_samples_split=2,
# min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', max_leaf_nodes=None, min_impurity_decrease=0.0,
# min_impurity_split=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False,
# class_weight=None)

# Some of methods of RandomForestClassifier()
# fit(X, y, [sample_weight])   Fit the SVM model according to the given training data.
```

```

# predict(X)      Perform classification on samples in X.
# predict_proba (X)    Perform classification on samples in X.

# some of attributes of RandomForestClassifier()
# feature_importances_ : array of shape = [n_features]
# The feature importances (the higher, the more important the feature).

# -----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/
Lessons/random-forest-and-their-construction-2/
# -----


clf1 = SGDClassifier(alpha=0.001, penalty='l2', loss='log', class_weight='balanced', random_state=0)
clf1.fit(train_x_onehotCoding, train_y)
sig_clf1 = CalibratedClassifierCV(clf1, method="sigmoid")

clf2 = SGDClassifier(alpha=1, penalty='l2', loss='hinge', class_weight='balanced', random_state=0)
clf2.fit(train_x_onehotCoding, train_y)
sig_clf2 = CalibratedClassifierCV(clf2, method="sigmoid")

clf3 = MultinomialNB(alpha=0.001)
clf3.fit(train_x_onehotCoding, train_y)
sig_clf3 = CalibratedClassifierCV(clf3, method="sigmoid")

sig_clf1.fit(train_x_onehotCoding, train_y)
print("Logistic Regression : Log Loss: %0.2f" % (log_loss(cv_y, sig_clf1.predict_proba(cv_x_onehotCoding))))
sig_clf2.fit(train_x_onehotCoding, train_y)
print("Support vector machines : Log Loss: %0.2f" % (log_loss(cv_y, sig_clf2.predict_proba(cv_x_onehotCoding))))
sig_clf3.fit(train_x_onehotCoding, train_y)
print("Naive Bayes : Log Loss: %0.2f" % (log_loss(cv_y, sig_clf3.predict_proba(cv_x_onehotCoding))))
print("-"*50)
alpha = [0.0001, 0.001, 0.01, 0.1, 1, 10]
best_alpha = 999
for i in alpha:
    lr = LogisticRegression(C=i)
    sclf = StackingClassifier(classifiers=[sig_clf1, sig_clf2, sig_clf3], meta_classifier=lr, use_probas=True)
    sclf.fit(train_x_onehotCoding, train_y)
    print("Stacking Classifier : for the value of alpha: %f Log Loss: %0.3f" %
(i, log_loss(cv_y, sclf.predict_proba(cv_x_onehotCoding))))
    log_error = log_loss(cv_y, sclf.predict_proba(cv_x_onehotCoding))
    if best_alpha > log_error:
        best_alpha = log_error

```

Logistic Regression : Log Loss: 1.06
Support vector machines : Log Loss: 1.45
Naive Bayes : Log Loss: 1.21

Stacking Classifier : for the value of alpha: 0.000100 Log Loss: 2.178
Stacking Classifier : for the value of alpha: 0.001000 Log Loss: 2.040
Stacking Classifier : for the value of alpha: 0.010000 Log Loss: 1.547
Stacking Classifier : for the value of alpha: 0.100000 Log Loss: 1.204
Stacking Classifier : for the value of alpha: 1.000000 Log Loss: 1.344
Stacking Classifier : for the value of alpha: 10.000000 Log Loss: 1.713

4.7.2 testing the model with the best hyper parameters

```
In [61]: lr = LogisticRegression(C=0.1)
sclf = StackingClassifier(classifiers=[sig_clf1, sig_clf2, sig_clf3], meta_classifier=lr, use_probas=True)
sclf.fit(train_x_onehotCoding, train_y)

log_error = log_loss(train_y, sclf.predict_proba(train_x_onehotCoding))
print("Log loss (train) on the stacking classifier :",log_error)

log_error = log_loss(cv_y, sclf.predict_proba(cv_x_onehotCoding))
print("Log loss (CV) on the stacking classifier :",log_error)

log_error = log_loss(test_y, sclf.predict_proba(test_x_onehotCoding))
print("Log loss (test) on the stacking classifier :",log_error)

print("Number of missclassified point :", np.count_nonzero((sclf.predict(test_x_onehotCoding)- test_y))/test_y.shape[0])
plot_confusion_matrix(test_y=test_y, predict_y=sclf.predict(test_x_onehotCoding))
```

Log loss (train) on the stacking classifier : 0.5983067620836117

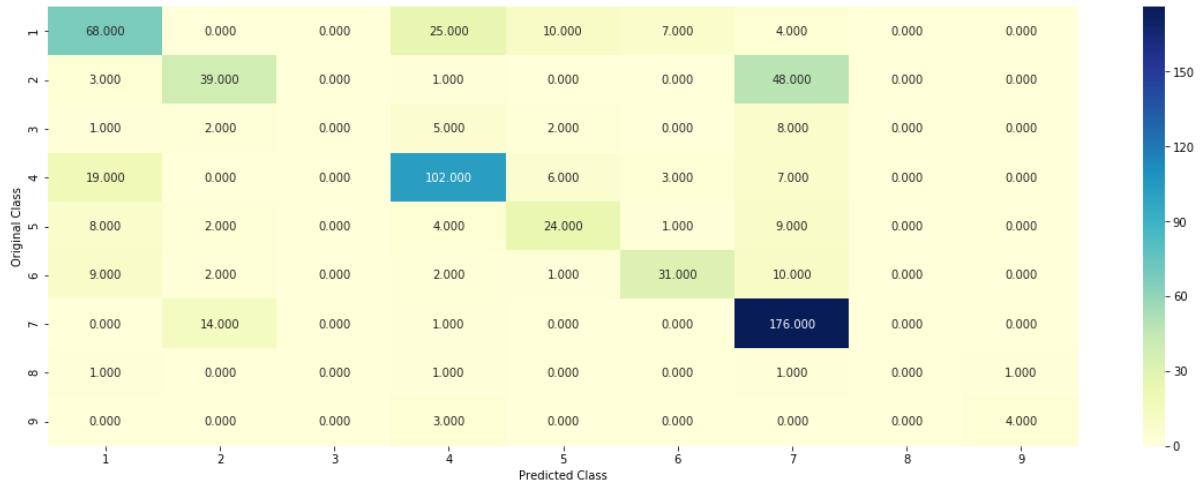
Log loss (CV) on the stacking classifier : 1.2044848236198258

Log loss (test) on the stacking classifier : 1.0948953363351406

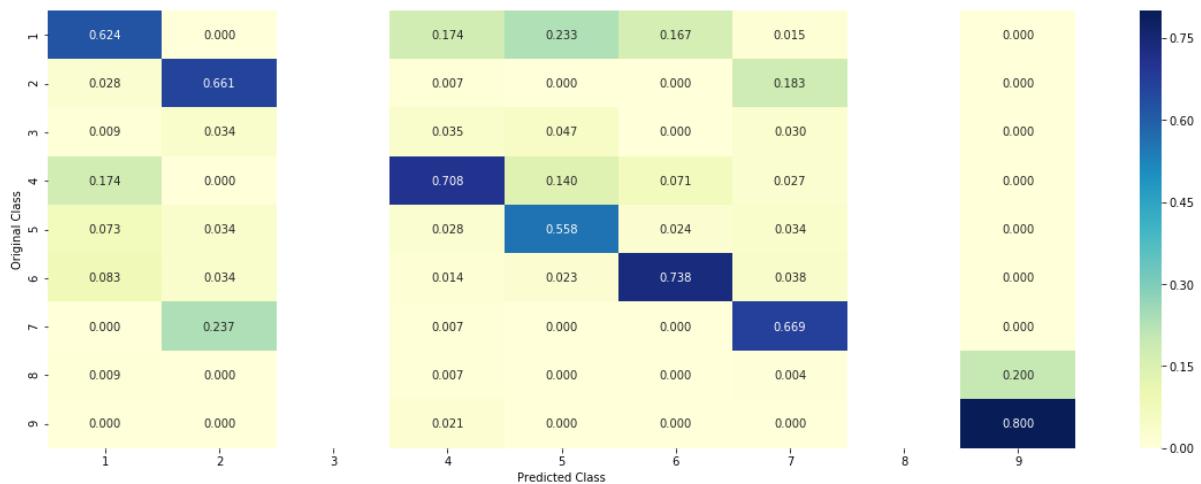
Number of missclassified point : 0.3323308270676692

----- Confusion matrix -----

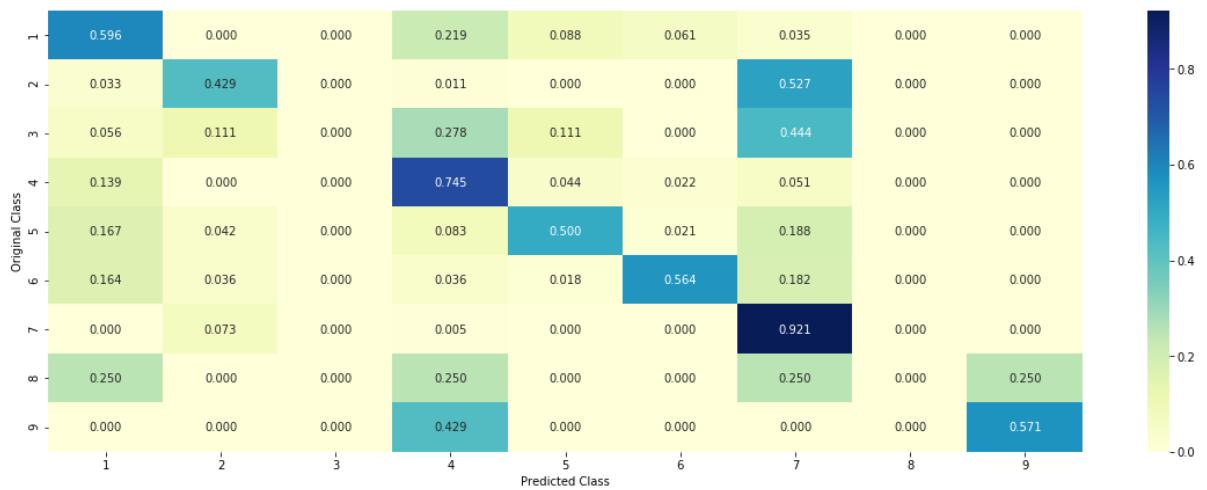
C:\Users\LENOVO\Anaconda3\lib\site-packages\ipykernel_launcher.py:22: RuntimeWarning: invalid value encountered in true_divide



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.7.3 Maximum Voting classifier

```
In [62]: #Refer:http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.VotingClassifier.html
from sklearn.ensemble import VotingClassifier
vclf = VotingClassifier(estimators=[('lr', sig_clf1), ('svc', sig_clf2), ('rf', sig_clf3)], voting='soft')
vclf.fit(train_x_onehotCoding, train_y)
print("Log loss (train) on the VotingClassifier :", log_loss(train_y, vclf.predict_proba(train_x_onehotCoding)))
print("Log loss (CV) on the VotingClassifier :", log_loss(cv_y, vclf.predict_proba(cv_x_onehotCoding)))
print("Log loss (test) on the VotingClassifier :", log_loss(test_y, vclf.predict_proba(test_x_onehotCoding)))
print("Number of missclassified point :", np.count_nonzero((vclf.predict(test_x_onehotCoding)- test_y)/test_y.shape[0]))
plot_confusion_matrix(test_y=test_y, predict_y=vclf.predict(test_x_onehotCoding))
```

Log loss (train) on the VotingClassifier : 0.8162958134944885

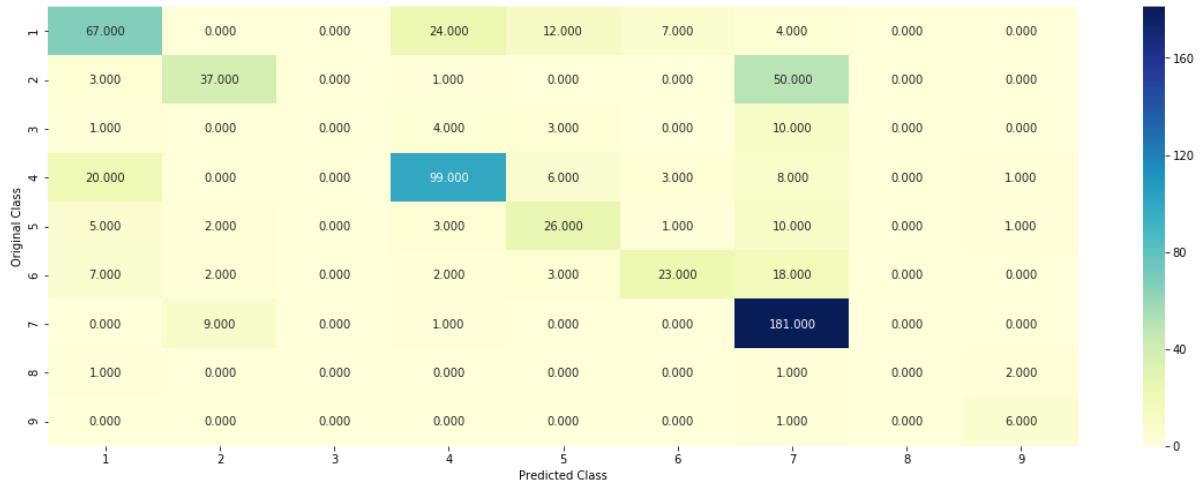
Log loss (CV) on the VotingClassifier : 1.1749084556760052

Log loss (test) on the VotingClassifier : 1.1234565574021964

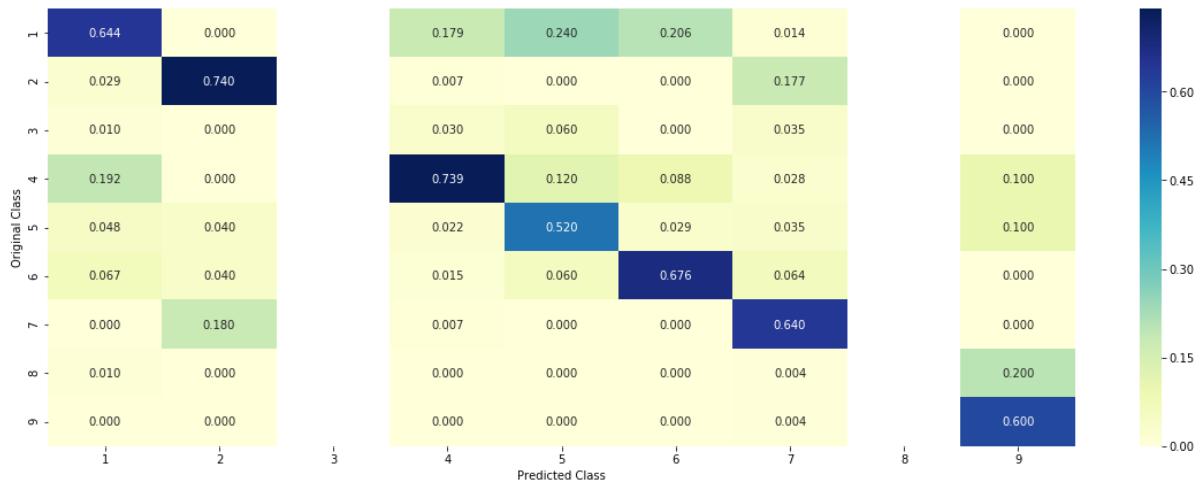
Number of missclassified point : 0.3398496240601504

----- Confusion matrix -----

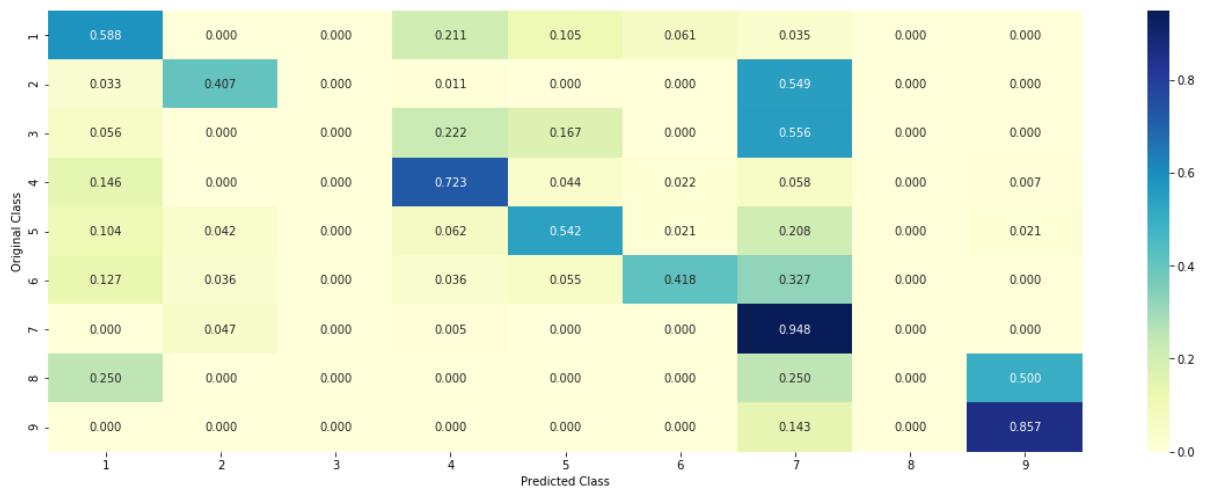
C:\Users\LENOVO\Anaconda3\lib\site-packages\ipykernel_launcher.py:22: RuntimeWarning: invalid value encountered in true_divide



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



3.LR:

In []:

```
In [9]: # https://machinelearningmastery.com/how-to-fix-futurewarning-messages-in-scikit-learn/
from warnings import simplefilter
# ignore all future warnings
simplefilter(action='ignore', category=FutureWarning)

import pandas as pd
import matplotlib.pyplot as plt
import re
import time
import warnings
import numpy as np
from nltk.corpus import stopwords
from sklearn.decomposition import TruncatedSVD
from sklearn.preprocessing import normalize
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.manifold import TSNE
import seaborn as sns
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix
from sklearn.metrics.classification import accuracy_score, log_loss
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import SGDClassifier
from imblearn.over_sampling import SMOTE
from collections import Counter
from scipy.sparse import hstack
from sklearn.multiclass import OneVsRestClassifier
from sklearn.svm import SVC
from sklearn.model_selection import StratifiedKFold
from collections import Counter, defaultdict
from sklearn.calibration import CalibratedClassifierCV
from sklearn.naive_bayes import MultinomialNB
from sklearn.naive_bayes import GaussianNB
from sklearn.model_selection import train_test_split
from sklearn.model_selection import GridSearchCV
import math
from sklearn.metrics import normalized_mutual_info_score
from sklearn.ensemble import RandomForestClassifier
warnings.filterwarnings("ignore")

from mlxtend.classifier import StackingClassifier

from sklearn import model_selection
from sklearn.linear_model import LogisticRegression
```

3.1. Reading Data

3.1.1. Reading Gene and Variation Data

```
In [2]: data = pd.read_csv('training/training_variants')
print('Number of data points : ', data.shape[0])
print('Number of features : ', data.shape[1])
print('Features : ', data.columns.values)
data.head()
```

Number of data points : 3321
 Number of features : 4
 Features : ['ID' 'Gene' 'Variation' 'Class']

Out[2]:

	ID	Gene	Variation	Class
0	0	FAM58A	Truncating Mutations	1
1	1	CBL	W802*	2
2	2	CBL	Q249E	2
3	3	CBL	N454D	3
4	4	CBL	L399V	4

training/training_variants is a comma separated file containing the description of the genetic mutations used for training.

Fields are

- **ID** : the id of the row used to link the mutation to the clinical evidence
- **Gene** : the gene where this genetic mutation is located
- **Variation** : the aminoacid change for this mutations
- **Class** : 1-9 the class this genetic mutation has been classified on

3.1.2. Reading Text Data

```
In [3]: # note the separator in this file
data_text = pd.read_csv("training/training_text", sep="\|\|", engine="python", names=["ID", "TEXT"], skiprows=1)
print('Number of data points : ', data_text.shape[0])
print('Number of features : ', data_text.shape[1])
print('Features : ', data_text.columns.values)
data_text.head()
```

Number of data points : 3321
 Number of features : 2
 Features : ['ID' 'TEXT']

Out[3]:

	ID	TEXT
0	0	Cyclin-dependent kinases (CDKs) regulate a var...
1	1	Abstract Background Non-small cell lung canc...
2	2	Abstract Background Non-small cell lung canc...
3	3	Recent evidence has demonstrated that acquired...
4	4	Oncogenic mutations in the monomeric Casitas B...

3.1.3. Preprocessing of text

```
In [4]: # Loading stop words from nltk library
stop_words = set(stopwords.words('english'))

def nlp_preprocessing(total_text, index, column):
    if type(total_text) is not int:
        string = ""
        # replace every special char with space
        total_text = re.sub('[^a-zA-Z0-9\n]', ' ', total_text)
        # replace multiple spaces with single space
        total_text = re.sub('\s+', ' ', total_text)
        # converting all the chars into lower-case.
        total_text = total_text.lower()

        for word in total_text.split():
            # if the word is a not a stop word then retain that word from the data
            if not word in stop_words:
                string += word + " "

        data_text[column][index] = string
```

```
In [5]: #text processing stage.
start_time = time.clock()
for index, row in data_text.iterrows():
    if type(row['TEXT']) is str:
        nlp_preprocessing(row['TEXT'], index, 'TEXT')
    else:
        print("there is no text description for id:",index)
print('Time took for preprocessing the text :',time.clock() - start_time, "seconds")
```

there is no text description for id: 1109
 there is no text description for id: 1277
 there is no text description for id: 1407
 there is no text description for id: 1639
 there is no text description for id: 2755
 Time took for preprocessing the text : 152.977922774 seconds

```
In [6]: #merging both gene_variations and text data based on ID
result = pd.merge(data, data_text, on='ID', how='left')
result.head()
```

Out[6]:

	ID	Gene	Variation	Class	TEXT
0	0	FAM58A	Truncating Mutations	1	cyclin dependent kinases cdks regulate variety...
1	1	CBL	W802*	2	abstract background non small cell lung cancer...
2	2	CBL	Q249E	2	abstract background non small cell lung cancer...
3	3	CBL	N454D	3	recent evidence demonstrated acquired uniparen...
4	4	CBL	L399V	4	oncogenic mutations monomeric casitas b lineage...

```
In [7]: result[result.isnull().any(axis=1)]
```

Out[7]:

	ID	Gene	Variation	Class	TEXT
1109	1109	FANCA	S1088F	1	NaN
1277	1277	ARID5B	Truncating Mutations	1	NaN
1407	1407	FGFR3	K508M	6	NaN
1639	1639	FLT1	Amplification	6	NaN
2755	2755	BRAF	G596C	7	NaN

```
In [8]: result.loc[result['TEXT'].isnull(),'TEXT'] = result['Gene'] + ' '+result['Variation']
```

```
In [9]: result[result['ID']==1109]
```

Out[9]:

	ID	Gene	Variation	Class	TEXT
1109	1109	FANCA	S1088F	1	FANCA S1088F

3.1.4. Test, Train and Cross Validation Split

3.1.4.1. Splitting data into train, test and cross validation (64:20:16)

```
In [10]: y_true = result['Class'].values
result.Gene      = result.Gene.str.replace('\s+', '_')
result.Variation = result.Variation.str.replace('\s+', '_')

# split the data into test and train by maintaining same distribution of output variable 'y_true' [stratify=y_true]
X_train, test_df, y_train, y_test = train_test_split(result, y_true, stratify=y_true, test_size=0.2)
# split the train data into train and cross validation by maintaining same distribution of output variable 'y_train' [stratify=y_train]
train_df, cv_df, y_train, y_cv = train_test_split(X_train, y_train, stratify=y_train, test_size=0.2)
```

```
In [11]: y_train
```

```
Out[11]: array([4, 1, 4, ..., 9, 4, 7], dtype=int64)
```

We split the data into train, test and cross validation data sets, preserving the ratio of class distribution in the original data set

```
In [12]: print('Number of data points in train data:', train_df.shape[0])
print('Number of data points in test data:', test_df.shape[0])
print('Number of data points in cross validation data:', cv_df.shape[0])
```

```
Number of data points in train data: 2124
Number of data points in test data: 665
Number of data points in cross validation data: 532
```

```
In [13]: # X_train = X_train.drop(['Class'], axis=1)
# test_df = test_df.drop(['Class'], axis=1)
# cv_df = cv_df.drop(['Class'], axis=1)
```

3.2 Prediction using a 'Random' Model

In a 'Random' Model, we generate the NINE class probabilities randomly such that they sum to 1.

```
In [14]: # This function plots the confusion matrices given y_i, y_i_hat.

def plot_confusion_matrix(test_y, predict_y):
    C = confusion_matrix(test_y, predict_y)
    # C = 9,9 matrix, each cell (i,j) represents number of points of class i a
    re predicted class j

    A =(((C.T)/(C.sum(axis=1))).T)
    #divid each element of the confusion matrix with the sum of elements in th
    at column

    # C = [[1, 2],
    #      [3, 4]]
    # C.T = [[1, 3],
    #          [2, 4]]
    # C.sum(axis = 1) axis=0 corresonds to columns and axis=1 corresponds to
    rows in two dimensional array
    # C.sum(axis =1) = [[3, 7]]
    # ((C.T)/(C.sum(axis=1))) = [[1/3, 3/7]
    #                               [2/3, 4/7]]

    # ((C.T)/(C.sum(axis=1))).T = [[1/3, 2/3]
    #                               [3/7, 4/7]]
    # sum of row elements = 1

    B =(C/C.sum(axis=0))
    #divid each element of the confusion matrix with the sum of elements in th
    at row

    # C = [[1, 2],
    #      [3, 4]]
    # C.sum(axis = 0) axis=0 corresonds to columns and axis=1 corresponds to
    rows in two dimensional array
    # C.sum(axis =0) = [[4, 6]]
    # (C/C.sum(axis=0)) = [[1/4, 2/6],
    #                      [3/4, 4/6]]

    labels = [1,2,3,4,5,6,7,8,9]
    # representing A in heatmap format
    print("-"*20, "Confusion matrix", "*"-20)
    plt.figure(figsize=(20,7))
    sns.heatmap(C, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y
    ticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.show()

    print("-"*20, "Precision matrix (Column Sum=1)", "*"-20)
    plt.figure(figsize=(20,7))
    sns.heatmap(B, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y
    ticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.show()

    # representing B in heatmap format
    print("-"*20, "Recall matrix (Row sum=1)", "*"-20)
    plt.figure(figsize=(20,7))
```

```
    sns.heatmap(A, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y  
    ticklabels=labels)  
    plt.xlabel('Predicted Class')  
    plt.ylabel('Original Class')  
    plt.show()
```

```
In [15]: # we need to generate 9 numbers and the sum of numbers should be 1
# one solution is to generate 9 numbers and divide each of the numbers by their sum
# ref: https://stackoverflow.com/a/18662466/4084039
test_data_len = test_df.shape[0]
cv_data_len = cv_df.shape[0]

# we create a output array that has exactly same size as the CV data
cv_predicted_y = np.zeros((cv_data_len,9))
for i in range(cv_data_len):
    rand_probs = np.random.rand(1,9)
    cv_predicted_y[i] = ((rand_probs/sum(sum(rand_probs)))[0])
print("Log loss on Cross Validation Data using Random Model",log_loss(y_cv,cv_predicted_y, eps=1e-15))

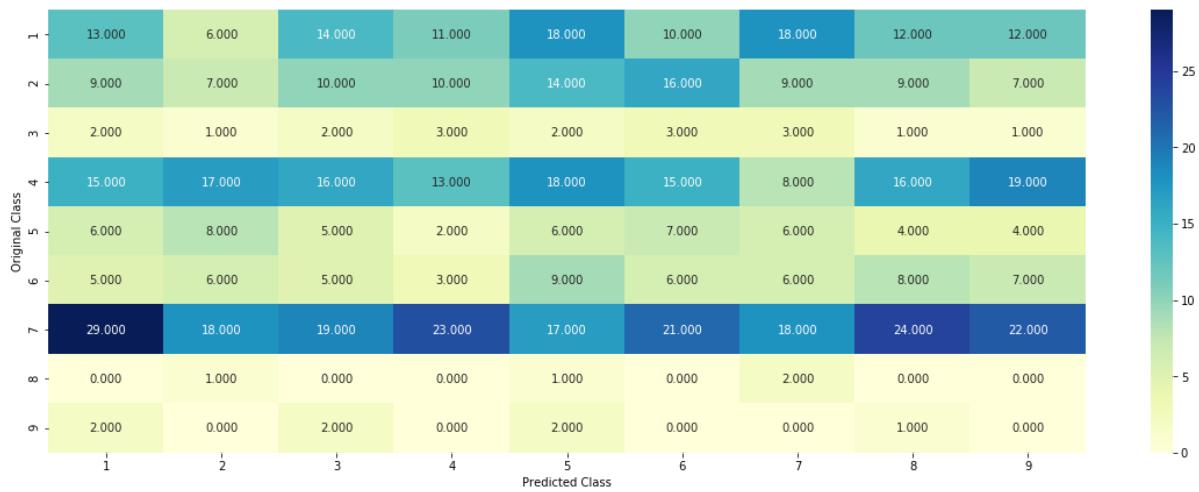
# Test-Set error.
#we create a output array that has exactly same as the test data
test_predicted_y = np.zeros((test_data_len,9))
for i in range(test_data_len):
    rand_probs = np.random.rand(1,9)
    test_predicted_y[i] = ((rand_probs/sum(sum(rand_probs)))[0])
print("Log loss on Test Data using Random Model",log_loss(y_test,test_predicted_y, eps=1e-15))

predicted_y =np.argmax(test_predicted_y, axis=1)
plot_confusion_matrix(y_test, predicted_y+1)
```

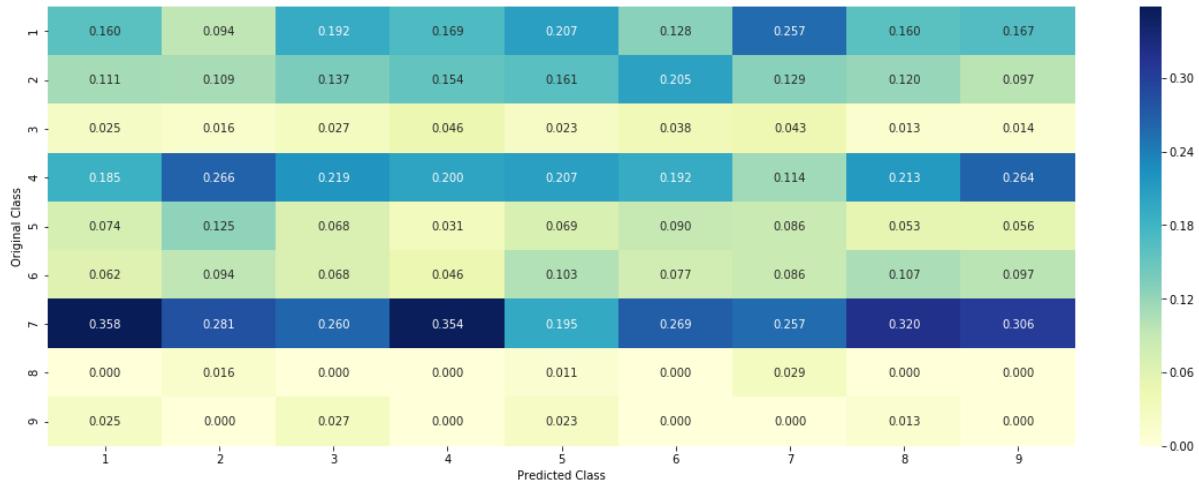
Log loss on Cross Validation Data using Random Model 2.4642508877445484

Log loss on Test Data using Random Model 2.5634994141509297

----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



3.3 Univariate Analysis

```
In [16]: # code for response coding with Laplace smoothing.
# alpha : used for Laplace smoothing
# feature: ['gene', 'variation']
# df: ['train_df', 'test_df', 'cv_df']
# algorithm
# -----
# Consider all unique values and the number of occurrences of given feature in
# train data dataframe
# build a vector (1*9) , the first element = (number of times it occurred in cl
ass1 + 10*alpha / number of time it occurred in total data+90*alpha)
# gv_dict is like a look up table, for every gene it store a (1*9) representat
ion of it
# for a value of feature in df:
# if it is in train data:
# we add the vector that was stored in 'gv_dict' Look up table to 'gv_fea'
# if it is not there is train:
# we add [1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9] to 'gv_fea'
# return 'gv_fea'
# -----
# get_gv_fea_dict: Get Gene variation Feature Dict
def get_gv_fea_dict(alpha, feature, df):
    # value_count: it contains a dict like
    # print(train_df['Gene'].value_counts())
    # output:
    #     {BRCA1      174
    #      TP53      106
    #      EGFR      86
    #      BRCA2      75
    #      PTEN      69
    #      KIT       61
    #      BRAF      60
    #      ERBB2      47
    #      PDGFRA     46
    #      ...}
    # print(train_df['Variation'].value_counts())
    # output:
    # {
    # Truncating_Mutations      63
    # Deletion                  43
    # Amplification              43
    # Fusions                   22
    # Overexpression              3
    # E17K                      3
    # Q61L                      3
    # S222D                     2
    # P130S                     2
    # ...
    # }
    value_count = train_df[feature].value_counts()

    # gv_dict : Gene Variation Dict, which contains the probability array for
    # each gene/variation
    gv_dict = dict()

    # denominator will contain the number of time that particular feature occu
```

```

red in whole data
for i, denominator in value_count.items():
    # vec will contain ( $p(y_i=1/G_i)$ ) probability of gene/variation belongs
    to particular class
    # vec is 9 dimensional vector
    vec = []
    for k in range(1,10):
        # print(train_df.loc[(train_df['Class']==1) & (train_df['Gene']=
        = 'BRCA1')])]

        #           ID   Gene          Variation Class
        # 2470  2470  BRCA1      S1715C     1
        # 2486  2486  BRCA1      S1841R     1
        # 2614  2614  BRCA1      M1R       1
        # 2432  2432  BRCA1      L1657P     1
        # 2567  2567  BRCA1      T1685A     1
        # 2583  2583  BRCA1      E1660G     1
        # 2634  2634  BRCA1      W1718L     1
        # cls_cnt.shape[0] will return the number of rows

    cls_cnt = train_df.loc[(train_df['Class']==k) & (train_df[feature]==i)]

    # cls_cnt.shape[0](numerator) will contain the number of time that
    particular feature occurred in whole data
    vec.append((cls_cnt.shape[0] + alpha*10)/ (denominator + 90*alpha))
)

# we are adding the gene/variation to the dict as key and vec as value
gv_dict[i]=vec
return gv_dict

# Get Gene variation feature
def get_gv_feature(alpha, feature, df):
    # print(gv_dict)
    # {'BRCA1': [0.200757575757575, 0.03787878787878788, 0.068181818181818177, 0.13636363636363635, 0.25, 0.1931818181818181, 0.03787878787878788, 0.03787878787878788, 0.03787878787878788], 'TP53': [0.32142857142857145, 0.061224489795918366, 0.061224489795918366, 0.27040816326530615, 0.061224489795918366, 0.066326530612244902, 0.051020408163265307, 0.051020408163265307, 0.056122448979591837], 'EGFR': [0.0568181818181816, 0.21590909090909091, 0.0625, 0.06818181818177, 0.06818181818177, 0.0625, 0.34659090909090912, 0.0625, 0.0568181818181816], 'BRCA2': [0.1333333333333333, 0.060606060606060608, 0.060606060606060608, 0.0787878787878782, 0.1393939393939394, 0.34545454545454546, 0.060606060606060608, 0.060606060606060608, 0.060606060606060608], 'PTEN': [0.06918238937106917, 0.062893081761006289, 0.069182389937106917, 0.46540880503144655, 0.075471698113207544, 0.062893081761006289, 0.069182389937106917, 0.062893081761006289, 0.062893081761006289], 'KIT': [0.066225165562913912, 0.25165562913907286, 0.072847682119205295, 0.072847682119205295, 0.066225165562913912, 0.066225165562913912, 0.066225165562913912, 0.27152317880794702, 0.066225165562913912, 0.066225165562913912], 'BRAF': [0.066666666666666666, 0.1799999999999999, 0.073333333333333334, 0.07333333333333334, 0.09333333333333338, 0.08000000000000002, 0.2999999999999999, 0.066666666666666666, 0.066666666666666666], '# ...'
# }

```

```

gv_dict = get_gv_fea_dict(alpha, feature, df)
# value_count is similar in get_gv_fea_dict
value_count = train_df[feature].value_counts()

# gv_fea: Gene_variation feature, it will contain the feature for each feature value in the data
gv_fea = []
# for every feature values in the given data frame we will check if it is there in the train data then we will add the feature to gv_fea
# if not we will add [1/9,1/9,1/9,1/9,1/9,1/9,1/9,1/9] to gv_fea
for index, row in df.iterrows():
    if row[feature] in dict(value_count).keys():
        gv_fea.append(gv_dict[row[feature]])
    else:
        gv_fea.append([1/9,1/9,1/9,1/9,1/9,1/9,1/9,1/9])
#
return gv_fea

```

when we calculate the probability of a feature belongs to any particular class, we apply laplace smoothing

- $(\text{numerator} + 10^{10}\alpha) / (\text{denominator} + 90^{10}\alpha)$

3.2.1 Univariate Analysis on Gene Feature

Q1. Gene, What type of feature it is ?

Ans. Gene is a categorical variable

Q2. How many categories are there and How they are distributed?

Q3. How to featurize this Gene feature ?

Ans.there are two ways we can featurize this variable check out this video:

<https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/handling-categorical-and-numerical-features/>

1. One hot Encoding
2. Response coding

We will choose the appropriate featurization based on the ML model we use. For this problem of multi-class classification with categorical features, one-hot encoding is better for Logistic regression while response coding is better for Random Forests.

```
In [17]: #response-coding of the Gene feature
# alpha is used for Laplace smoothing
alpha = 1
# train gene feature
train_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", train_df))
# test gene feature
test_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", test_df))
# cross validation gene feature
cv_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", cv_df))
```

```
In [18]: print("train_gene_feature_responseCoding is converted feature using response coding method. The shape of gene feature:", train_gene_feature_responseCoding.shape)
```

train_gene_feature_responseCoding is converted feature using response coding method. The shape of gene feature: (2124, 9)

```
In [19]: # one-hot encoding of Gene feature.
gene_vectorizer = CountVectorizer()
train_gene_feature_onehotCoding = gene_vectorizer.fit_transform(train_df['Gene'])
test_gene_feature_onehotCoding = gene_vectorizer.transform(test_df['Gene'])
cv_gene_feature_onehotCoding = gene_vectorizer.transform(cv_df['Gene'])
```

```
In [20]: train_df['Gene'].head()
```

```
Out[20]: 602      SMAD4
2906      NF2
823       BRIP1
256       EGFR
2230      PTEN
Name: Gene, dtype: object
```

```
In [21]: print("train_gene_feature_onehotCoding is converted feature using one-hot encoding method. The shape of gene feature:", train_gene_feature_onehotCoding.shape)
```

train_gene_feature_onehotCoding is converted feature using one-hot encoding method. The shape of gene feature: (2124, 231)

3.2.2 Univariate Analysis on Variation Feature

Q7. Variation, What type of feature is it ?

Ans. Variation is a categorical variable

Q8. How many categories are there?

Q9. How to featurize this Variation feature ?

Ans. There are two ways we can featurize this variable check out this video:

<https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/handling-categorical-and-numerical-features/>

1. One hot Encoding
2. Response coding

We will be using both these methods to featurize the Variation Feature

```
In [22]: # alpha is used for Laplace smoothing
alpha = 1
# train gene feature
train_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", train_df))
# test gene feature
test_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", test_df))
# cross validation gene feature
cv_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", cv_df))
```

```
In [23]: print("train_variation_feature_responseCoding is a converted feature using the response coding method. The shape of Variation feature:", train_variation_feature_responseCoding.shape)
```

train_variation_feature_responseCoding is a converted feature using the response coding method. The shape of Variation feature: (2124, 9)

```
In [24]: # one-hot encoding of variation feature.
variation_vectorizer = CountVectorizer()
train_variation_feature_onehotCoding = variation_vectorizer.fit_transform(train_df['Variation'])
test_variation_feature_onehotCoding = variation_vectorizer.transform(test_df['Variation'])
cv_variation_feature_onehotCoding = variation_vectorizer.transform(cv_df['Variation'])
```

```
In [25]: print("train_variation_feature_onehotEncoded is converted feature using the one-hot encoding method. The shape of Variation feature:", train_variation_feature_onehotCoding.shape)
```

train_variation_feature_onehotEncoded is converted feature using the one-hot encoding method. The shape of Variation feature: (2124, 1964)

3.2.3 Univariate Analysis on Text Feature

1. How many unique words are present in train data?
2. How are word frequencies distributed?
3. How to featurize text field?
4. Is the text feature useful in predicting y_i?
5. Is the text feature stable across train, test and CV datasets?

```
In [26]: # cls_text is a data frame
# for every row in data fram consider the 'TEXT'
# split the words by space
# make a dict with those words
# increment its count whenever we see that word

def extract_dictionary_paddle(cls_text):
    dictionary = defaultdict(int)
    for index, row in cls_text.iterrows():
        for word in row['TEXT'].split():
            dictionary[word] +=1
    return dictionary
```

```
In [27]: import math
#https://stackoverflow.com/a/1602964
def get_text_responsecoding(df):
    text_feature_responseCoding = np.zeros((df.shape[0],9))
    for i in range(0,9):
        row_index = 0
        for index, row in df.iterrows():
            sum_prob = 0
            for word in row['TEXT'].split():
                sum_prob += math.log(((dict_list[i].get(word,0)+10 )/(total_di
ct.get(word,0)+90)))
            text_feature_responseCoding[row_index][i] = math.exp(sum_prob/len(
row['TEXT'].split()))
        row_index += 1
    return text_feature_responseCoding
```

```
In [28]: # building a CountVectorizer with all the words that occurred minimum 3 times in train data
text_vectorizer = CountVectorizer(min_df=3,ngram_range=(1,2))
train_text_feature_onehotCoding = text_vectorizer.fit_transform(train_df['TEXT'])
# getting all the feature names (words)
train_text_features= text_vectorizer.get_feature_names()

# train_text_feature_onehotCoding.sum(axis=0).A1 will sum every row and returns (1*number of features) vector
train_textfea_counts = train_text_feature_onehotCoding.sum(axis=0).A1

# zip(list(text_features),text_fea_counts) will zip a word with its number of times it occurred
text_fea_dict = dict(zip(list(train_text_features),train_textfea_counts))

print("Total number of unique words in train data :", len(train_text_features))
```

Total number of unique words in train data : 778766

```
In [29]: dict_list = []
# dict_list =[] contains 9 dictionaries each corresponds to a class
for i in range(1,10):
    cls_text = train_df[train_df['Class']==i]
    # build a word dict based on the words in that class
    dict_list.append(extract_dictionary_paddle(cls_text))
    # append it to dict_list

# dict_list[i] is build on i'th class text data
# total_dict is build on whole training text data
total_dict = extract_dictionary_paddle(train_df)

confuse_array = []
for i in train_text_features:
    ratios = []
    max_val = -1
    for j in range(0,9):
        ratios.append((dict_list[j][i]+10)/(total_dict[i]+90))
    confuse_array.append(ratios)
confuse_array = np.array(confuse_array)
```

```
In [30]: #response coding of text features
train_text_feature_responseCoding = get_text_responsecoding(train_df)
test_text_feature_responseCoding = get_text_responsecoding(test_df)
cv_text_feature_responseCoding = get_text_responsecoding(cv_df)
```

```
In [31]: # https://stackoverflow.com/a/16202486
# we convert each row values such that they sum to 1
train_text_feature_responseCoding = (train_text_feature_responseCoding.T/train_text_feature_responseCoding.sum(axis=1)).T
test_text_feature_responseCoding = (test_text_feature_responseCoding.T/test_text_feature_responseCoding.sum(axis=1)).T
cv_text_feature_responseCoding = (cv_text_feature_responseCoding.T/cv_text_feature_responseCoding.sum(axis=1)).T
```

```
In [32]: train_text_feature_responseCoding[0].sum()
```

```
Out[32]: 1.0000000000000002
```

```
In [33]: # don't forget to normalize every feature
train_text_feature_onehotCoding = normalize(train_text_feature_onehotCoding, axis=0)

# we use the same vectorizer that was trained on train data
test_text_feature_onehotCoding = text_vectorizer.transform(test_df['TEXT'])
# don't forget to normalize every feature
test_text_feature_onehotCoding = normalize(test_text_feature_onehotCoding, axis=0)

# we use the same vectorizer that was trained on train data
cv_text_feature_onehotCoding = text_vectorizer.transform(cv_df['TEXT'])
# don't forget to normalize every feature
cv_text_feature_onehotCoding = normalize(cv_text_feature_onehotCoding, axis=0)
```

```
In [34]: #https://stackoverflow.com/a/2258273/4084039
sorted_text_fea_dict = dict(sorted(text_fea_dict.items(), key=lambda x: x[1] , reverse=True))
sorted_text_occur = np.array(list(sorted_text_fea_dict.values()))
```

```
In [35]: # Number of words for a given frequency.  
print(Counter(sorted_text_occur))
```

Counter({3: 154525, 4: 97649, 5: 72826, 6: 53879, 8: 48381, 7: 43720, 9: 2889, 9, 11: 25697, 10: 24687, 12: 21449, 13: 17174, 14: 12399, 16: 11537, 15: 1096, 5, 18: 8624, 20: 7878, 17: 7680, 21: 6299, 19: 5538, 22: 5047, 23: 4966, 24: 4638, 30: 4215, 25: 3941, 39: 3553, 28: 3447, 34: 3365, 26: 3239, 27: 3171, 3, 6: 3011, 29: 2380, 32: 2353, 49: 2308, 31: 2300, 33: 2180, 35: 2021, 40: 193, 6, 37: 1728, 38: 1483, 41: 1463, 42: 1452, 70: 1373, 44: 1252, 43: 1246, 57: 1228, 45: 1156, 50: 1084, 48: 1046, 46: 1033, 51: 972, 47: 952, 52: 860, 56: 825, 53: 817, 72: 791, 55: 758, 54: 744, 60: 737, 58: 715, 59: 665, 61: 648, 63: 616, 64: 566, 65: 558, 62: 557, 71: 539, 66: 526, 68: 503, 78: 492, 69: 483, 73: 478, 74: 459, 67: 458, 77: 434, 75: 434, 80: 411, 79: 406, 76: 388, 8, 1: 365, 82: 325, 84: 323, 83: 311, 85: 310, 88: 309, 98: 287, 90: 287, 86: 28, 6, 87: 267, 93: 265, 96: 251, 91: 251, 92: 249, 94: 247, 100: 245, 89: 238, 9, 5: 231, 99: 229, 104: 224, 108: 217, 105: 215, 97: 212, 109: 205, 106: 201, 1, 14: 198, 102: 197, 112: 196, 110: 196, 101: 194, 107: 191, 117: 185, 125: 17, 1, 113: 167, 103: 164, 111: 161, 119: 157, 140: 155, 116: 155, 118: 152, 115: 149, 123: 142, 120: 141, 122: 137, 136: 134, 124: 134, 138: 133, 135: 133, 12, 6: 133, 121: 133, 131: 127, 147: 126, 129: 126, 128: 126, 145: 124, 127: 123, 134: 121, 133: 120, 130: 120, 132: 117, 156: 114, 157: 107, 142: 106, 141: 10, 5, 144: 104, 143: 101, 139: 101, 148: 99, 137: 97, 162: 94, 170: 93, 150: 93, 152: 92, 155: 91, 153: 91, 146: 90, 159: 89, 149: 88, 151: 86, 171: 85, 154: 85, 168: 84, 172: 83, 161: 82, 165: 79, 160: 79, 163: 78, 164: 77, 169: 76, 1, 80: 75, 166: 75, 158: 71, 192: 69, 167: 69, 179: 67, 178: 67, 175: 67, 181: 6, 5, 198: 64, 196: 64, 184: 64, 210: 63, 176: 63, 214: 62, 190: 62, 189: 62, 21, 2: 61, 185: 61, 174: 61, 199: 59, 195: 59, 204: 58, 231: 57, 173: 57, 183: 5, 6, 186: 55, 187: 54, 217: 53, 208: 53, 207: 53, 177: 53, 202: 51, 205: 49, 18, 8: 49, 209: 48, 206: 48, 197: 48, 194: 48, 193: 48, 182: 48, 216: 47, 201: 4, 6, 213: 45, 240: 44, 225: 44, 256: 43, 247: 43, 237: 43, 226: 43, 252: 42, 24, 8: 42, 245: 42, 220: 42, 233: 41, 200: 41, 224: 40, 222: 40, 218: 40, 291: 3, 9, 211: 39, 241: 38, 234: 38, 221: 38, 191: 38, 275: 37, 258: 37, 244: 37, 21, 9: 37, 215: 37, 203: 37, 253: 36, 242: 36, 238: 36, 236: 36, 230: 36, 223: 3, 6, 273: 35, 266: 35, 265: 35, 227: 35, 284: 34, 254: 34, 229: 34, 304: 33, 28, 6: 33, 277: 33, 257: 33, 239: 33, 235: 33, 228: 33, 250: 32, 243: 32, 232: 3, 2, 294: 31, 251: 31, 302: 30, 279: 30, 274: 30, 267: 30, 259: 30, 288: 29, 28, 5: 29, 270: 29, 328: 28, 318: 28, 297: 28, 283: 28, 272: 28, 262: 28, 261: 2, 8, 326: 27, 278: 27, 255: 27, 306: 26, 282: 26, 280: 26, 269: 26, 336: 25, 32, 7: 25, 317: 25, 359: 24, 358: 24, 307: 24, 300: 24, 293: 24, 287: 24, 271: 2, 4, 263: 24, 260: 24, 339: 23, 338: 23, 333: 23, 319: 23, 311: 23, 308: 23, 29, 9: 23, 289: 23, 268: 23, 367: 22, 348: 22, 341: 22, 340: 22, 330: 22, 315: 2, 2, 313: 22, 276: 22, 264: 22, 249: 22, 409: 21, 343: 21, 323: 21, 301: 21, 29, 8: 21, 246: 21, 402: 20, 376: 20, 361: 20, 331: 20, 309: 20, 296: 20, 281: 2, 0, 357: 19, 350: 19, 337: 19, 334: 19, 329: 19, 305: 19, 290: 19, 380: 18, 37, 9: 18, 368: 18, 362: 18, 353: 18, 347: 18, 345: 18, 342: 18, 332: 18, 324: 1, 8, 295: 18, 415: 17, 414: 17, 396: 17, 392: 17, 370: 17, 369: 17, 354: 17, 34, 6: 17, 316: 17, 314: 17, 312: 17, 310: 17, 303: 17, 292: 17, 447: 16, 431: 1, 6, 405: 16, 383: 16, 365: 16, 349: 16, 503: 15, 484: 15, 440: 15, 437: 15, 43, 0: 15, 424: 15, 420: 15, 418: 15, 412: 15, 411: 15, 397: 15, 391: 15, 390: 1, 5, 375: 15, 364: 15, 360: 15, 355: 15, 325: 15, 322: 15, 493: 14, 465: 14, 44, 6: 14, 442: 14, 441: 14, 422: 14, 419: 14, 407: 14, 398: 14, 387: 14, 351: 1, 4, 335: 14, 321: 14, 663: 13, 566: 13, 522: 13, 491: 13, 472: 13, 471: 13, 45, 6: 13, 451: 13, 433: 13, 432: 13, 381: 13, 373: 13, 363: 13, 344: 13, 320: 1, 3, 659: 12, 581: 12, 571: 12, 556: 12, 535: 12, 533: 12, 490: 12, 470: 12, 46, 6: 12, 462: 12, 453: 12, 429: 12, 427: 12, 413: 12, 395: 12, 393: 12, 388: 1, 2, 378: 12, 366: 12, 356: 12, 1139: 11, 673: 11, 604: 11, 559: 11, 532: 11, 5, 31: 11, 515: 11, 497: 11, 495: 11, 494: 11, 468: 11, 460: 11, 459: 11, 457: 1, 1, 417: 11, 410: 11, 403: 11, 400: 11, 394: 11, 372: 11, 352: 11, 768: 10, 68, 8: 10, 593: 10, 582: 10, 560: 10, 555: 10, 552: 10, 539: 10, 525: 10, 514: 1, 0, 508: 10, 506: 10, 480: 10, 455: 10, 445: 10, 435: 10, 426: 10, 408: 10, 40}

6: 10, 404: 10, 377: 10, 1047: 9, 657: 9, 650: 9, 621: 9, 597: 9, 587: 9, 57
 3: 9, 562: 9, 547: 9, 545: 9, 534: 9, 526: 9, 501: 9, 498: 9, 478: 9, 477: 9,
 473: 9, 463: 9, 452: 9, 450: 9, 443: 9, 436: 9, 423: 9, 399: 9, 386: 9, 385:
 9, 384: 9, 371: 9, 902: 8, 779: 8, 776: 8, 713: 8, 700: 8, 684: 8, 670: 8, 66
 1: 8, 654: 8, 627: 8, 622: 8, 620: 8, 614: 8, 606: 8, 602: 8, 600: 8, 572: 8,
 567: 8, 557: 8, 524: 8, 523: 8, 517: 8, 502: 8, 500: 8, 496: 8, 489: 8, 467:
 8, 461: 8, 444: 8, 434: 8, 416: 8, 389: 8, 1201: 7, 1107: 7, 1102: 7, 1063:
 7, 872: 7, 844: 7, 803: 7, 798: 7, 774: 7, 764: 7, 760: 7, 709: 7, 671: 7, 66
 8: 7, 666: 7, 664: 7, 652: 7, 646: 7, 644: 7, 638: 7, 637: 7, 625: 7, 613: 7,
 609: 7, 603: 7, 595: 7, 591: 7, 578: 7, 568: 7, 553: 7, 538: 7, 537: 7, 530:
 7, 529: 7, 527: 7, 511: 7, 509: 7, 505: 7, 488: 7, 485: 7, 483: 7, 482: 7, 47
 9: 7, 476: 7, 475: 7, 464: 7, 458: 7, 454: 7, 428: 7, 401: 7, 1344: 6, 1285:
 6, 1156: 6, 1088: 6, 1029: 6, 1025: 6, 1010: 6, 992: 6, 980: 6, 958: 6, 943:
 6, 929: 6, 919: 6, 911: 6, 909: 6, 843: 6, 819: 6, 813: 6, 808: 6, 807: 6, 79
 0: 6, 786: 6, 773: 6, 761: 6, 757: 6, 722: 6, 707: 6, 706: 6, 699: 6, 677: 6,
 674: 6, 669: 6, 667: 6, 660: 6, 648: 6, 636: 6, 635: 6, 634: 6, 633: 6, 629:
 6, 628: 6, 618: 6, 617: 6, 607: 6, 590: 6, 589: 6, 588: 6, 585: 6, 576: 6, 57
 0: 6, 558: 6, 546: 6, 541: 6, 540: 6, 521: 6, 520: 6, 519: 6, 518: 6, 516: 6,
 513: 6, 510: 6, 492: 6, 487: 6, 481: 6, 438: 6, 421: 6, 374: 6, 1926: 5, 166
 1: 5, 1641: 5, 1380: 5, 1378: 5, 1301: 5, 1298: 5, 1271: 5, 1126: 5, 1110: 5,
 1052: 5, 1007: 5, 960: 5, 946: 5, 940: 5, 936: 5, 933: 5, 928: 5, 917: 5, 87
 3: 5, 871: 5, 867: 5, 863: 5, 861: 5, 851: 5, 850: 5, 838: 5, 836: 5, 827: 5,
 815: 5, 795: 5, 788: 5, 780: 5, 777: 5, 775: 5, 772: 5, 769: 5, 759: 5, 754:
 5, 751: 5, 750: 5, 748: 5, 742: 5, 739: 5, 738: 5, 733: 5, 727: 5, 719: 5, 71
 8: 5, 715: 5, 703: 5, 702: 5, 701: 5, 687: 5, 680: 5, 649: 5, 639: 5, 626: 5,
 616: 5, 608: 5, 605: 5, 601: 5, 598: 5, 583: 5, 580: 5, 575: 5, 569: 5, 565:
 5, 564: 5, 563: 5, 551: 5, 549: 5, 544: 5, 543: 5, 528: 5, 504: 5, 486: 5, 47
 4: 5, 469: 5, 425: 5, 2458: 4, 2220: 4, 2042: 4, 1973: 4, 1969: 4, 1819: 4, 1
 725: 4, 1648: 4, 1630: 4, 1597: 4, 1475: 4, 1350: 4, 1330: 4, 1311: 4, 1284:
 4, 1269: 4, 1263: 4, 1254: 4, 1230: 4, 1223: 4, 1207: 4, 1204: 4, 1199: 4, 11
 95: 4, 1182: 4, 1169: 4, 1164: 4, 1142: 4, 1127: 4, 1119: 4, 1089: 4, 1071:
 4, 1068: 4, 1045: 4, 1043: 4, 1042: 4, 1015: 4, 1014: 4, 1011: 4, 1004: 4, 10
 01: 4, 1000: 4, 998: 4, 997: 4, 988: 4, 969: 4, 964: 4, 955: 4, 954: 4, 938:
 4, 932: 4, 913: 4, 910: 4, 906: 4, 898: 4, 897: 4, 892: 4, 890: 4, 888: 4, 88
 6: 4, 881: 4, 874: 4, 869: 4, 865: 4, 853: 4, 840: 4, 834: 4, 828: 4, 825: 4,
 824: 4, 817: 4, 805: 4, 804: 4, 797: 4, 781: 4, 778: 4, 765: 4, 763: 4, 758:
 4, 752: 4, 749: 4, 741: 4, 740: 4, 730: 4, 729: 4, 728: 4, 724: 4, 720: 4, 71
 7: 4, 712: 4, 711: 4, 710: 4, 704: 4, 698: 4, 697: 4, 696: 4, 693: 4, 692: 4,
 689: 4, 683: 4, 678: 4, 676: 4, 672: 4, 653: 4, 647: 4, 645: 4, 643: 4, 631:
 4, 623: 4, 612: 4, 611: 4, 610: 4, 596: 4, 594: 4, 577: 4, 561: 4, 548: 4, 54
 2: 4, 536: 4, 512: 4, 507: 4, 449: 4, 448: 4, 439: 4, 382: 4, 4746: 3, 3483:
 3, 3107: 3, 2496: 3, 2391: 3, 2389: 3, 2364: 3, 2347: 3, 2338: 3, 2327: 3, 21
 88: 3, 2175: 3, 2122: 3, 2091: 3, 2009: 3, 1919: 3, 1918: 3, 1913: 3, 1882:
 3, 1859: 3, 1837: 3, 1807: 3, 1801: 3, 1789: 3, 1743: 3, 1724: 3, 1659: 3, 16
 18: 3, 1615: 3, 1614: 3, 1603: 3, 1592: 3, 1587: 3, 1584: 3, 1577: 3, 1555:
 3, 1529: 3, 1528: 3, 1526: 3, 1514: 3, 1494: 3, 1493: 3, 1482: 3, 1458: 3, 14
 46: 3, 1430: 3, 1427: 3, 1404: 3, 1401: 3, 1398: 3, 1394: 3, 1391: 3, 1387:
 3, 1377: 3, 1374: 3, 1373: 3, 1372: 3, 1371: 3, 1362: 3, 1351: 3, 1348: 3, 13
 46: 3, 1342: 3, 1340: 3, 1339: 3, 1324: 3, 1308: 3, 1307: 3, 1295: 3, 1289:
 3, 1280: 3, 1278: 3, 1276: 3, 1264: 3, 1262: 3, 1252: 3, 1235: 3, 1225: 3, 12
 24: 3, 1222: 3, 1215: 3, 1203: 3, 1197: 3, 1179: 3, 1173: 3, 1172: 3, 1160:
 3, 1157: 3, 1153: 3, 1151: 3, 1148: 3, 1143: 3, 1137: 3, 1136: 3, 1135: 3, 11
 34: 3, 1132: 3, 1131: 3, 1124: 3, 1123: 3, 1114: 3, 1109: 3, 1104: 3, 1100:
 3, 1095: 3, 1084: 3, 1077: 3, 1048: 3, 1041: 3, 1034: 3, 1023: 3, 1022: 3, 10
 05: 3, 1002: 3, 999: 3, 996: 3, 994: 3, 993: 3, 983: 3, 970: 3, 965: 3, 963:
 3, 961: 3, 952: 3, 948: 3, 926: 3, 924: 3, 921: 3, 916: 3, 912: 3, 907: 3, 90
 0: 3, 894: 3, 891: 3, 887: 3, 883: 3, 877: 3, 876: 3, 868: 3, 862: 3, 857: 3,

856: 3, 855: 3, 847: 3, 835: 3, 833: 3, 832: 3, 826: 3, 811: 3, 809: 3, 801: 3, 799: 3, 793: 3, 792: 3, 791: 3, 789: 3, 771: 3, 766: 3, 762: 3, 756: 3, 747: 3, 746: 3, 736: 3, 734: 3, 732: 3, 731: 3, 726: 3, 721: 3, 714: 3, 695: 3, 694: 3, 691: 3, 690: 3, 682: 3, 681: 3, 665: 3, 662: 3, 658: 3, 656: 3, 655: 3, 651: 3, 641: 3, 615: 3, 599: 3, 592: 3, 584: 3, 574: 3, 554: 3, 550: 3, 12594: 2, 10736: 2, 8017: 2, 6528: 2, 6317: 2, 6147: 2, 5366: 2, 5062: 2, 4926: 2, 4592: 2, 4549: 2, 4534: 2, 4475: 2, 4474: 2, 4471: 2, 4395: 2, 4332: 2, 4241: 2, 4204: 2, 4196: 2, 4176: 2, 4099: 2, 4085: 2, 3988: 2, 3956: 2, 3914: 2, 3877: 2, 3775: 2, 3753: 2, 3730: 2, 3638: 2, 3619: 2, 3601: 2, 3514: 2, 3441: 2, 3436: 2, 3423: 2, 3394: 2, 3390: 2, 3379: 2, 3371: 2, 3340: 2, 3303: 2, 3264: 2, 3214: 2, 3178: 2, 3142: 2, 3125: 2, 3080: 2, 2984: 2, 2982: 2, 2950: 2, 2935: 2, 2922: 2, 2921: 2, 2835: 2, 2830: 2, 2802: 2, 2766: 2, 2736: 2, 2692: 2, 2668: 2, 2655: 2, 2624: 2, 2609: 2, 2604: 2, 2587: 2, 2582: 2, 2568: 2, 2552: 2, 2518: 2, 2493: 2, 2492: 2, 2485: 2, 2481: 2, 2462: 2, 2456: 2, 2433: 2, 2430: 2, 2429: 2, 2424: 2, 2415: 2, 2414: 2, 2405: 2, 2403: 2, 2390: 2, 2388: 2, 2387: 2, 2357: 2, 2336: 2, 2326: 2, 2310: 2, 2299: 2, 2284: 2, 2281: 2, 2271: 2, 2255: 2, 2218: 2, 2209: 2, 2201: 2, 2200: 2, 2194: 2, 2171: 2, 2164: 2, 2159: 2, 2157: 2, 2153: 2, 2146: 2, 2145: 2, 2139: 2, 2118: 2, 2113: 2, 2107: 2, 2104: 2, 2095: 2, 2089: 2, 2084: 2, 2083: 2, 2065: 2, 2061: 2, 2057: 2, 2034: 2, 2032: 2, 2018: 2, 2008: 2, 2005: 2, 2001: 2, 1999: 2, 1997: 2, 1992: 2, 1990: 2, 1981: 2, 1974: 2, 1961: 2, 1960: 2, 1952: 2, 1949: 2, 1944: 2, 1938: 2, 1931: 2, 1900: 2, 1874: 2, 1872: 2, 1867: 2, 1864: 2, 1843: 2, 1832: 2, 1817: 2, 1816: 2, 1815: 2, 1813: 2, 1806: 2, 1795: 2, 1788: 2, 1787: 2, 1780: 2, 1726: 2, 1721: 2, 1713: 2, 1707: 2, 1705: 2, 1699: 2, 1687: 2, 1685: 2, 1682: 2, 1672: 2, 1670: 2, 1667: 2, 1655: 2, 1653: 2, 1652: 2, 1650: 2, 1647: 2, 1636: 2, 1634: 2, 1626: 2, 1610: 2, 1606: 2, 1600: 2, 1598: 2, 1581: 2, 1579: 2, 1576: 2, 1574: 2, 1570: 2, 1563: 2, 1561: 2, 1560: 2, 1551: 2, 1546: 2, 1545: 2, 1543: 2, 1539: 2, 1538: 2, 1535: 2, 1533: 2, 1527: 2, 1525: 2, 1524: 2, 1523: 2, 1516: 2, 1511: 2, 1510: 2, 1509: 2, 1496: 2, 1492: 2, 1478: 2, 1477: 2, 1471: 2, 1461: 2, 1459: 2, 1457: 2, 1454: 2, 1453: 2, 1448: 2, 1443: 2, 1440: 2, 1435: 2, 1433: 2, 1431: 2, 1428: 2, 1426: 2, 1425: 2, 1416: 2, 1406: 2, 1402: 2, 1397: 2, 1388: 2, 1386: 2, 1385: 2, 1383: 2, 1382: 2, 1379: 2, 1367: 2, 1359: 2, 1354: 2, 1353: 2, 1343: 2, 1337: 2, 1336: 2, 1335: 2, 1329: 2, 1320: 2, 1316: 2, 1313: 2, 1312: 2, 1310: 2, 1305: 2, 1300: 2, 1294: 2, 1292: 2, 1291: 2, 1288: 2, 1286: 2, 1277: 2, 1275: 2, 1273: 2, 1272: 2, 1255: 2, 1249: 2, 1245: 2, 1244: 2, 1243: 2, 1240: 2, 1239: 2, 1238: 2, 1236: 2, 1234: 2, 1233: 2, 1231: 2, 1226: 2, 1220: 2, 1217: 2, 1212: 2, 1205: 2, 1200: 2, 1196: 2, 1189: 2, 1186: 2, 1185: 2, 1181: 2, 1177: 2, 1175: 2, 1174: 2, 1171: 2, 1170: 2, 1168: 2, 1166: 2, 1162: 2, 1152: 2, 1150: 2, 1145: 2, 1141: 2, 1130: 2, 1129: 2, 1128: 2, 1122: 2, 1121: 2, 1120: 2, 1118: 2, 1116: 2, 1108: 2, 1103: 2, 1098: 2, 1093: 2, 1079: 2, 1076: 2, 1072: 2, 1062: 2, 1061: 2, 1057: 2, 1056: 2, 1055: 2, 1054: 2, 1053: 2, 1051: 2, 1039: 2, 1028: 2, 1026: 2, 1021: 2, 1020: 2, 1019: 2, 1018: 2, 1009: 2, 1003: 2, 990: 2, 987: 2, 986: 2, 985: 2, 984: 2, 977: 2, 976: 2, 975: 2, 966: 2, 962: 2, 959: 2, 957: 2, 947: 2, 944: 2, 941: 2, 935: 2, 934: 2, 931: 2, 927: 2, 918: 2, 914: 2, 904: 2, 899: 2, 896: 2, 895: 2, 893: 2, 889: 2, 885: 2, 882: 2, 880: 2, 875: 2, 860: 2, 859: 2, 858: 2, 849: 2, 848: 2, 845: 2, 841: 2, 839: 2, 831: 2, 823: 2, 821: 2, 818: 2, 816: 2, 814: 2, 806: 2, 802: 2, 800: 2, 796: 2, 784: 2, 770: 2, 767: 2, 755: 2, 753: 2, 744: 2, 737: 2, 735: 2, 725: 2, 723: 2, 716: 2, 708: 2, 705: 2, 685: 2, 675: 2, 640: 2, 632: 2, 630: 2, 624: 2, 619: 2, 586: 2, 579: 2, 499: 2, 155172: 1, 117350: 1, 81000: 1, 68577: 1, 68390: 1, 68385: 1, 67725: 1, 66804: 1, 65330: 1, 63861: 1, 56722: 1, 53939: 1, 49464: 1, 48787: 1, 47859: 1, 47186: 1, 45792: 1, 44081: 1, 43026: 1, 42807: 1, 42230: 1, 41275: 1, 39692: 1, 39309: 1, 38870: 1, 38692: 1, 37750: 1, 37103: 1, 36751: 1, 36557: 1, 36287: 1, 35351: 1, 34997: 1, 34746: 1, 33945: 1, 33555: 1, 33062: 1, 31854: 1, 29532: 1, 28390: 1, 28362: 1, 26276: 1, 25873: 1, 25839: 1, 25609: 1, 25519: 1, 25146: 1, 25074: 1, 2477

9: 1, 24370: 1, 24328: 1, 24090: 1, 24051: 1, 23897: 1, 23722: 1, 22701: 1, 2
 2666: 1, 22252: 1, 21818: 1, 21787: 1, 21548: 1, 21464: 1, 21156: 1, 21044:
 1, 20455: 1, 20167: 1, 19990: 1, 19877: 1, 19858: 1, 19771: 1, 19584: 1, 1956
 4: 1, 19313: 1, 19149: 1, 19044: 1, 18858: 1, 18800: 1, 18779: 1, 18766: 1, 1
 8728: 1, 18498: 1, 18466: 1, 18079: 1, 17958: 1, 17824: 1, 17707: 1, 17614:
 1, 17490: 1, 17484: 1, 17476: 1, 17465: 1, 17391: 1, 17197: 1, 17167: 1, 1707
 4: 1, 17063: 1, 17021: 1, 16959: 1, 16867: 1, 16829: 1, 16713: 1, 16386: 1, 1
 6379: 1, 15896: 1, 15781: 1, 15761: 1, 15756: 1, 15713: 1, 15635: 1, 15631:
 1, 15602: 1, 15550: 1, 15484: 1, 15166: 1, 15131: 1, 15117: 1, 14901: 1, 1489
 7: 1, 14896: 1, 14672: 1, 14552: 1, 14434: 1, 14267: 1, 14082: 1, 14070: 1, 1
 4065: 1, 14013: 1, 13921: 1, 13894: 1, 13816: 1, 13787: 1, 13698: 1, 13657:
 1, 13469: 1, 13340: 1, 13302: 1, 13131: 1, 13036: 1, 13001: 1, 13000: 1, 1274
 6: 1, 12725: 1, 12722: 1, 12709: 1, 12663: 1, 12661: 1, 12626: 1, 12588: 1, 1
 2552: 1, 12381: 1, 12322: 1, 12306: 1, 12294: 1, 12279: 1, 12230: 1, 12209:
 1, 12206: 1, 12177: 1, 12151: 1, 12132: 1, 12127: 1, 12055: 1, 12047: 1, 1200
 4: 1, 11974: 1, 11911: 1, 11893: 1, 11888: 1, 11852: 1, 11813: 1, 11780: 1, 1
 1731: 1, 11723: 1, 11706: 1, 11653: 1, 11558: 1, 11557: 1, 11544: 1, 11516:
 1, 11482: 1, 11165: 1, 11103: 1, 11091: 1, 11066: 1, 10971: 1, 10835: 1, 1082
 0: 1, 10735: 1, 10664: 1, 10641: 1, 10570: 1, 10540: 1, 10471: 1, 10468: 1, 1
 0432: 1, 10404: 1, 10388: 1, 10377: 1, 10368: 1, 10362: 1, 10305: 1, 10182:
 1, 10156: 1, 10117: 1, 10115: 1, 10099: 1, 10052: 1, 10039: 1, 10034: 1, 999
 6: 1, 9972: 1, 9937: 1, 9880: 1, 9759: 1, 9743: 1, 9707: 1, 9574: 1, 9508: 1,
 9489: 1, 9461: 1, 9431: 1, 9410: 1, 9406: 1, 9372: 1, 9357: 1, 9318: 1, 9281:
 1, 9257: 1, 9240: 1, 9226: 1, 9224: 1, 9199: 1, 9181: 1, 9173: 1, 9147: 1, 91
 42: 1, 9030: 1, 9028: 1, 9010: 1, 8970: 1, 8932: 1, 8920: 1, 8881: 1, 8871:
 1, 8840: 1, 8824: 1, 8781: 1, 8679: 1, 8546: 1, 8539: 1, 8521: 1, 8503: 1, 85
 00: 1, 8494: 1, 8491: 1, 8484: 1, 8481: 1, 8460: 1, 8432: 1, 8400: 1, 8388:
 1, 8360: 1, 8320: 1, 8294: 1, 8274: 1, 8272: 1, 8249: 1, 8219: 1, 8218: 1, 81
 65: 1, 8154: 1, 8150: 1, 8142: 1, 8140: 1, 8085: 1, 8057: 1, 8041: 1, 7984:
 1, 7966: 1, 7946: 1, 7916: 1, 7856: 1, 7849: 1, 7841: 1, 7830: 1, 7828: 1, 78
 18: 1, 7813: 1, 7775: 1, 7771: 1, 7743: 1, 7696: 1, 7689: 1, 7661: 1, 7639:
 1, 7631: 1, 7605: 1, 7604: 1, 7579: 1, 7568: 1, 7557: 1, 7542: 1, 7541: 1, 75
 22: 1, 7461: 1, 7439: 1, 7420: 1, 7411: 1, 7409: 1, 7341: 1, 7328: 1, 7324:
 1, 7298: 1, 7296: 1, 7285: 1, 7273: 1, 7267: 1, 7252: 1, 7197: 1, 7184: 1, 71
 61: 1, 7159: 1, 7158: 1, 7154: 1, 7136: 1, 7116: 1, 7113: 1, 7112: 1, 7085:
 1, 7084: 1, 7060: 1, 7034: 1, 7029: 1, 7007: 1, 6976: 1, 6965: 1, 6958: 1, 69
 55: 1, 6938: 1, 6930: 1, 6916: 1, 6914: 1, 6913: 1, 6910: 1, 6902: 1, 6867:
 1, 6843: 1, 6829: 1, 6819: 1, 6808: 1, 6757: 1, 6724: 1, 6720: 1, 6696: 1, 66
 82: 1, 6677: 1, 6674: 1, 6664: 1, 6633: 1, 6626: 1, 6621: 1, 6598: 1, 6585:
 1, 6572: 1, 6532: 1, 6525: 1, 6518: 1, 6511: 1, 6504: 1, 6481: 1, 6473: 1, 64
 53: 1, 6444: 1, 6442: 1, 6421: 1, 6401: 1, 6399: 1, 6374: 1, 6372: 1, 6331:
 1, 6318: 1, 6291: 1, 6279: 1, 6273: 1, 6260: 1, 6259: 1, 6251: 1, 6247: 1, 62
 44: 1, 6205: 1, 6198: 1, 6183: 1, 6178: 1, 6172: 1, 6153: 1, 6129: 1, 6088:
 1, 6083: 1, 6082: 1, 6079: 1, 6074: 1, 6060: 1, 6042: 1, 6028: 1, 6019: 1, 60
 09: 1, 6005: 1, 6000: 1, 5984: 1, 5973: 1, 5967: 1, 5961: 1, 5943: 1, 5934:
 1, 5932: 1, 5912: 1, 5903: 1, 5895: 1, 5888: 1, 5871: 1, 5859: 1, 5858: 1, 58
 03: 1, 5783: 1, 5775: 1, 5771: 1, 5766: 1, 5743: 1, 5731: 1, 5715: 1, 5707:
 1, 5705: 1, 5695: 1, 5658: 1, 5639: 1, 5618: 1, 5617: 1, 5612: 1, 5597: 1, 55
 57: 1, 5547: 1, 5527: 1, 5526: 1, 5517: 1, 5516: 1, 5511: 1, 5501: 1, 5500:
 1, 5487: 1, 5479: 1, 5478: 1, 5467: 1, 5451: 1, 5437: 1, 5436: 1, 5426: 1, 54
 10: 1, 5406: 1, 5405: 1, 5404: 1, 5403: 1, 5402: 1, 5393: 1, 5391: 1, 5361:
 1, 5359: 1, 5350: 1, 5298: 1, 5289: 1, 5265: 1, 5255: 1, 5250: 1, 5222: 1, 52
 21: 1, 5215: 1, 5206: 1, 5193: 1, 5190: 1, 5177: 1, 5175: 1, 5167: 1, 5165:
 1, 5154: 1, 5144: 1, 5134: 1, 5123: 1, 5111: 1, 5100: 1, 5088: 1, 5082: 1, 50
 76: 1, 5064: 1, 5054: 1, 5053: 1, 5052: 1, 5037: 1, 5033: 1, 5023: 1, 5019:
 1, 5015: 1, 5014: 1, 5008: 1, 5005: 1, 4988: 1, 4975: 1, 4971: 1, 4970: 1, 49
 38: 1, 4911: 1, 4901: 1, 4896: 1, 4890: 1, 4887: 1, 4867: 1, 4865: 1, 4841:

1, 4829: 1, 4823: 1, 4817: 1, 4803: 1, 4787: 1, 4781: 1, 4780: 1, 4774: 1, 47
66: 1, 4752: 1, 4751: 1, 4741: 1, 4730: 1, 4708: 1, 4704: 1, 4688: 1, 4684:
1, 4673: 1, 4668: 1, 4667: 1, 4657: 1, 4639: 1, 4636: 1, 4613: 1, 4612: 1, 46
04: 1, 4575: 1, 4566: 1, 4557: 1, 4511: 1, 4506: 1, 4501: 1, 4498: 1, 4488:
1, 4481: 1, 4459: 1, 4458: 1, 4453: 1, 4448: 1, 4446: 1, 4443: 1, 4438: 1, 44
36: 1, 4426: 1, 4420: 1, 4419: 1, 4417: 1, 4406: 1, 4399: 1, 4379: 1, 4369:
1, 4367: 1, 4363: 1, 4356: 1, 4334: 1, 4316: 1, 4309: 1, 4306: 1, 4294: 1, 42
91: 1, 4286: 1, 4280: 1, 4276: 1, 4273: 1, 4267: 1, 4253: 1, 4249: 1, 4248:
1, 4236: 1, 4226: 1, 4225: 1, 4221: 1, 4201: 1, 4200: 1, 4188: 1, 4163: 1, 41
44: 1, 4141: 1, 4139: 1, 4126: 1, 4125: 1, 4115: 1, 4112: 1, 4097: 1, 4089:
1, 4076: 1, 4071: 1, 4069: 1, 4068: 1, 4057: 1, 4051: 1, 4040: 1, 4038: 1, 40
37: 1, 4027: 1, 4026: 1, 4025: 1, 4018: 1, 4017: 1, 3991: 1, 3985: 1, 3977:
1, 3975: 1, 3971: 1, 3967: 1, 3963: 1, 3962: 1, 3954: 1, 3952: 1, 3950: 1, 39
35: 1, 3930: 1, 3919: 1, 3916: 1, 3915: 1, 3906: 1, 3901: 1, 3900: 1, 3899:
1, 3891: 1, 3876: 1, 3875: 1, 3863: 1, 3861: 1, 3849: 1, 3836: 1, 3834: 1, 38
31: 1, 3828: 1, 3825: 1, 3824: 1, 3822: 1, 3813: 1, 3803: 1, 3788: 1, 3787:
1, 3783: 1, 3779: 1, 3777: 1, 3776: 1, 3773: 1, 3761: 1, 3760: 1, 3746: 1, 37
39: 1, 3736: 1, 3733: 1, 3728: 1, 3727: 1, 3720: 1, 3719: 1, 3712: 1, 3707:
1, 3698: 1, 3687: 1, 3685: 1, 3682: 1, 3674: 1, 3666: 1, 3661: 1, 3659: 1, 36
54: 1, 3652: 1, 3644: 1, 3642: 1, 3622: 1, 3618: 1, 3609: 1, 3607: 1, 3602:
1, 3600: 1, 3598: 1, 3593: 1, 3588: 1, 3584: 1, 3581: 1, 3578: 1, 3576: 1, 35
75: 1, 3566: 1, 3564: 1, 3560: 1, 3559: 1, 3557: 1, 3550: 1, 3546: 1, 3542:
1, 3538: 1, 3533: 1, 3522: 1, 3519: 1, 3515: 1, 3511: 1, 3510: 1, 3505: 1, 35
02: 1, 3501: 1, 3500: 1, 3496: 1, 3494: 1, 3491: 1, 3489: 1, 3488: 1, 3485:
1, 3482: 1, 3479: 1, 3475: 1, 3462: 1, 3452: 1, 3450: 1, 3449: 1, 3447: 1, 34
46: 1, 3439: 1, 3428: 1, 3425: 1, 3424: 1, 3421: 1, 3420: 1, 3415: 1, 3413:
1, 3410: 1, 3408: 1, 3407: 1, 3406: 1, 3397: 1, 3376: 1, 3373: 1, 3370: 1, 33
54: 1, 3351: 1, 3342: 1, 3337: 1, 3335: 1, 3331: 1, 3325: 1, 3316: 1, 3313:
1, 3305: 1, 3299: 1, 3297: 1, 3295: 1, 3289: 1, 3284: 1, 3282: 1, 3281: 1, 32
79: 1, 3278: 1, 3277: 1, 3270: 1, 3269: 1, 3265: 1, 3259: 1, 3256: 1, 3255:
1, 3253: 1, 3252: 1, 3242: 1, 3238: 1, 3236: 1, 3234: 1, 3232: 1, 3231: 1, 32
30: 1, 3225: 1, 3222: 1, 3221: 1, 3215: 1, 3206: 1, 3205: 1, 3204: 1, 3203:
1, 3199: 1, 3198: 1, 3194: 1, 3193: 1, 3191: 1, 3183: 1, 3174: 1, 3158: 1, 31
49: 1, 3144: 1, 3139: 1, 3137: 1, 3135: 1, 3134: 1, 3133: 1, 3127: 1, 3123:
1, 3122: 1, 3096: 1, 3084: 1, 3078: 1, 3077: 1, 3070: 1, 3067: 1, 3064: 1, 30
63: 1, 3062: 1, 3059: 1, 3058: 1, 3049: 1, 3048: 1, 3044: 1, 3043: 1, 3038:
1, 3030: 1, 3027: 1, 3026: 1, 3025: 1, 3024: 1, 3015: 1, 3011: 1, 2981: 1, 29
69: 1, 2968: 1, 2965: 1, 2964: 1, 2963: 1, 2962: 1, 2956: 1, 2953: 1, 2951:
1, 2949: 1, 2947: 1, 2944: 1, 2938: 1, 2932: 1, 2920: 1, 2918: 1, 2916: 1, 29
09: 1, 2906: 1, 2901: 1, 2899: 1, 2897: 1, 2891: 1, 2889: 1, 2885: 1, 2884:
1, 2880: 1, 2876: 1, 2873: 1, 2870: 1, 2867: 1, 2858: 1, 2850: 1, 2849: 1, 28
44: 1, 2841: 1, 2827: 1, 2824: 1, 2821: 1, 2820: 1, 2815: 1, 2814: 1, 2811:
1, 2805: 1, 2804: 1, 2799: 1, 2797: 1, 2791: 1, 2789: 1, 2788: 1, 2769: 1, 27
64: 1, 2763: 1, 2762: 1, 2758: 1, 2756: 1, 2754: 1, 2750: 1, 2747: 1, 2746:
1, 2744: 1, 2742: 1, 2740: 1, 2729: 1, 2728: 1, 2722: 1, 2716: 1, 2708: 1, 27
06: 1, 2702: 1, 2700: 1, 2693: 1, 2688: 1, 2685: 1, 2683: 1, 2682: 1, 2681:
1, 2680: 1, 2678: 1, 2675: 1, 2674: 1, 2672: 1, 2663: 1, 2660: 1, 2656: 1, 26
54: 1, 2650: 1, 2647: 1, 2643: 1, 2638: 1, 2637: 1, 2634: 1, 2632: 1, 2630:
1, 2626: 1, 2625: 1, 2621: 1, 2619: 1, 2608: 1, 2592: 1, 2591: 1, 2590: 1, 25
88: 1, 2581: 1, 2580: 1, 2577: 1, 2575: 1, 2569: 1, 2567: 1, 2561: 1, 2560:
1, 2559: 1, 2556: 1, 2550: 1, 2548: 1, 2544: 1, 2543: 1, 2538: 1, 2537: 1, 25
36: 1, 2534: 1, 2532: 1, 2530: 1, 2529: 1, 2525: 1, 2511: 1, 2501: 1, 2491:
1, 2488: 1, 2487: 1, 2484: 1, 2482: 1, 2479: 1, 2477: 1, 2476: 1, 2474: 1, 24
73: 1, 2463: 1, 2461: 1, 2457: 1, 2453: 1, 2452: 1, 2450: 1, 2445: 1, 2443:
1, 2431: 1, 2428: 1, 2426: 1, 2425: 1, 2423: 1, 2422: 1, 2419: 1, 2417: 1, 24
16: 1, 2412: 1, 2411: 1, 2410: 1, 2408: 1, 2404: 1, 2401: 1, 2398: 1, 2393:
1, 2386: 1, 2383: 1, 2382: 1, 2379: 1, 2378: 1, 2373: 1, 2365: 1, 2361: 1, 23

60: 1, 2359: 1, 2356: 1, 2350: 1, 2343: 1, 2342: 1, 2341: 1, 2335: 1, 2331: 1, 2330: 1, 2322: 1, 2321: 1, 2315: 1, 2311: 1, 2308: 1, 2307: 1, 2305: 1, 2304: 1, 2302: 1, 2298: 1, 2294: 1, 2291: 1, 2287: 1, 2279: 1, 2276: 1, 2275: 1, 2274: 1, 2268: 1, 2264: 1, 2263: 1, 2261: 1, 2254: 1, 2248: 1, 2235: 1, 2234: 1, 2233: 1, 2232: 1, 2231: 1, 2230: 1, 2225: 1, 2224: 1, 2223: 1, 2216: 1, 2214: 1, 2206: 1, 2204: 1, 2197: 1, 2195: 1, 2191: 1, 2185: 1, 2183: 1, 2182: 1, 2177: 1, 2176: 1, 2174: 1, 2173: 1, 2169: 1, 2168: 1, 2166: 1, 2165: 1, 2163: 1, 2156: 1, 2151: 1, 2150: 1, 2147: 1, 2143: 1, 2141: 1, 2140: 1, 2137: 1, 2135: 1, 2134: 1, 2132: 1, 2129: 1, 2126: 1, 2124: 1, 2120: 1, 2116: 1, 2115: 1, 2108: 1, 2102: 1, 2099: 1, 2087: 1, 2081: 1, 2080: 1, 2077: 1, 2072: 1, 2071: 1, 2070: 1, 2069: 1, 2066: 1, 2064: 1, 2060: 1, 2055: 1, 2053: 1, 2051: 1, 2050: 1, 2048: 1, 2046: 1, 2045: 1, 2044: 1, 2041: 1, 2039: 1, 2038: 1, 2037: 1, 2033: 1, 2025: 1, 2024: 1, 2022: 1, 2020: 1, 2015: 1, 2014: 1, 2010: 1, 2006: 1, 1991: 1, 1989: 1, 1982: 1, 1980: 1, 1978: 1, 1975: 1, 1968: 1, 1967: 1, 1965: 1, 1964: 1, 1957: 1, 1955: 1, 1954: 1, 1950: 1, 1947: 1, 1943: 1, 1936: 1, 1924: 1, 1923: 1, 1921: 1, 1920: 1, 1917: 1, 1912: 1, 1911: 1, 1908: 1, 1907: 1, 1906: 1, 1905: 1, 1901: 1, 1899: 1, 1898: 1, 1894: 1, 1893: 1, 1892: 1, 1891: 1, 1890: 1, 1887: 1, 1884: 1, 1883: 1, 1881: 1, 1880: 1, 1879: 1, 1877: 1, 1871: 1, 1869: 1, 1865: 1, 1863: 1, 1858: 1, 1857: 1, 1856: 1, 1855: 1, 1854: 1, 1851: 1, 1842: 1, 1841: 1, 1839: 1, 1838: 1, 1834: 1, 1833: 1, 1831: 1, 1830: 1, 1829: 1, 1826: 1, 1822: 1, 1821: 1, 1820: 1, 1814: 1, 1811: 1, 1810: 1, 1809: 1, 1808: 1, 1805: 1, 1804: 1, 1803: 1, 1802: 1, 1800: 1, 1799: 1, 1798: 1, 1796: 1, 1794: 1, 1792: 1, 1790: 1, 1786: 1, 1785: 1, 1783: 1, 1782: 1, 1777: 1, 1776: 1, 1774: 1, 1771: 1, 1770: 1, 1768: 1, 1767: 1, 1766: 1, 1765: 1, 1762: 1, 1761: 1, 1759: 1, 1756: 1, 1755: 1, 1754: 1, 1752: 1, 1749: 1, 1747: 1, 1746: 1, 1745: 1, 1742: 1, 1738: 1, 1736: 1, 1733: 1, 1732: 1, 1729: 1, 1723: 1, 1722: 1, 1711: 1, 1704: 1, 1703: 1, 1702: 1, 1700: 1, 1698: 1, 1696: 1, 1695: 1, 1694: 1, 1688: 1, 1686: 1, 1684: 1, 1680: 1, 1678: 1, 1674: 1, 1671: 1, 1669: 1, 1665: 1, 1664: 1, 1663: 1, 1660: 1, 1654: 1, 1643: 1, 1642: 1, 1639: 1, 1637: 1, 1635: 1, 1632: 1, 1629: 1, 1627: 1, 1625: 1, 1623: 1, 1621: 1, 1619: 1, 1616: 1, 1613: 1, 1612: 1, 1608: 1, 1607: 1, 1605: 1, 1602: 1, 1599: 1, 1596: 1, 1594: 1, 1593: 1, 1590: 1, 1589: 1, 1588: 1, 1583: 1, 1578: 1, 1572: 1, 1571: 1, 1569: 1, 1566: 1, 1564: 1, 1559: 1, 1557: 1, 1556: 1, 1554: 1, 1550: 1, 1549: 1, 1542: 1, 1540: 1, 1537: 1, 1534: 1, 1532: 1, 1530: 1, 1522: 1, 1520: 1, 1519: 1, 1513: 1, 1508: 1, 1507: 1, 1506: 1, 1503: 1, 1501: 1, 1500: 1, 1499: 1, 1498: 1, 1497: 1, 1491: 1, 1490: 1, 1488: 1, 1487: 1, 1486: 1, 1485: 1, 1483: 1, 1479: 1, 1476: 1, 1474: 1, 1472: 1, 1470: 1, 1469: 1, 1468: 1, 1463: 1, 1462: 1, 1456: 1, 1444: 1, 1442: 1, 1439: 1, 1434: 1, 1432: 1, 1429: 1, 1424: 1, 1423: 1, 1422: 1, 1421: 1, 1420: 1, 1418: 1, 1411: 1, 1405: 1, 1403: 1, 1393: 1, 1392: 1, 1390: 1, 1381: 1, 1376: 1, 1365: 1, 1363: 1, 1361: 1, 1358: 1, 1357: 1, 1356: 1, 1355: 1, 1352: 1, 1349: 1, 1347: 1, 1345: 1, 1332: 1, 1331: 1, 1327: 1, 1326: 1, 1325: 1, 1323: 1, 1322: 1, 1319: 1, 1318: 1, 1315: 1, 1314: 1, 1309: 1, 1306: 1, 1304: 1, 1302: 1, 1297: 1, 1296: 1, 1293: 1, 1290: 1, 1287: 1, 1282: 1, 1279: 1, 1274: 1, 1270: 1, 1267: 1, 1265: 1, 1261: 1, 1260: 1, 1257: 1, 1256: 1, 1253: 1, 1251: 1, 1250: 1, 1247: 1, 1246: 1, 1241: 1, 1232: 1, 1229: 1, 1228: 1, 1227: 1, 1218: 1, 1216: 1, 1214: 1, 1209: 1, 1206: 1, 1194: 1, 1191: 1, 1190: 1, 1188: 1, 1183: 1, 1176: 1, 1167: 1, 1165: 1, 1161: 1, 1158: 1, 1155: 1, 1154: 1, 1147: 1, 1146: 1, 1144: 1, 1140: 1, 1133: 1, 1125: 1, 1117: 1, 1115: 1, 1113: 1, 1106: 1, 1105: 1, 1097: 1, 1096: 1, 1094: 1, 1090: 1, 1087: 1, 1086: 1, 1085: 1, 1083: 1, 1081: 1, 1080: 1, 1078: 1, 1070: 1, 1069: 1, 1067: 1, 1065: 1, 1064: 1, 1058: 1, 1050: 1, 1049: 1, 1046: 1, 1038: 1, 1037: 1, 1036: 1, 1035: 1, 1033: 1, 1032: 1, 1031: 1, 1030: 1, 1027: 1, 1024: 1, 1016: 1, 1013: 1, 1012: 1, 991: 1, 989: 1, 982: 1, 979: 1, 978: 1, 974: 1, 973: 1, 972: 1, 971: 1, 968: 1, 967: 1, 956: 1, 953: 1, 951: 1, 950: 1, 939: 1, 930: 1, 925: 1, 923: 1, 922: 1, 920: 1, 908: 1, 905: 1, 901: 1, 884: 1, 879: 1, 878: 1, 866: 1, 864: 1, 854: 1, 852: 1, 842: 1, 837: 1

1, 830: 1, 829: 1, 822: 1, 812: 1, 810: 1, 787: 1, 785: 1, 783: 1, 782: 1, 745: 1, 743: 1, 686: 1, 679: 1})

```
In [36]: # Train a Logistic regression+Calibration model using text features which are
# on-hot encoded
alpha = [10 ** x for x in range(-5, 1)]

# read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generate
# d/sklearn.Linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_i
# ntercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_
# rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochast
# ic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link:
#-----
```

cv_log_error_array=[]
for i in alpha:
 clf = SGDClassifier(alpha=i, penalty='l2', loss='log', random_state=42)
 clf.fit(train_text_feature_onehotCoding, y_train)

 sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
 sig_clf.fit(train_text_feature_onehotCoding, y_train)
 predict_y = sig_clf.predict_proba(cv_text_feature_onehotCoding)
 cv_log_error_array.append(log_loss(y_cv, predict_y, labels=clf.classes_, e
ps=1e-15))
 print('For values of alpha = ', i, "The log loss is:", log_loss(y_cv, predi
ct_y, labels=clf.classes_, eps=1e-15))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
 ax.annotate((alpha[i],np.round(txt,3)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

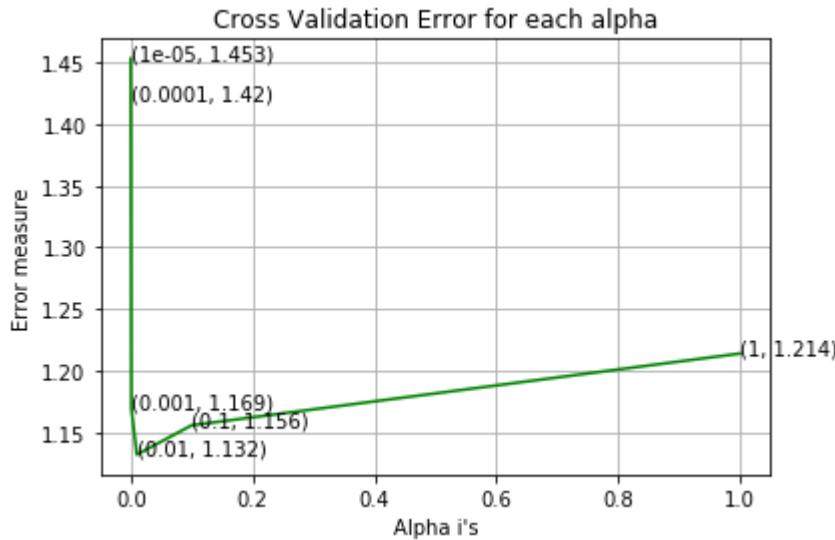
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_
state=42)
clf.fit(train_text_feature_onehotCoding, y_train)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_text_feature_onehotCoding, y_train)

predict_y = sig_clf.predict_proba(train_text_feature_onehotCoding)

```

print('For values of best alpha = ', alpha[best_alpha], "The train log loss is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_text_feature_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_text_feature_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
For values of alpha = 1e-05 The log loss is: 1.4532819300768252
For values of alpha = 0.0001 The log loss is: 1.4202733206787634
For values of alpha = 0.001 The log loss is: 1.169278090692579
For values of alpha = 0.01 The log loss is: 1.1317098499669203
For values of alpha = 0.1 The log loss is: 1.155545558427755
For values of alpha = 1 The log loss is: 1.2138125028557656

```



```

For values of best alpha = 0.01 The train log loss is: 0.7276377766441388
For values of best alpha = 0.01 The cross validation log loss is: 1.1317098499669203
For values of best alpha = 0.01 The test log loss is: 1.1669888052217867

```

Q. Is the Text feature stable across all the data sets (Test, Train, Cross validation)?

Ans. Yes, it seems like!

```

In [37]: def get_intersec_text(df):
    df_text_vec = CountVectorizer(min_df=3)
    df_text_fea = df_text_vec.fit_transform(df['TEXT'])
    df_text_features = df_text_vec.get_feature_names()

    df_text_fea_counts = df_text_fea.sum(axis=0).A1
    df_text_fea_dict = dict(zip(list(df_text_features),df_text_fea_counts))
    len1 = len(set(df_text_features))
    len2 = len(set(train_text_features) & set(df_text_features))
    return len1,len2

```

```
In [38]: len1,len2 = get_intersec_text(test_df)
print(np.round((len2/len1)*100, 3), "% of word of test data appeared in train
      data")
len1,len2 = get_intersec_text(cv_df)
print(np.round((len2/len1)*100, 3), "% of word of Cross Validation appeared in
      train data")
```

97.065 % of word of test data appeared in train data
 98.643 % of word of Cross Validation appeared in train data

```
In [1]: import dill
# dill.dump_session('notebook_envassign3.db')
# dill.load_session('notebook_envassign3.db')
```

In [2]: #Data preparation for ML models.

```
#Misc. functionns for ML models

def predict_and_plot_confusion_matrix(train_x, train_y,test_x, test_y, clf):
    clf.fit(train_x, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x, train_y)
    pred_y = sig_clf.predict(test_x)

    # for calculating Log_loss we will provide the array of probabilities belongs to each class
    print("Log loss :",log_loss(test_y, sig_clf.predict_proba(test_x)))
    # calculating the number of data points that are misclassified
    print("Number of mis-classified points :", np.count_nonzero((pred_y- test_y))/test_y.shape[0])
    plot_confusion_matrix(test_y, pred_y)
```

```
In [3]: def report_log_loss(train_x, train_y, test_x, test_y, clf):
    clf.fit(train_x, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x, train_y)
    sig_clf_probs = sig_clf.predict_proba(test_x)
    return log_loss(test_y, sig_clf_probs, eps=1e-15)
```

```
In [4]: # this function will be used just for naive bayes
# for the given indices, we will print the name of the features
# and we will check whether the feature present in the test point text or not
def get_impfeature_names(indices, text, gene, var, no_features):
    gene_count_vec = CountVectorizer()
    var_count_vec = CountVectorizer()
    text_count_vec = CountVectorizer(min_df=3,ngram_range=(1,2))

    gene_vec = gene_count_vec.fit(train_df['Gene'])
    var_vec = var_count_vec.fit(train_df['Variation'])
    text_vec = text_count_vec.fit(train_df['TEXT'])

    fea1_len = len(gene_vec.get_feature_names())
    fea2_len = len(var_count_vec.get_feature_names())

    word_present = 0
    for i,v in enumerate(indices):
        if (v < fea1_len):
            word = gene_vec.get_feature_names()[v]
            yes_no = True if word == gene else False
            if yes_no:
                word_present += 1
                print(i, "Gene feature [{}] present in test data point [{}].format(word,yes_no))
            elif (v < fea1_len+fea2_len):
                word = var_vec.get_feature_names()[v-(fea1_len)]
                yes_no = True if word == var else False
                if yes_no:
                    word_present += 1
                    print(i, "variation feature [{}] present in test data point [{}].format(word,yes_no))
            else:
                word = text_vec.get_feature_names()[v-(fea1_len+fea2_len)]
                yes_no = True if word in text.split() else False
                if yes_no:
                    word_present += 1
                    print(i, "Text feature [{}] present in test data point [{}].format(word,yes_no))

    print("Out of the top ",no_features," features ", word_present, "are present in query point")
```

Stacking the three types of features

In [5]: # merging gene, variance and text features

```
# building train, test and cross validation data sets
# a = [[1, 2],
#       [3, 4]]
# b = [[4, 5],
#       [6, 7]]
# hstack(a, b) = [[1, 2, 4, 5],
#                  [3, 4, 6, 7]]

train_gene_var_onehotCoding = hstack((train_gene_feature_onehotCoding,train_variation_feature_onehotCoding))
test_gene_var_onehotCoding = hstack((test_gene_feature_onehotCoding,test_variation_feature_onehotCoding))
cv_gene_var_onehotCoding = hstack((cv_gene_feature_onehotCoding, cv_variation_feature_onehotCoding))

train_x_onehotCoding = hstack((train_gene_var_onehotCoding, train_text_feature_onehotCoding)).tocsr()

test_x_onehotCoding = hstack((test_gene_var_onehotCoding, test_text_feature_on
ehotCoding)).tocsr()

cv_x_onehotCoding = hstack((cv_gene_var_onehotCoding, cv_text_feature_onehotCo
ding)).tocsr()

train_gene_var_responseCoding = np.hstack((train_gene_feature_responseCoding,t
rain_variation_feature_responseCoding))
test_gene_var_responseCoding = np.hstack((test_gene_feature_responseCoding,tes
t_variation_feature_responseCoding))
cv_gene_var_responseCoding = np.hstack((cv_gene_feature_responseCoding, cv_vari
ation_feature_responseCoding))

train_x_responseCoding = np.hstack((train_gene_var_responseCoding, train_text_
feature_responseCoding))
test_x_responseCoding = np.hstack((test_gene_var_responseCoding, test_text_fea
ture_responseCoding))
cv_x_responseCoding = np.hstack((cv_gene_var_responseCoding, cv_text_feature_r
esponseCoding))

train_y = np.array(list(train_df['Class']))
test_y = np.array(list(test_df['Class']))
cv_y = np.array(list(cv_df['Class']))
```

```
In [6]: print("One hot encoding features :")
print("(number of data points * number of features) in train data = ", train_x_
_onehotCoding.shape)
print("(number of data points * number of features) in test data = ", test_x_o
nehotCoding.shape)
print("(number of data points * number of features) in cross validation data
 =", cv_x_onehotCoding.shape)
```

One hot encoding features :
(number of data points * number of features) in train data = (2124, 780961)
(number of data points * number of features) in test data = (665, 780961)
(number of data points * number of features) in cross validation data = (532,
780961)

```
In [7]: print(" Response encoding features :")
print("(number of data points * number of features) in train data = ", train_x_
_responseCoding.shape)
print("(number of data points * number of features) in test data = ", test_x_r
esponseCoding.shape)
print("(number of data points * number of features) in cross validation data
 =", cv_x_responseCoding.shape)
```

Response encoding features :
(number of data points * number of features) in train data = (2124, 27)
(number of data points * number of features) in test data = (665, 27)
(number of data points * number of features) in cross validation data = (532,
27)

In []:

4. Machine Learning Models

In []:

4.3. Logistic Regression

4.3.1. With Class balancing

4.3.1.1. Hyper parameter tuning

```
In [10]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight])      Fit the calibrated model
# get_params([deep])      Get parameters for this estimator.
# predict(X)      Predict the target of new samples.
# predict_proba(X)    Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10 ** x for x in range(-6, 3)]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = SGDClassifier(class_weight='balanced', alpha=i, penalty='L2', loss='log', random_state=42)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
        # to avoid rounding error while multiplying probabilites we use log-probability estimates
    print("Log Loss :",log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
```

```
    ax.annotate((alpha[i],str(txt)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

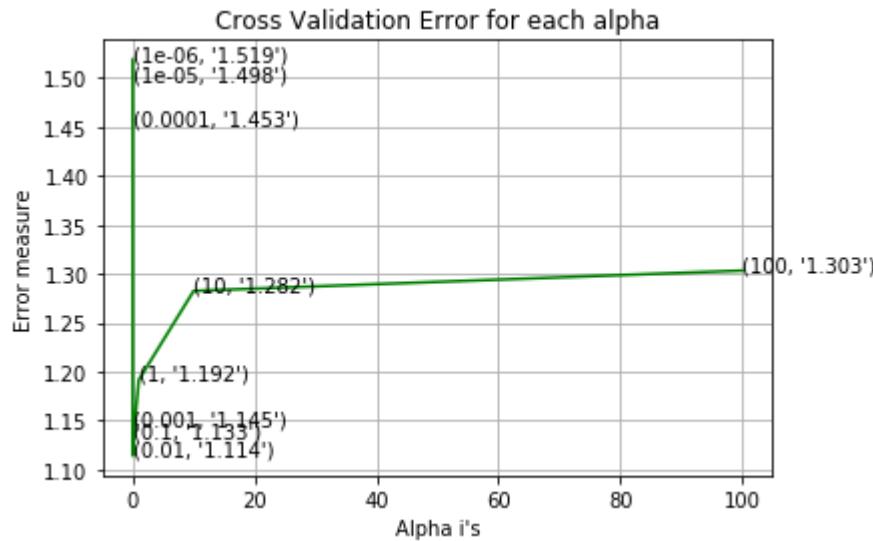
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty=
'12', loss='log', random_state=42)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss i
s:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation
log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss i
s:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 1e-06
Log Loss : 1.518667464100214
for alpha = 1e-05
Log Loss : 1.4981216275856408
for alpha = 0.0001
Log Loss : 1.453148235596527
for alpha = 0.001
Log Loss : 1.1445450670837332
for alpha = 0.01
Log Loss : 1.114187623675461
for alpha = 0.1
Log Loss : 1.132690923178632
for alpha = 1
Log Loss : 1.1916442482956104
for alpha = 10
Log Loss : 1.2824656367429352
for alpha = 100
Log Loss : 1.3031686952735435

```



For values of best alpha = 0.01 The train log loss is: 0.7138899640045442
 For values of best alpha = 0.01 The cross validation log loss is: 1.114187623675461
 For values of best alpha = 0.01 The test log loss is: 1.153420009891365

4.3.1.2. Testing the model with best hyper paramters

```
In [11]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

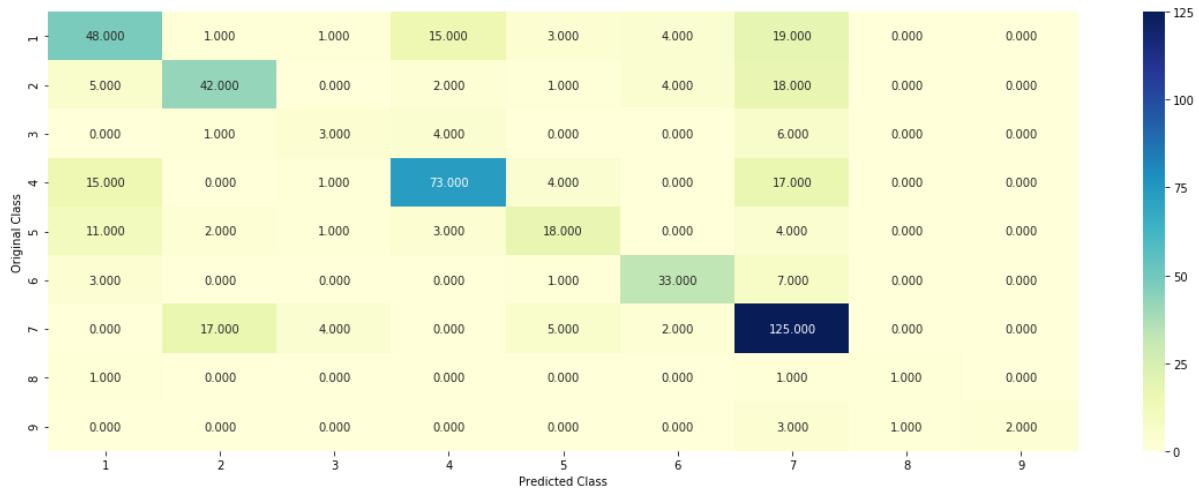
# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCoding, cv_y, clf)
```

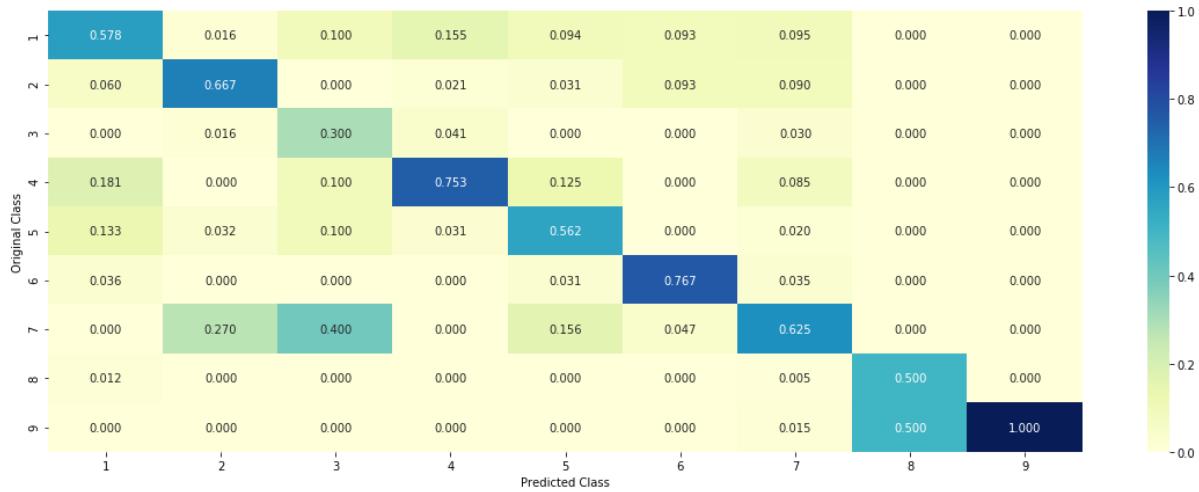
Log loss : 1.114187623675461

Number of mis-classified points : 0.35150375939849626

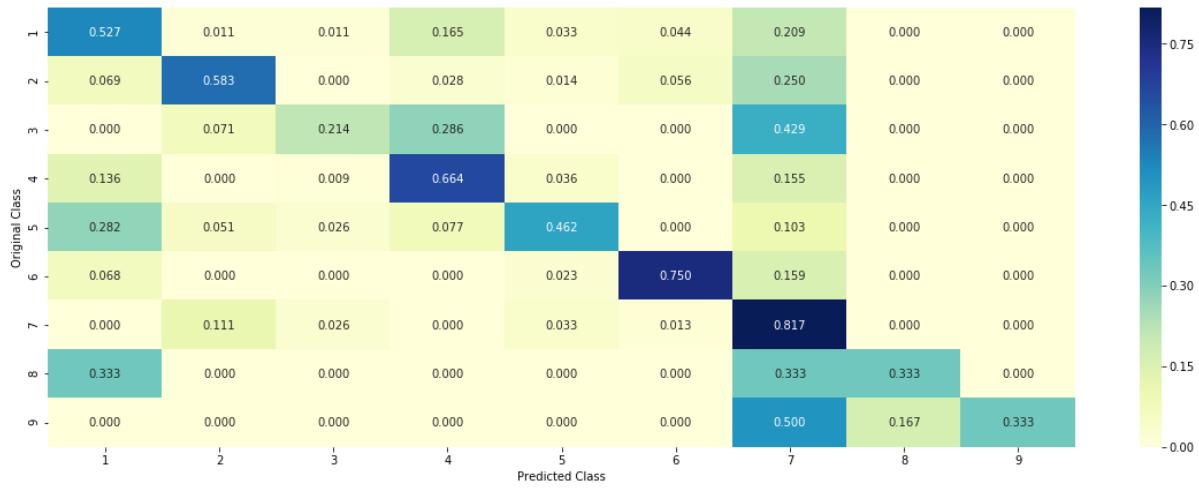
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.3.1.3. Feature Importance

```
In [12]: def get_imp_feature_names(text, indices, removed_ind = []):
    word_present = 0
    tabulte_list = []
    incresingorder_ind = 0
    for i in indices:
        if i < train_gene_feature_onehotCoding.shape[1]:
            tabulte_list.append([incresingorder_ind, "Gene", "Yes"])
        elif i < 18:
            tabulte_list.append([incresingorder_ind, "Variation", "Yes"])
        if ((i > 17) & (i not in removed_ind)):
            word = train_text_features[i]
            yes_no = True if word in text.split() else False
            if yes_no:
                word_present += 1
            tabulte_list.append([incresingorder_ind, train_text_features[i], yes_no])
        incresingorder_ind += 1
    print(word_present, "most important features are present in our query point")
    print("-"*50)
    print("The features that are most important of the ",predicted_cls[0]," class:")
    print (tabulate(tabulte_list, headers=["Index", 'Feature name', 'Present or Not']))
```

4.3.1.3.1. Correctly Classified point

```
In [13]: # from tabulate import tabulate
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
clf.fit(train_x_onehotCoding,train_y)
test_point_index = 1
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imffeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)
```

```
Predicted Class : 7
Predicted Class Probabilities: [[4.700e-03 3.705e-01 5.000e-04 2.610e-02 2.00
0e-03 1.000e-04 5.842e-01
8.800e-03 3.100e-03]]
Actual Class : 7
-----
74 Text feature [function] present in test data point [True]
89 Text feature [affected] present in test data point [True]
Out of the top 100 features 2 are present in query point
```

4.3.1.3.2. Incorrectly Classified point

```
In [14]: test_point_index = 361
no_feature = 100
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)

Predicted Class : 5
Predicted Class Probabilities: [[0.0827 0.1132 0.0145 0.0673 0.4725 0.0362 0.
1983 0.007 0.0084]]
Actual Class : 6
-----
Out of the top 100 features 0 are present in query point
```

4.3.2. Without Class balancing

4.3.2.1. Hyper parameter tuning

```
In [15]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight])  Fit the calibrated model
# get_params([deep])  Get parameters for this estimator.
# predict(X)  Predict the target of new samples.
# predict_proba(X)  Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10 ** x for x in range(-6, 1)]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = SGDClassifier(alpha=i, penalty='l2', loss='log', random_state=42)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
    print("Log Loss :", log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array, c='g')
for i, txt in enumerate(np.round(cv_log_error_array, 3)):
    ax.annotate((alpha[i], str(txt)), (alpha[i], cv_log_error_array[i]))
plt.grid()
```

```
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

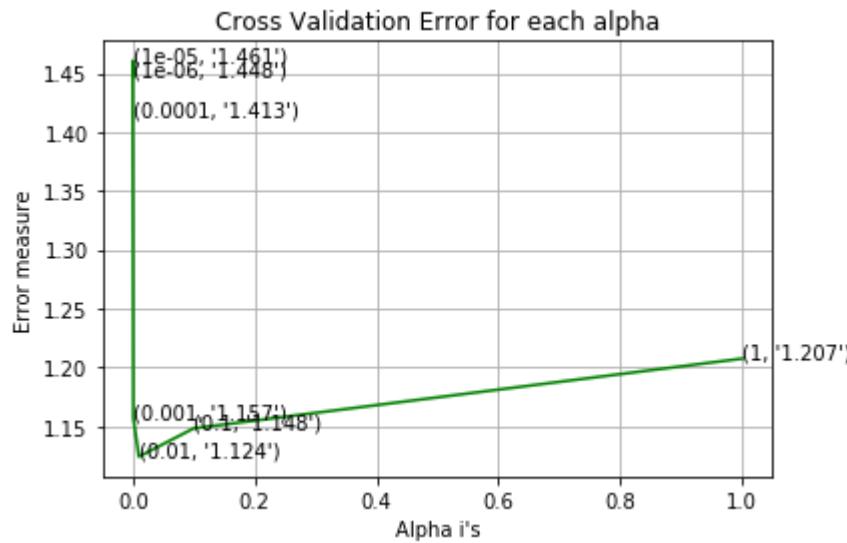
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_
state=42)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss i
s:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation
log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss i
s:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 1e-06
Log Loss : 1.4478812536450172
for alpha = 1e-05
Log Loss : 1.4607028515361526
for alpha = 0.0001
Log Loss : 1.413281343594843
for alpha = 0.001
Log Loss : 1.1567469114390878
for alpha = 0.01
Log Loss : 1.1240181919494745
for alpha = 0.1
Log Loss : 1.1480509388484441
for alpha = 1
Log Loss : 1.207367148982691

```



For values of best alpha = 0.01 The train log loss is: 0.702603375246738
 For values of best alpha = 0.01 The cross validation log loss is: 1.1240181919494745
 For values of best alpha = 0.01 The test log loss is: 1.1485067320074351

4.3.2.2. Testing model with best hyper parameters

```
In [16]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)    Predict class labels for samples in X.

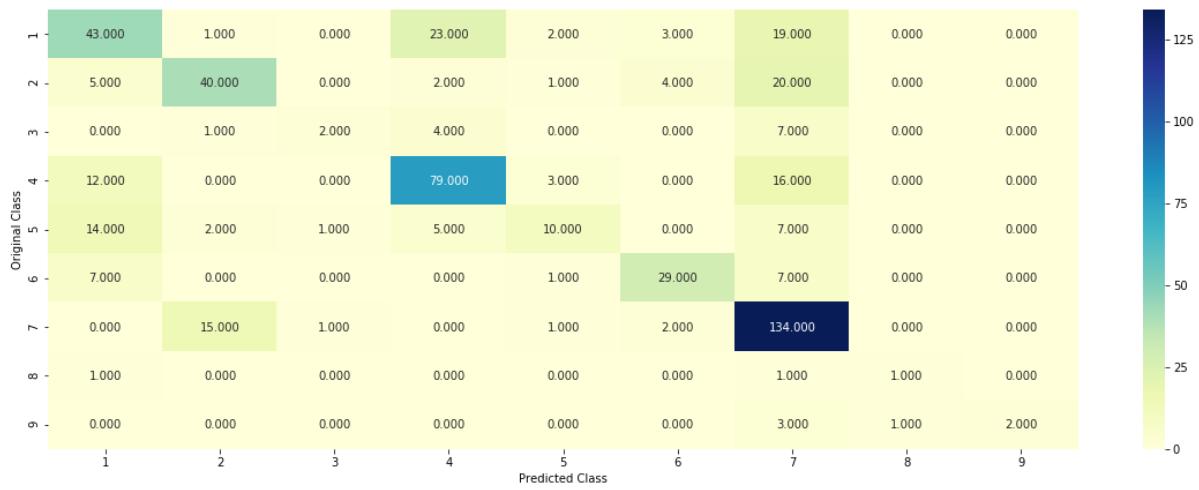
#-----
# video link:
#-----
```

```
clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCoding, cv_y, clf)
```

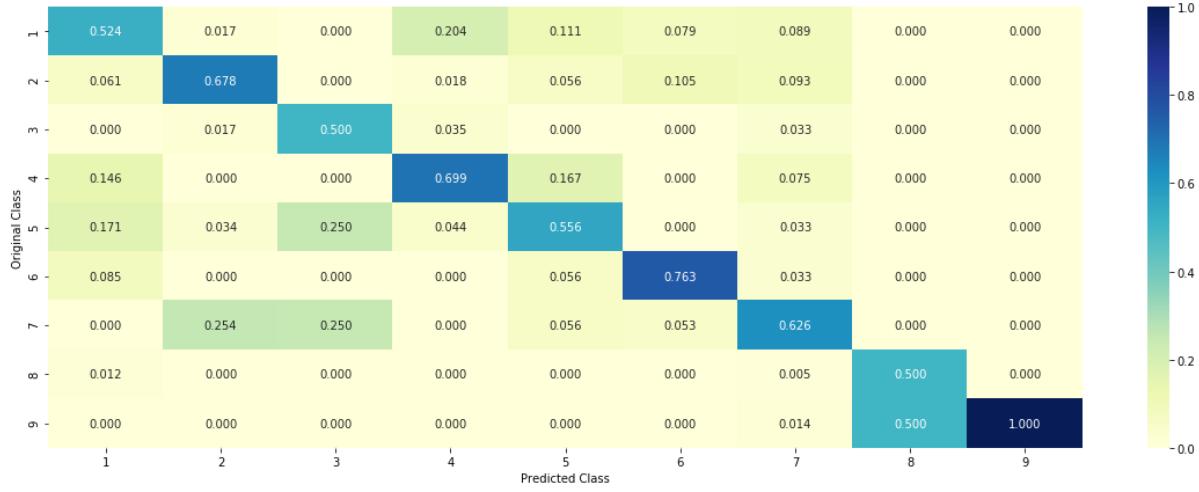
Log loss : 1.1240181919494745

Number of mis-classified points : 0.3609022556390977

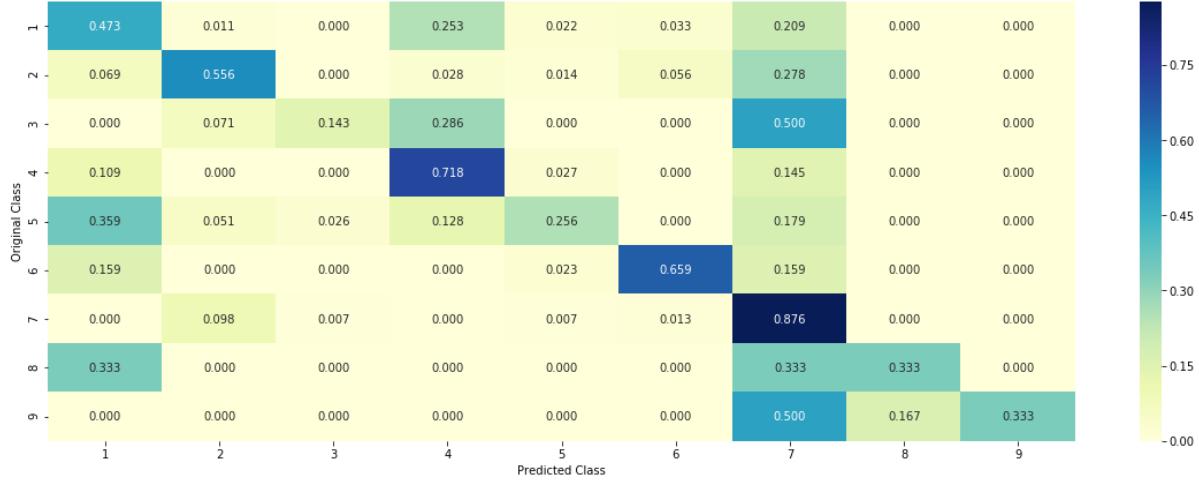
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.3.2.3. Feature Importance, Correctly Classified point

```
In [18]: clf = SGDClassifier(alpha=alpha[best_alpha], penalty='l2', loss='log', random_
state=42)
clf.fit(train_x_onehotCoding,train_y)
test_point_index = 1
no_feature = 250
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

Predicted Class : 7
Predicted Class Probabilities: [[3.400e-03 2.733e-01 2.000e-04 2.640e-02 1.20
0e-03 1.000e-04 6.923e-01
3.100e-03 0.000e+00]]
Actual Class : 7

Out of the top 250 features 0 are present in query point

4.3.2.4. Feature Importance, Inorrectly Classified point

```
In [19]: test_point_index = 125
no_feature = 250
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imptfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

Predicted Class : 7
Predicted Class Probabilities: [[0.1268 0.1334 0.0208 0.1314 0.0501 0.0365 0.
4898 0.0073 0.004]]
Actual Class : 2

236 Text feature [individuals] present in test data point [True]
Out of the top 250 features 1 are present in query point

In []:

In []:

```
In [1]: # https://machinelearningmastery.com/how-to-fix-futurewarning-messages-in-scikit-learn/
from warnings import simplefilter
# ignore all future warnings
simplefilter(action='ignore', category=FutureWarning)

import pandas as pd
import matplotlib.pyplot as plt
import re
import time
import warnings
import numpy as np
from nltk.corpus import stopwords
from sklearn.decomposition import TruncatedSVD
from sklearn.preprocessing import normalize
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.manifold import TSNE
import seaborn as sns
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix
from sklearn.metrics.classification import accuracy_score, log_loss
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import SGDClassifier
from imblearn.over_sampling import SMOTE
from collections import Counter
from scipy.sparse import hstack
from sklearn.multiclass import OneVsRestClassifier
from sklearn.svm import SVC
from sklearn.model_selection import StratifiedKFold
from collections import Counter, defaultdict
from sklearn.calibration import CalibratedClassifierCV
from sklearn.naive_bayes import MultinomialNB
from sklearn.naive_bayes import GaussianNB
from sklearn.model_selection import train_test_split
from sklearn.model_selection import GridSearchCV
import math
from sklearn.metrics import normalized_mutual_info_score
from sklearn.ensemble import RandomForestClassifier
warnings.filterwarnings("ignore")

from mlxtend.classifier import StackingClassifier

from sklearn import model_selection
from sklearn.linear_model import LogisticRegression
```

3.1. Reading Data

3.1.1. Reading Gene and Variation Data

```
In [2]: data = pd.read_csv('training/training_variants')
print('Number of data points : ', data.shape[0])
print('Number of features : ', data.shape[1])
print('Features : ', data.columns.values)
data.head()
```

Number of data points : 3321
 Number of features : 4
 Features : ['ID' 'Gene' 'Variation' 'Class']

Out[2]:

	ID	Gene	Variation	Class
0	0	FAM58A	Truncating Mutations	1
1	1	CBL	W802*	2
2	2	CBL	Q249E	2
3	3	CBL	N454D	3
4	4	CBL	L399V	4

training/training_variants is a comma separated file containing the description of the genetic mutations used for training.

Fields are

- **ID** : the id of the row used to link the mutation to the clinical evidence
- **Gene** : the gene where this genetic mutation is located
- **Variation** : the aminoacid change for this mutations
- **Class** : 1-9 the class this genetic mutation has been classified on

3.1.2. Reading Text Data

```
In [3]: # note the separator in this file
data_text = pd.read_csv("training/training_text", sep="\|\|", engine="python", names=["ID", "TEXT"], skiprows=1)
print('Number of data points : ', data_text.shape[0])
print('Number of features : ', data_text.shape[1])
print('Features : ', data_text.columns.values)
data_text.head()
```

Number of data points : 3321
 Number of features : 2
 Features : ['ID' 'TEXT']

Out[3]:

	ID	TEXT
0	0	Cyclin-dependent kinases (CDKs) regulate a var...
1	1	Abstract Background Non-small cell lung canc...
2	2	Abstract Background Non-small cell lung canc...
3	3	Recent evidence has demonstrated that acquired...
4	4	Oncogenic mutations in the monomeric Casitas B...

3.1.3. Preprocessing of text

```
In [4]: # Loading stop words from nltk library
stop_words = set(stopwords.words('english'))

def nlp_preprocessing(total_text, index, column):
    if type(total_text) is not int:
        string = ""
        # replace every special char with space
        total_text = re.sub('[^a-zA-Z0-9\n]', ' ', total_text)
        # replace multiple spaces with single space
        total_text = re.sub('\s+', ' ', total_text)
        # converting all the chars into lower-case.
        total_text = total_text.lower()

        for word in total_text.split():
            # if the word is a not a stop word then retain that word from the data
            if not word in stop_words:
                string += word + " "

        data_text[column][index] = string
```

```
In [5]: #text processing stage.
start_time = time.clock()
for index, row in data_text.iterrows():
    if type(row['TEXT']) is str:
        nlp_preprocessing(row['TEXT'], index, 'TEXT')
    else:
        print("there is no text description for id:",index)
print('Time took for preprocessing the text :',time.clock() - start_time, "seconds")
```

there is no text description for id: 1109
 there is no text description for id: 1277
 there is no text description for id: 1407
 there is no text description for id: 1639
 there is no text description for id: 2755
 Time took for preprocessing the text : 154.838072291 seconds

```
In [6]: #merging both gene_variations and text data based on ID
result = pd.merge(data, data_text, on='ID', how='left')
result.head()
```

Out[6]:

	ID	Gene	Variation	Class	TEXT
0	0	FAM58A	Truncating Mutations	1	cyclin dependent kinases cdks regulate variety...
1	1	CBL	W802*	2	abstract background non small cell lung cancer...
2	2	CBL	Q249E	2	abstract background non small cell lung cancer...
3	3	CBL	N454D	3	recent evidence demonstrated acquired uniparen...
4	4	CBL	L399V	4	oncogenic mutations monomeric casitas b lineage...

```
In [7]: result[result.isnull().any(axis=1)]
```

Out[7]:

	ID	Gene	Variation	Class	TEXT
1109	1109	FANCA	S1088F	1	NaN
1277	1277	ARID5B	Truncating Mutations	1	NaN
1407	1407	FGFR3	K508M	6	NaN
1639	1639	FLT1	Amplification	6	NaN
2755	2755	BRAF	G596C	7	NaN

```
In [8]: result.loc[result['TEXT'].isnull(),'TEXT'] = result['Gene'] + ' '+result['Variation']
```

```
In [9]: result[result['ID']==1109]
```

Out[9]:

	ID	Gene	Variation	Class	TEXT
1109	1109	FANCA	S1088F	1	FANCA S1088F

3.1.4. Test, Train and Cross Validation Split

3.1.4.1. Splitting data into train, test and cross validation (64:20:16)

```
In [10]: y_true = result['Class'].values
result.Gene      = result.Gene.str.replace('\s+', '_')
result.Variation = result.Variation.str.replace('\s+', '_')

# split the data into test and train by maintaining same distribution of output variable 'y_true' [stratify=y_true]
X_train, test_df, y_train, y_test = train_test_split(result, y_true, stratify=y_true, test_size=0.2)
# split the train data into train and cross validation by maintaining same distribution of output variable 'y_train' [stratify=y_train]
train_df, cv_df, y_train, y_cv = train_test_split(X_train, y_train, stratify=y_train, test_size=0.2)
```

```
In [11]: y_train
```

```
Out[11]: array([6, 7, 4, ..., 7, 7, 5], dtype=int64)
```

We split the data into train, test and cross validation data sets, preserving the ratio of class distribution in the original data set

```
In [12]: print('Number of data points in train data:', train_df.shape[0])
print('Number of data points in test data:', test_df.shape[0])
print('Number of data points in cross validation data:', cv_df.shape[0])
```

```
Number of data points in train data: 2124
Number of data points in test data: 665
Number of data points in cross validation data: 532
```

3.1.4.2. Distribution of y_i's in Train, Test and Cross Validation datasets

```
In [13]: pd.__version__
```

```
Out[13]: '0.24.2'
```

```
In [14]: # X_train = X_train.drop(['Class'],axis=1)
# test_df = test_df.drop(['Class'],axis=1)
# cv_df = cv_df.drop(['Class'],axis=1)
```

3.2 Prediction using a 'Random' Model

In a 'Random' Model, we generate the NINE class probabilites randomly such that they sum to 1.

```
In [15]: # This function plots the confusion matrices given y_i, y_i_hat.

def plot_confusion_matrix(test_y, predict_y):
    C = confusion_matrix(test_y, predict_y)
    # C = 9,9 matrix, each cell (i,j) represents number of points of class i a
    re predicted class j

    A =(((C.T)/(C.sum(axis=1))).T)
    #divid each element of the confusion matrix with the sum of elements in th
    at column

    # C = [[1, 2],
    #      [3, 4]]
    # C.T = [[1, 3],
    #          [2, 4]]
    # C.sum(axis = 1) axis=0 corresonds to columns and axis=1 corresponds to
    rows in two dimensional array
    # C.sum(axis =1) = [[3, 7]]
    # ((C.T)/(C.sum(axis=1))) = [[1/3, 3/7]
    #                               [2/3, 4/7]]

    # ((C.T)/(C.sum(axis=1))).T = [[1/3, 2/3]
    #                               [3/7, 4/7]]
    # sum of row elements = 1

    B =(C/C.sum(axis=0))
    #divid each element of the confusion matrix with the sum of elements in th
    at row

    # C = [[1, 2],
    #      [3, 4]]
    # C.sum(axis = 0) axis=0 corresonds to columns and axis=1 corresponds to
    rows in two dimensional array
    # C.sum(axis =0) = [[4, 6]]
    # (C/C.sum(axis=0)) = [[1/4, 2/6],
    #                      [3/4, 4/6]]

    labels = [1,2,3,4,5,6,7,8,9]
    # representing A in heatmap format
    print("-"*20, "Confusion matrix", "*"-20)
    plt.figure(figsize=(20,7))
    sns.heatmap(C, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y
    ticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.show()

    print("-"*20, "Precision matrix (Column Sum=1)", "*"-20)
    plt.figure(figsize=(20,7))
    sns.heatmap(B, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y
    ticklabels=labels)
    plt.xlabel('Predicted Class')
    plt.ylabel('Original Class')
    plt.show()

    # representing B in heatmap format
    print("-"*20, "Recall matrix (Row sum=1)", "*"-20)
    plt.figure(figsize=(20,7))
```

```
    sns.heatmap(A, annot=True, cmap="YlGnBu", fmt=".3f", xticklabels=labels, y  
    ticklabels=labels)  
    plt.xlabel('Predicted Class')  
    plt.ylabel('Original Class')  
    plt.show()
```

```
In [16]: # we need to generate 9 numbers and the sum of numbers should be 1
# one solution is to generate 9 numbers and divide each of the numbers by their sum
# ref: https://stackoverflow.com/a/18662466/4084039
test_data_len = test_df.shape[0]
cv_data_len = cv_df.shape[0]

# we create a output array that has exactly same size as the CV data
cv_predicted_y = np.zeros((cv_data_len,9))
for i in range(cv_data_len):
    rand_probs = np.random.rand(1,9)
    cv_predicted_y[i] = ((rand_probs/sum(sum(rand_probs)))[0])
print("Log loss on Cross Validation Data using Random Model",log_loss(y_cv,cv_predicted_y, eps=1e-15))

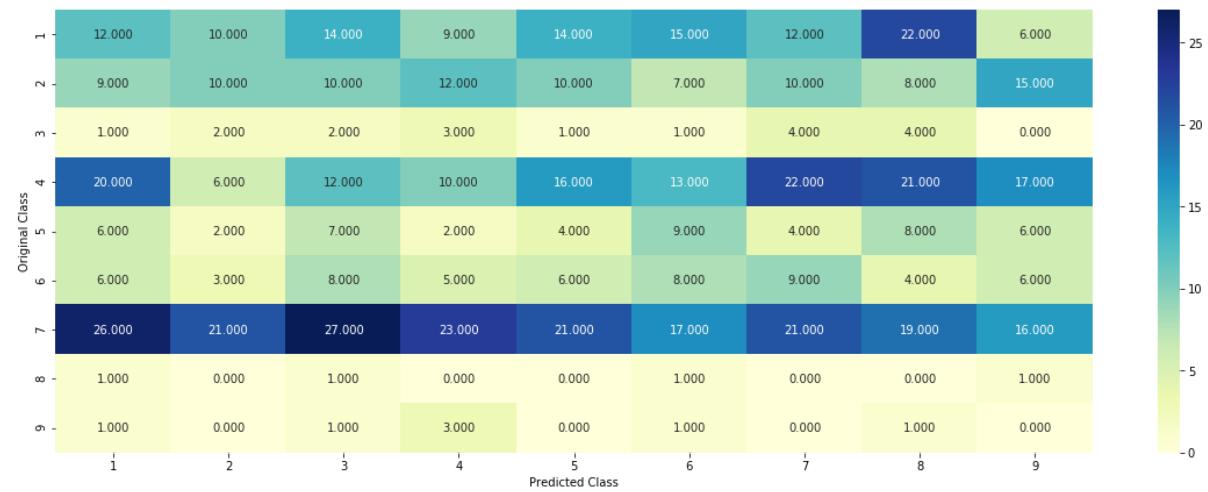
# Test-Set error.
#we create a output array that has exactly same as the test data
test_predicted_y = np.zeros((test_data_len,9))
for i in range(test_data_len):
    rand_probs = np.random.rand(1,9)
    test_predicted_y[i] = ((rand_probs/sum(sum(rand_probs)))[0])
print("Log loss on Test Data using Random Model",log_loss(y_test,test_predicted_y, eps=1e-15))

predicted_y =np.argmax(test_predicted_y, axis=1)
plot_confusion_matrix(y_test, predicted_y+1)
```

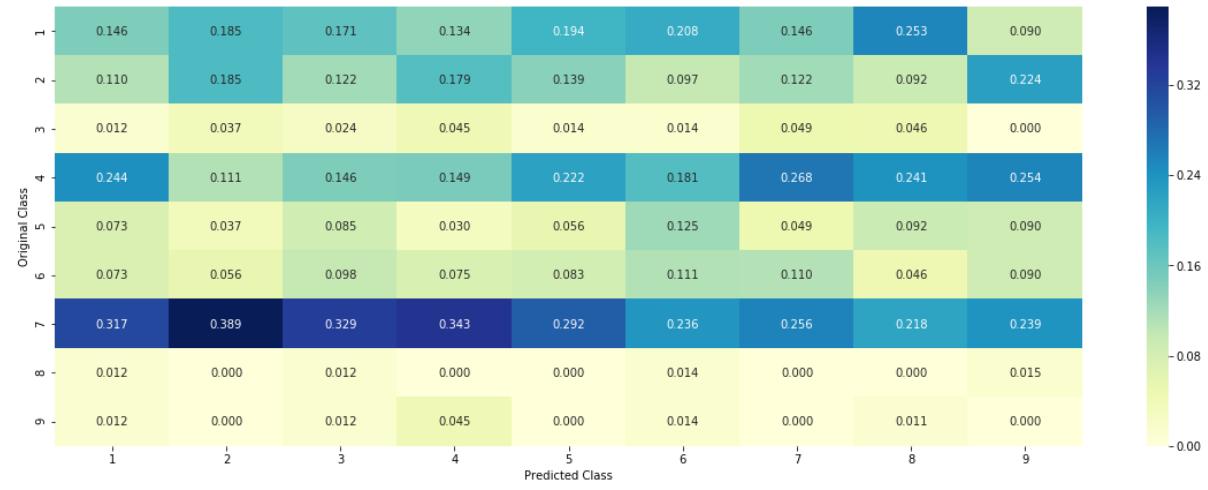
Log loss on Cross Validation Data using Random Model 2.508275926496927

Log loss on Test Data using Random Model 2.505389279903556

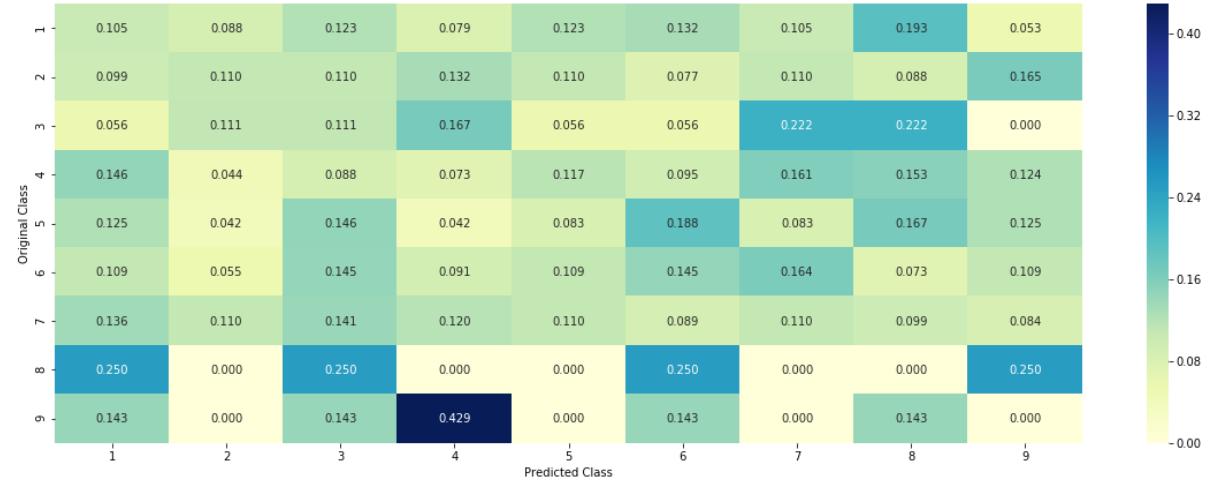
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



3.3 Univariate Analysis


```
In [17]: # code for response coding with Laplace smoothing.
# alpha : used for Laplace smoothing
# feature: ['gene', 'variation']
# df: ['train_df', 'test_df', 'cv_df']
# algorithm
# -----
# Consider all unique values and the number of occurrences of given feature in
# train data dataframe
# build a vector (1*9) , the first element = (number of times it occurred in cl
ass1 + 10*alpha / number of time it occurred in total data+90*alpha)
# gv_dict is like a look up table, for every gene it store a (1*9) representat
ion of it
# for a value of feature in df:
# if it is in train data:
# we add the vector that was stored in 'gv_dict' Look up table to 'gv_fea'
# if it is not there is train:
# we add [1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9, 1/9] to 'gv_fea'
# return 'gv_fea'
# -----
# get_gv_fea_dict: Get Gene variation Feature Dict
def get_gv_fea_dict(alpha, feature, df):
    # value_count: it contains a dict like
    # print(train_df['Gene'].value_counts())
    # output:
    #     {BRCA1      174
    #      TP53      106
    #      EGFR      86
    #      BRCA2      75
    #      PTEN      69
    #      KIT       61
    #      BRAF      60
    #      ERBB2      47
    #      PDGFRA     46
    #      ...}
    # print(train_df['Variation'].value_counts())
    # output:
    # {
    # Truncating_Mutations      63
    # Deletion                  43
    # Amplification              43
    # Fusions                   22
    # Overexpression              3
    # E17K                      3
    # Q61L                      3
    # S222D                     2
    # P130S                     2
    # ...
    # }
    value_count = train_df[feature].value_counts()

    # gv_dict : Gene Variation Dict, which contains the probability array for
    # each gene/variation
    gv_dict = dict()

    # denominator will contain the number of time that particular feature occu
```

```

red in whole data
for i, denominator in value_count.items():
    # vec will contain ( $p(y_i=1/G_i)$ ) probability of gene/variation belongs
    to particular class
    # vec is 9 dimensional vector
    vec = []
    for k in range(1,10):
        # print(train_df.loc[(train_df['Class']==1) & (train_df['Gene']=
        = 'BRCA1')])]

        #           ID   Gene          Variation Class
        # 2470  2470  BRCA1      S1715C     1
        # 2486  2486  BRCA1      S1841R     1
        # 2614  2614  BRCA1      M1R       1
        # 2432  2432  BRCA1      L1657P     1
        # 2567  2567  BRCA1      T1685A     1
        # 2583  2583  BRCA1      E1660G     1
        # 2634  2634  BRCA1      W1718L     1
        # cls_cnt.shape[0] will return the number of rows

    cls_cnt = train_df.loc[(train_df['Class']==k) & (train_df[feature]==i)]

    # cls_cnt.shape[0](numerator) will contain the number of time that
    # particular feature occurred in whole data
    vec.append((cls_cnt.shape[0] + alpha*10)/ (denominator + 90*alpha))
)

# we are adding the gene/variation to the dict as key and vec as value
gv_dict[i]=vec
return gv_dict

# Get Gene variation feature
def get_gv_feature(alpha, feature, df):
    # print(gv_dict)
    # {'BRCA1': [0.200757575757575, 0.03787878787878788, 0.068181818181818177, 0.13636363636363635, 0.25, 0.1931818181818181, 0.03787878787878788, 0.03787878787878788, 0.03787878787878788], 'TP53': [0.32142857142857145, 0.061224489795918366, 0.061224489795918366, 0.27040816326530615, 0.061224489795918366, 0.066326530612244902, 0.051020408163265307, 0.051020408163265307, 0.056122448979591837], 'EGFR': [0.0568181818181816, 0.21590909090909091, 0.0625, 0.06818181818177, 0.06818181818177, 0.0625, 0.34659090909090912, 0.0625, 0.0568181818181816], 'BRCA2': [0.1333333333333333, 0.060606060606060608, 0.060606060606060608, 0.0787878787878782, 0.1393939393939394, 0.34545454545454546, 0.060606060606060608, 0.060606060606060608, 0.060606060606060608], 'PTEN': [0.06918238937106917, 0.062893081761006289, 0.069182389937106917, 0.46540880503144655, 0.075471698113207544, 0.062893081761006289, 0.069182389937106917, 0.062893081761006289, 0.062893081761006289], 'KIT': [0.066225165562913912, 0.25165562913907286, 0.072847682119205295, 0.072847682119205295, 0.066225165562913912, 0.066225165562913912, 0.066225165562913912, 0.27152317880794702, 0.066225165562913912, 0.066225165562913912], 'BRAF': [0.066666666666666666, 0.1799999999999999, 0.073333333333333334, 0.07333333333333334, 0.09333333333333338, 0.08000000000000002, 0.2999999999999999, 0.066666666666666666, 0.066666666666666666], '# ...'
# }

```

```

gv_dict = get_gv_fea_dict(alpha, feature, df)
# value_count is similar in get_gv_fea_dict
value_count = train_df[feature].value_counts()

# gv_fea: Gene_variation feature, it will contain the feature for each feature value in the data
gv_fea = []
# for every feature values in the given data frame we will check if it is there in the train data then we will add the feature to gv_fea
# if not we will add [1/9,1/9,1/9,1/9,1/9,1/9,1/9,1/9] to gv_fea
for index, row in df.iterrows():
    if row[feature] in dict(value_count).keys():
        gv_fea.append(gv_dict[row[feature]])
    else:
        gv_fea.append([1/9,1/9,1/9,1/9,1/9,1/9,1/9,1/9])
    gv_fea.append([-1,-1,-1,-1,-1,-1,-1,-1])
return gv_fea

```

when we calculate the probability of a feature belongs to any particular class, we apply laplace smoothing

- $(\text{numerator} + 10^{10} \alpha) / (\text{denominator} + 90^{10} \alpha)$

Q3. How to featurize this Gene feature ?

Ans.there are two ways we can featurize this variable check out this video:

<https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/handling-categorical-and-numerical-features/>

1. One hot Encoding
2. Response coding

We will choose the appropriate featurization based on the ML model we use. For this problem of multi-class classification with categorical features, one-hot encoding is better for Logistic regression while response coding is better for Random Forests.

```

In [18]: # #response-coding of the Gene feature
# # alpha is used for Laplace smoothing
# alpha = 1
# # train gene feature
# train_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", train_df))
# # test gene feature
# test_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", test_df))
# # cross validation gene feature
# cv_gene_feature_responseCoding = np.array(get_gv_feature(alpha, "Gene", cv_df))

```

```

In [19]: # print("train_gene_feature_responseCoding is converted feature using response
          coding method. The shape of gene feature:", train_gene_feature_responseCoding.shape)

```

```
In [20]: # one-hot encoding of Gene feature.
gene_vectorizer = CountVectorizer()
train_gene_feature_onehotCoding = gene_vectorizer.fit_transform(train_df[ 'Gene'])
test_gene_feature_onehotCoding = gene_vectorizer.transform(test_df[ 'Gene'])
cv_gene_feature_onehotCoding = gene_vectorizer.transform(cv_df[ 'Gene'])
```

```
In [21]: print("train_gene_feature_onehotCoding is converted feature using one-hot encoding method. The shape of gene feature:", train_gene_feature_onehotCoding.shape)
```

train_gene_feature_onehotCoding is converted feature using one-hot encoding method. The shape of gene feature: (2124, 229)

Q9. How to featurize this Variation feature ?

Ans. There are two ways we can featurize this variable check out this video:

<https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/handling-categorical-and-numerical-features/>

1. One hot Encoding
2. Response coding

We will be using both these methods to featurize the Variation Feature

```
In [22]: # # alpha is used for Laplace smoothing
# alpha = 1
# # train gene feature
# train_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", train_df))
# # test gene feature
# test_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", test_df))
# # cross validation gene feature
# cv_variation_feature_responseCoding = np.array(get_gv_feature(alpha, "Variation", cv_df))
```

```
In [23]: # print("train_variation_feature_responseCoding is a converted feature using the response coding method. The shape of Variation feature:", train_variation_feature_responseCoding.shape)
```

```
In [24]: # one-hot encoding of variation feature.
variation_vectorizer = CountVectorizer()
train_variation_feature_onehotCoding = variation_vectorizer.fit_transform(train_df[ 'Variation'])
test_variation_feature_onehotCoding = variation_vectorizer.transform(test_df[ 'Variation'])
cv_variation_feature_onehotCoding = variation_vectorizer.transform(cv_df[ 'Variation'])
```

```
In [25]: print("train_variation_feature_onehotEncoded is converted feature using the on  
ne-hot encoding method. The shape of Variation feature:", train_variation_fea  
ture_onehotCoding.shape)
```

train_variation_feature_onehotEncoded is converted feature using the onne-hot encoding method. The shape of Variation feature: (2124, 1978)

3.2.3 Univariate Analysis on Text Feature

```
In [1]: import dill  
# dill.dump_session('notebook_env_assinm4final.db')  
dill.load_session('notebook_env_assinm4final.db')
```

1. How many unique words are present in train data?
2. How are word frequencies distributed?
3. How to featurize text field?
4. Is the text feature useful in predicitng y_i?
5. Is the text feature stable across train, test and CV datasets?

```
In [2]: # building a CountVectorizer with all the words that occured minimum 3 times i  
n train data  
text_vectorizer = TfidfVectorizer(ngram_range=(2,2))  
  
train_text_feature_tfidf_Coding = text_vectorizer.fit_transform(train_df['TEX  
T'])  
test_text_feature_tfidf_Coding = text_vectorizer.transform(test_df['TEXT'])  
cv_text_feature_tfidf_Coding = text_vectorizer.transform(cv_df['TEXT'])
```

```
In [3]: print(train_text_feature_tfidf_Coding.shape)  
print(test_text_feature_tfidf_Coding.shape)  
print(cv_text_feature_tfidf_Coding.shape)
```

(2124, 2181232)
(665, 2181232)
(532, 2181232)

```
In [4]: # from sklearn.feature_selection import SelectKBest  
# from sklearn.feature_selection import chi2  
  
# chi = SelectKBest(chi2,k=1000).fit(train_text_feature_tfidf_Coding, y_train)  
  
# train_text_feature_tfidf_Coding = chi.transform(train_text_feature_tfidf_Cod  
ing)  
# test_text_feature_tfidf_Coding = chi.transform(test_text_feature_tfidf_Codin  
g)  
# cv_text_feature_tfidf_Coding = chi.transform(cv_text_feature_tfidf_Coding)
```

```
In [5]: print(train_text_feature_tfidf_Coding.shape)
print(test_text_feature_tfidf_Coding.shape)
print(cv_text_feature_tfidf_Coding.shape)
```

```
(2124, 2181232)
(665, 2181232)
(532, 2181232)
```

```
In [6]: #####
```

```
In [7]: #####
```

```
In [8]: # # # don't forget to normalize every feature
# train_text_feature_tfidf_Coding = normalize(train_text_feature_tfidf_Coding,
axis=0)

# test_text_feature_tfidf_Coding = normalize(test_text_feature_tfidf_Coding, a
xis=0)

# cv_text_feature_tfidf_Coding = normalize(cv_text_feature_tfidf_Coding, axis=
0)
```

```
In [9]: # # Train a Logistic regression+Calibration model using text features which are on-hot encoded
# alpha = [10 ** x for x in range(-5, 1)]

# # read more about SGDClassifier() at http://scikit-Learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# # -----
# # default parameters
# # SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# # shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# # class_weight=None, warm_start=False, average=False, n_iter=None)

# # some of methods
# # fit(X, y[, coef_init, intercept_init, ...]) Fit linear model with Stochastic Gradient Descent.
# # predict(X) Predict class labels for samples in X.

# #-----
# # video Link:
# #-----


# cv_Log_error_array=[]
# for i in alpha:
#     clf = SGDClassifier(alpha=i, class_weight='balanced', penalty='L2', loss='log', random_state=42)
#     clf.fit(train_text_feature_tfidf_Coding, y_train)

#     sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
#     sig_clf.fit(train_text_feature_tfidf_Coding, y_train)
#     predict_y = sig_clf.predict_proba(cv_text_feature_tfidf_Coding)
#     cv_Log_error_array.append(log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
#     print('For values of alpha = ', i, "The Log Loss is:", log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))

# fig, ax = plt.subplots()
# ax.plot(alpha, cv_Log_error_array, c='g')
# for i, txt in enumerate(np.round(cv_Log_error_array, 3)):
#     ax.annotate((alpha[i], np.round(txt, 3)), (alpha[i], cv_Log_error_array[i]))
# plt.grid()
# plt.title("Cross Validation Error for each alpha")
# plt.xlabel("Alpha i's")
# plt.ylabel("Error measure")
# plt.show()

# best_alpha = np.argmin(cv_Log_error_array)
# clf = SGDClassifier(alpha=alpha[best_alpha], class_weight='balanced', penalty='L2', loss='log', random_state=42)
# clf.fit(train_text_feature_tfidf_Coding, y_train)
# sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
# sig_clf.fit(train_text_feature_tfidf_Coding, y_train)
```

```
# predict_y = sig_clf.predict_proba(train_text_feature_tfidf_Coding)
# print('For values of best alpha = ', alpha[best_alpha], "The train Log Loss
# is:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
# predict_y = sig_clf.predict_proba(cv_text_feature_tfidf_Coding)
# print('For values of best alpha = ', alpha[best_alpha], "The cross validation Log Loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
# predict_y = sig_clf.predict_proba(test_text_feature_tfidf_Coding)
# print('For values of best alpha = ', alpha[best_alpha], "The test Log Loss is:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

Q. Is the Text feature stable across all the data sets (Test, Train, Cross validation)?

Ans. Yes, it seems like!

4. Machine Learning Models

```
In [10]: #Data preparation for ML models.

#Misc. functionns for ML models

def predict_and_plot_confusion_matrix(train_x, train_y,test_x, test_y, clf):
    clf.fit(train_x, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x, train_y)
    pred_y = sig_clf.predict(test_x)

    # for calculating log_loss we will provide the array of probabilities belongs to each class
    print("Log loss :",log_loss(test_y, sig_clf.predict_proba(test_x)))
    # calculating the number of data points that are misclassified
    print("Number of mis-classified points :", np.count_nonzero((pred_y- test_y))/test_y.shape[0])
    plot_confusion_matrix(test_y, pred_y)
```

```
In [11]: def report_log_loss(train_x, train_y, test_x, test_y,  clf):
    clf.fit(train_x, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x, train_y)
    sig_clf_probs = sig_clf.predict_proba(test_x)
    return log_loss(test_y, sig_clf_probs, eps=1e-15)
```

```
In [28]: # this function will be used just for naive bayes
# for the given indices, we will print the name of the features
# and we will check whether the feature present in the test point text or not
def get_impfeature_names(indices, text, gene, var, no_features):
    gene_count_vec = CountVectorizer()
    var_count_vec = CountVectorizer()
    text_count_vec = TfidfVectorizer()

    gene_vec = gene_count_vec.fit(train_df['Gene'])
    var_vec = var_count_vec.fit(train_df['Variation'])
    text_vec = text_count_vec.fit(train_df['TEXT'])

    fea1_len = len(gene_vec.vocabulary_.keys())
    fea2_len = len(var_count_vec.vocabulary_.keys())

    word_present = 0
    for i,v in enumerate(indices):
        if (v < fea1_len):
            word = gene_vec.get_feature_names()[v]
            yes_no = True if word == gene else False
            if yes_no:
                word_present += 1
                print(i, "Gene feature [{}] present in test data point [{}].format(word,yes_no))
            elif (v < fea1_len+fea2_len):
                word = var_vec.get_feature_names()[v-(fea1_len)]
                yes_no = True if word == var else False
                if yes_no:
                    word_present += 1
                    print(i, "variation feature [{}] present in test data point [{}].format(word,yes_no))
            else:
                word = text_vec.get_feature_names()[v-(fea1_len+fea2_len)]
                yes_no = True if word in text.split() else False
                if yes_no:
                    word_present += 1
                    print(i, "Text feature [{}] present in test data point [{}].format(word,yes_no))

    print("Out of the top ",no_features," features ", word_present, "are present in query point")
```

Stacking the three types of features

In [13]: # merging gene, variance and text features

```
# building train, test and cross validation data sets
# a = [[1, 2],
#       [3, 4]]
# b = [[4, 5],
#       [6, 7]]
# hstack(a, b) = [[1, 2, 4, 5],
#                  [3, 4, 6, 7]]

train_gene_var_onehotCoding = hstack((train_gene_feature_onehotCoding,train_variation_feature_onehotCoding))
test_gene_var_onehotCoding = hstack((test_gene_feature_onehotCoding,test_variation_feature_onehotCoding))
cv_gene_var_onehotCoding = hstack((cv_gene_feature_onehotCoding,cv_variation_feature_onehotCoding))

train_x_onehotCoding = hstack((train_gene_var_onehotCoding, train_text_feature_tfidf_Coding)).tocsr()

test_x_onehotCoding = hstack((test_gene_var_onehotCoding, test_text_feature_tfidf_Coding)).tocsr()

cv_x_onehotCoding = hstack((cv_gene_var_onehotCoding, cv_text_feature_tfidf_Coding)).tocsr()

# train_gene_var_responseCoding = np.hstack((train_gene_feature_responseCoding,train_variation_feature_responseCoding))
# test_gene_var_responseCoding = np.hstack((test_gene_feature_responseCoding,test_variation_feature_responseCoding))
# cv_gene_var_responseCoding = np.hstack((cv_gene_feature_responseCoding,cv_variation_feature_responseCoding))

# train_x_responseCoding = hstack((train_gene_var_responseCoding, train_text_feature_tfidf_Coding)).tocsr()
# test_x_responseCoding = hstack((test_gene_var_responseCoding, test_text_feature_tfidf_Coding)).tocsr()
# cv_x_responseCoding = hstack((cv_gene_var_responseCoding, cv_text_feature_tfidf_Coding)).tocsr()

train_y = np.array(list(train_df['Class']))
test_y = np.array(list(test_df['Class']))
cv_y = np.array(list(cv_df['Class']))
```

```
In [14]: print("One hot encoding features :")
print("(number of data points * number of features) in train data = ", train_x
      _onehotCoding.shape)
print("(number of data points * number of features) in test data = ", test_x_o
      nehotCoding.shape)
print("(number of data points * number of features) in cross validation data
      =", cv_x_onehotCoding.shape)
```

One hot encoding features :
 (number of data points * number of features) in train data = (2124, 2183439)
 (number of data points * number of features) in test data = (665, 2183439)
 (number of data points * number of features) in cross validation data = (532,
 2183439)

```
In [15]: # print(" Response encoding features :")
# print("(number of data points * number of features) in train data = ", train
      _x_responseCoding.shape)
# print("(number of data points * number of features) in test data = ", test_x
      _responseCoding.shape)
# print("(number of data points * number of features) in cross validation data
      =", cv_x_responseCoding.shape)
```

In []:

```
In [16]: # https://machinelearningmastery.com/how-to-fix-futurewarning-messages-in-scik
      it-Learn/
from warnings import simplefilter
# ignore all future warnings
simplefilter(action='ignore', category=FutureWarning)
```

In []:

4.3. Logistic Regression

4.3.1. With Class balancing

4.3.1.1. Hyper parameter tuning

```
In [17]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----


# find more about CalibratedClassifierCV here at http://scikit-learn.org/stable/modules/generated/sklearn.calibration.CalibratedClassifierCV.html
# -----
# default paramters
# sklearn.calibration.CalibratedClassifierCV(base_estimator=None, method='sigmoid', cv=3)
#
# some of the methods of CalibratedClassifierCV()
# fit(X, y[, sample_weight])      Fit the calibrated model
# get_params([deep])      Get parameters for this estimator.
# predict(X)      Predict the target of new samples.
# predict_proba(X)    Posterior probabilities of classification
#-----
# video link:
#-----


alpha = [10 ** x for x in range(-6, 3)]
cv_log_error_array = []
for i in alpha:
    print("for alpha =", i)
    clf = SGDClassifier(class_weight='balanced', alpha=i, penalty='L2', loss='log', random_state=42)
    clf.fit(train_x_onehotCoding, train_y)
    sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
    sig_clf.fit(train_x_onehotCoding, train_y)
    sig_clf_probs = sig_clf.predict_proba(cv_x_onehotCoding)
    cv_log_error_array.append(log_loss(cv_y, sig_clf_probs, labels=clf.classes_, eps=1e-15))
        # to avoid rounding error while multiplying probabilites we use log-probability estimates
    print("Log Loss :",log_loss(cv_y, sig_clf_probs))

fig, ax = plt.subplots()
ax.plot(alpha, cv_log_error_array,c='g')
for i, txt in enumerate(np.round(cv_log_error_array,3)):
```

```
        ax.annotate((alpha[i],str(txt)), (alpha[i],cv_log_error_array[i]))
plt.grid()
plt.title("Cross Validation Error for each alpha")
plt.xlabel("Alpha i's")
plt.ylabel("Error measure")
plt.show()

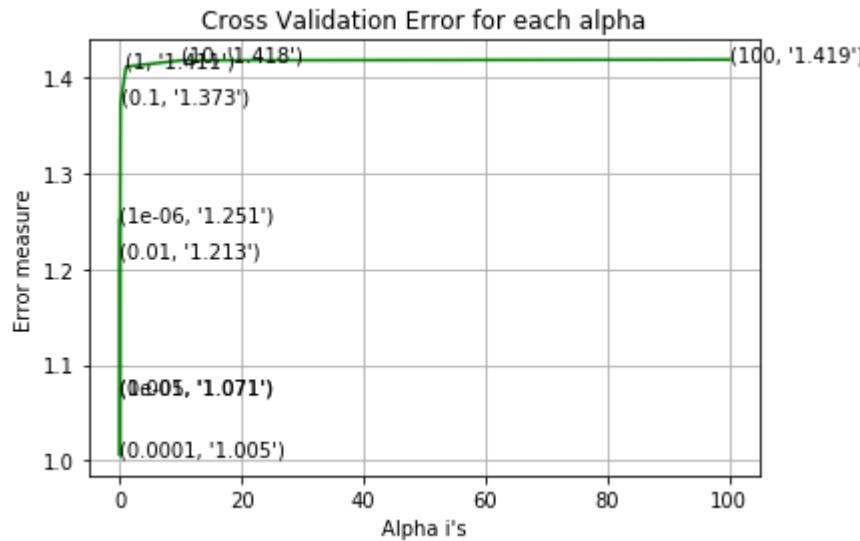
best_alpha = np.argmin(cv_log_error_array)
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty=
'12', loss='log', random_state=42)
clf.fit(train_x_onehotCoding, train_y)
sig_clf = CalibratedClassifierCV(clf, method="sigmoid")
sig_clf.fit(train_x_onehotCoding, train_y)

predict_y = sig_clf.predict_proba(train_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The train log loss i
s:",log_loss(y_train, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(cv_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The cross validation
log loss is:",log_loss(y_cv, predict_y, labels=clf.classes_, eps=1e-15))
predict_y = sig_clf.predict_proba(test_x_onehotCoding)
print('For values of best alpha = ', alpha[best_alpha], "The test log loss i
s:",log_loss(y_test, predict_y, labels=clf.classes_, eps=1e-15))
```

```

for alpha = 1e-06
Log Loss : 1.2513960386960545
for alpha = 1e-05
Log Loss : 1.071130733503459
for alpha = 0.0001
Log Loss : 1.0049903999838221
for alpha = 0.001
Log Loss : 1.0708855912179018
for alpha = 0.01
Log Loss : 1.212591212597999
for alpha = 0.1
Log Loss : 1.3733486441937253
for alpha = 1
Log Loss : 1.4112895398556229
for alpha = 10
Log Loss : 1.4178604898877982
for alpha = 100
Log Loss : 1.4186625402480884

```



For values of best alpha = 0.0001 The train log loss is: 0.3779257994483267
 For values of best alpha = 0.0001 The cross validation log loss is: 1.0049903999838221
 For values of best alpha = 0.0001 The test log loss is: 0.92941734371344

4.3.1.2. Testing the model with best hyper paramters

```
In [18]: # read more about SGDClassifier() at http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html
# -----
# default parameters
# SGDClassifier(loss='hinge', penalty='L2', alpha=0.0001, l1_ratio=0.15, fit_intercept=True, max_iter=None, tol=None,
# shuffle=True, verbose=0, epsilon=0.1, n_jobs=1, random_state=None, learning_rate='optimal', eta0=0.0, power_t=0.5,
# class_weight=None, warm_start=False, average=False, n_iter=None)

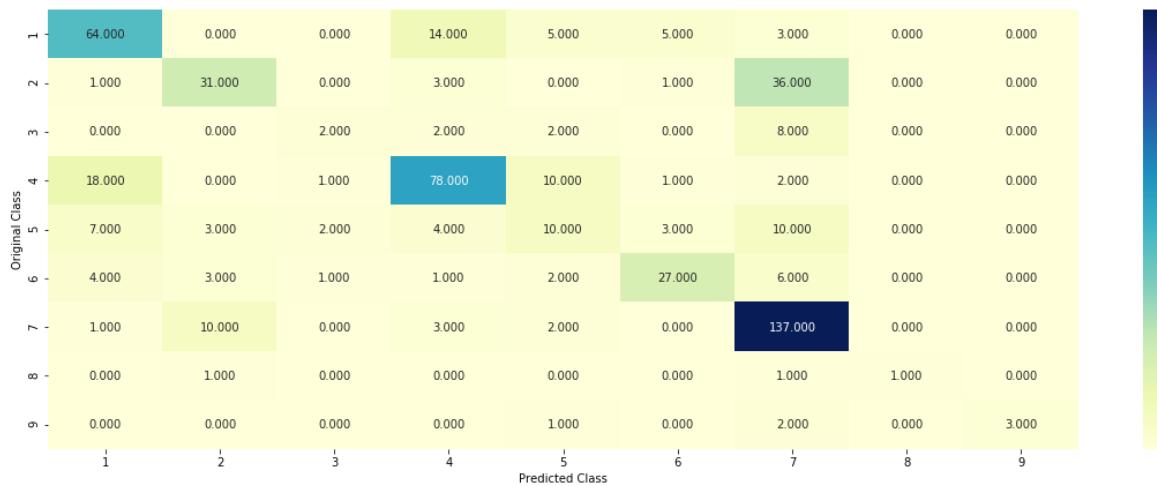
# some of methods
# fit(X, y[, coef_init, intercept_init, ...])      Fit linear model with Stochastic Gradient Descent.
# predict(X)   Predict class labels for samples in X.

#-----
# video link: https://www.appliedaicourse.com/course/applied-ai-course-online/lessons/geometric-intuition-1/
#-----
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
predict_and_plot_confusion_matrix(train_x_onehotCoding, train_y, cv_x_onehotCoding, cv_y, clf)
```

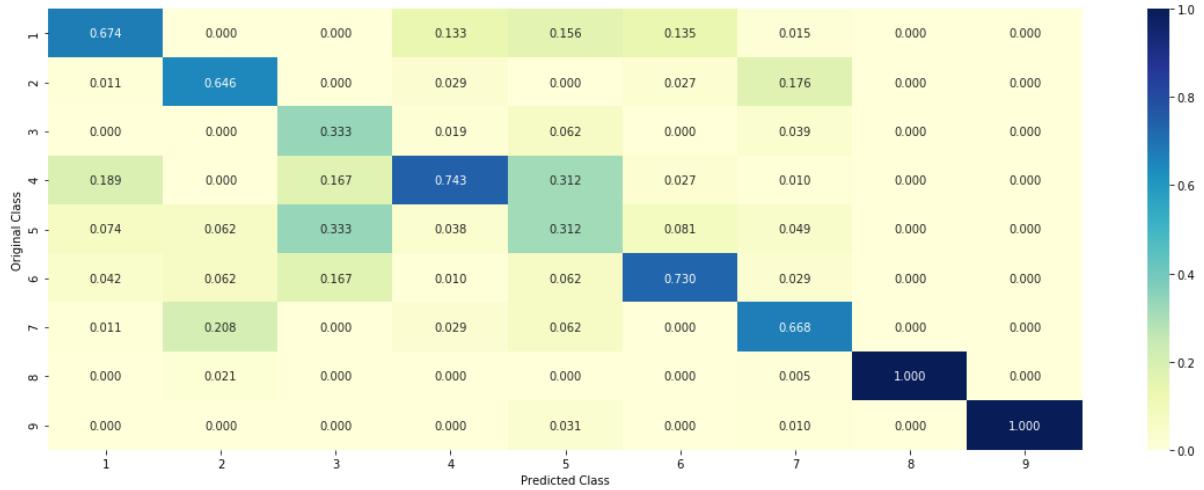
Log loss : 1.0049903999838221

Number of mis-classified points : 0.33646616541353386

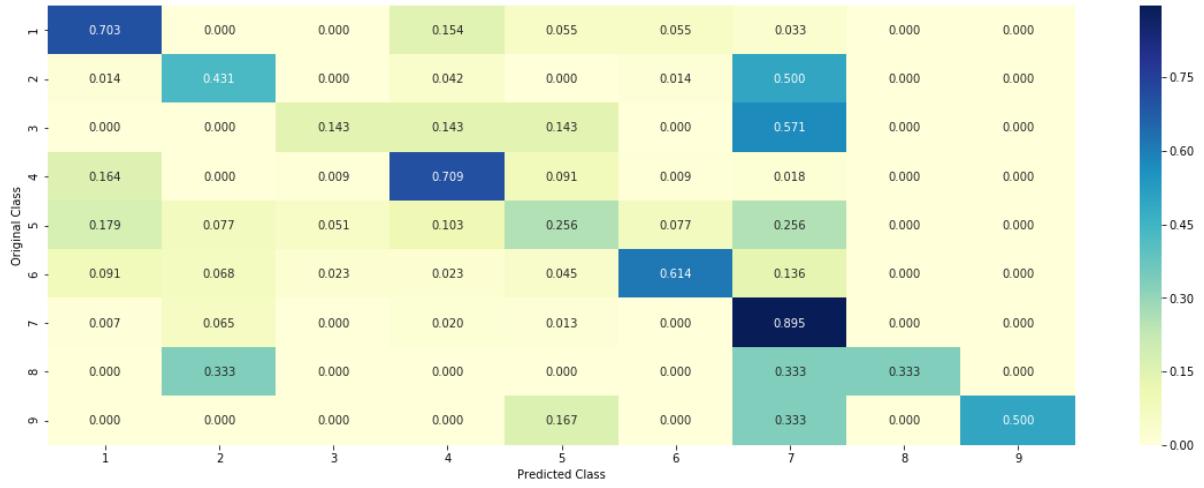
----- Confusion matrix -----



----- Precision matrix (Column Sum=1) -----



----- Recall matrix (Row sum=1) -----



4.3.1.3. Feature Importance

```
In [19]: def get_imp_feature_names(text, indices, removed_ind = []):
    word_present = 0
    tabulte_list = []
    incresingorder_ind = 0
    for i in indices:
        if i < train_gene_feature_onehotCoding.shape[1]:
            tabulte_list.append([incresingorder_ind, "Gene", "Yes"])
        elif i < 18:
            tabulte_list.append([incresingorder_ind, "Variation", "Yes"])
        if ((i > 17) & (i not in removed_ind)):
            word = train_text_features[i]
            yes_no = True if word in text.split() else False
            if yes_no:
                word_present += 1
            tabulte_list.append([incresingorder_ind, train_text_features[i], yes_no])
        incresingorder_ind += 1
    print(word_present, "most important features are present in our query point")
    print("-"*50)
    print("The features that are most important of the ",predicted_cls[0]," class:")
    print (tabulate(tabulte_list, headers=["Index", 'Feature name', 'Present or Not']))
```

4.3.1.3.1. Correctly Classified point

```
In [30]: # from tabulate import tabulate
clf = SGDClassifier(class_weight='balanced', alpha=alpha[best_alpha], penalty='l2', loss='log', random_state=42)
clf.fit(train_x_onehotCoding,train_y)
test_point_index = 50
no_feature = 10
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class : ", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_onehotCoding[test_point_index]),4))
print("Actual Class : ", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_imffeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_df['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index], no_feature)

Predicted Class : 7
Predicted Class Probabilities: [[0.0197 0.2715 0.009  0.0248 0.1142 0.0296 0.
5226 0.0044 0.0043]]
Actual Class : 7
-----
Out of the top 10 features 0 are present in query point
```

4.3.1.3.2. Incorrectly Classified point

```
In [31]: test_point_index = 361
no_feature = 10
predicted_cls = sig_clf.predict(test_x_onehotCoding[test_point_index])
print("Predicted Class :", predicted_cls[0])
print("Predicted Class Probabilities:", np.round(sig_clf.predict_proba(test_x_
onehotCoding[test_point_index]),4))
print("Actual Class :", test_y[test_point_index])
indices = np.argsort(-abs(clf.coef_))[predicted_cls-1][:,:no_feature]
print("-"*50)
get_impfeature_names(indices[0], test_df['TEXT'].iloc[test_point_index],test_d
f['Gene'].iloc[test_point_index],test_df['Variation'].iloc[test_point_index],
no_feature)
```

```
Predicted Class : 7
Predicted Class Probabilities: [[0.0446 0.3025 0.0145 0.0598 0.029 0.019 0.
5183 0.0056 0.0067]]
Actual Class : 2
-----
Out of the top 10 features 0 are present in query point
```

```
In [ ]:
```

Results(Pretty Table):

```
In [20]: from prettytable import PrettyTable
# Names of models
model=[ 'Naive Bayes ','KNN','Logistic Regression With Class balancing '
       , 'LogisticRegression Without Class balancing','Linear SVM '
       , 'Random Forest Classifier With One hot Encoding','Random Forest Classifier With Response Coding'
       , 'Stack Models:LR+NB+SVM','Maximum Voting classifier'
       , 'LR(BALANCED): CountVectorizer Features, including both unigrams and bigrams','LR(UNBALANCED): CountVectorizer Features, including both unigrams and bigrams'
       , 'LR: after feature engineering']
train =[0.67,0.82,0.44,0.44,0.40, 0.68,0.10,0.59,0.81,0.71,0.70,0.37]
cv=[1.18,1.15,1.01,1.01,1.08,1.08,0.99,1.20,1.17,1.11,1.12,1.00]
test = [1.11,1.07,0.93,0.93,0.99,1.01,0.89,1.09,1.12,1.15,1.14,0.92]
mp=[40,41,36,36,37,37,33,33,33,35,36,33]
numbering=['1 & 2','1 & 2','1 & 2','1 & 2','1 & 2','1 & 2','1 & 2','1 & 2','1 & 2',3,3,4]
# Initializing prettytable
ptable = PrettyTable()
# Adding columns
ptable.add_column("S.NO.",numbering)
ptable.add_column("model",model)
ptable.add_column("train",train)
ptable.add_column("cv",cv)
ptable.add_column("test",test)
ptable.add_column("% Misclassified Points",mp)
# Printing the Table
print(ptable)
```

S.NO.	model		
train	cv	test	% Misclassified Points
1 & 2			Naive Bayes
0.67	1.18	1.11	40
1 & 2			KNN
0.82	1.15	1.07	41
1 & 2			Logistic Regression With Class balancing
0.44	1.01	0.93	36
1 & 2			LogisticRegression Without Class balancing
0.44	1.01	0.93	36
1 & 2			Linear SVM
0.4	1.08	0.99	37
1 & 2			Random Forest Classifier With One hot Encoding
0.68	1.08	1.01	37
1 & 2			Random Forest Classifier With Response Coding
0.1	0.99	0.89	33
1 & 2			Stack Models:LR+NB+SVM
0.59	1.2	1.09	33
1 & 2			Maximum Voting classifier
0.81	1.17	1.12	33
3	LR(BALANCED): CountVectorizer Features, including both unigrams and bigrams		
	0.71	1.11	1.15
			35
3	LR(UNBALANCED): CountVectorizer Features, including both unigrams and bigrams		
	0.7	1.12	1.14
			36
4	LR: after feature engineering		
0.37	1.0	0.92	33

Procedure to Solve this case study:

1. For the 1 & 2 part of the assignment ,first i took TFIDF vectorization of text data and took top 1000 using k-best ,and also One hot encoding of Categorical features
2. For those futurization , for NB I got test log-loss of 1.11 and 40 % misclassified points
3. For Knn I got test log-loss of 1.07 and 41 % misclassified points
4. For LR with balanced I got test log-loss of 0.93 and 36 % misclassified points
5. For LR without balanced I got test log-loss of 0.93 and 36 % misclassified points
6. For SVM I got test log-loss of 0.99 and 37 % misclassified points
7. For RF with one-hot-encoding I got test log-loss of 1.01 and 37 % misclassified points
8. For RF with response-encoding I got test log-loss of 0.89 and 33 % misclassified points
9. For STACKING I got test log-loss of 1.09 and 33 % misclassified points
10. For VOTING I got test log-loss of 1.12 and 33 % misclassified points
11. Now for the 3rd part of the assignment, I took LR as a model and tested with count vectorization of features both on Balanced and unbalanced LR
12. For Balanced LR I got test Log-loss of 1.15 and 35 % misclassified points
13. For UN-Balanced LR I got test Log-loss of 1.14 and 36 % misclassified points
14. Now For the 4th part of the assignment TFIDF vectorization of text data with only BI-Grams ,one-hot of both Gene and Variation feature and used LR with balanced as the model , got a test log-loss of 0.92 and 33 % misclassified points

In []: