In [68]:
```
import requests
```

In [2]:
```
!pip install requests
```

```
Requirement already satisfied: requests in /Users/alexandroskanter
akis/anaconda3/lib/python3.7/site-packages (2.19.1)
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in /Users/ale
xandroskanterakis/anaconda3/lib/python3.7/site-packages (from requ
ests) (3.0.4)
Requirement already satisfied: idna<2.8,>=2.5 in /Users/alexandros
kanterakis/anaconda3/lib/python3.7/site-packages (from requests)
(2.7)
Requirement already satisfied: urllib3<1.24,>=1.21.1 in /Users/ale
xandroskanterakis/anaconda3/lib/python3.7/site-packages (from requ
ests) (1.23)
Requirement already satisfied: certifi>=2017.4.17 in /Users/alexan
droskanterakis/anaconda3/lib/python3.7/site-packages (from request
s) (2019.11.28)
WARNING: You are using pip version 20.1; however, version 20.3.3 i
s available.
You should consider upgrading via the '/Users/alexandroskanterakis
/anaconda3/bin/python -m pip install --upgrade pip' command.
```

In [4]:
```
import requests

url = "https://rest.ensembl.org/vep/human/hgvs/AGT:c.803T>C?"

r = requests.get(url, headers={ "Content-Type" : "application/jso
n"})

#decoded = r.json()
#print(repr(decoded))
```

In [7]:
```
j = r.json()
```

In [8]:
```
j
```

Out[8]:

```
[{'start': 230710048,
  'assembly_name': 'GRCh38',
  'strand': -1,
  'seq_region_name': '1',
  'allele_string': 'T/C',
  'most_severe_consequence': 'missense_variant',
  'transcript_consequences': [{'cds_start': 803,
    'impact': 'MODERATE',
    'hgnc_id': 'HGNC:333',
    'amino_acids': 'M/T',
    'codons': 'aTg/aCg',
    'cdna_end': 843,
    'gene_symbol_source': 'HGNC',
    'consequence_terms': ['missense_variant'],
    'cdna_start': 843,
    'strand': -1,
    'transcript_id': 'ENST00000366667',
    'variant_allele': 'C',
    'gene_id': 'ENSG00000135744',
    'biotype': 'protein_coding',
    'sift_prediction': 'tolerated',
    'polyphen_score': 0,
    'protein_start': 268,
    'polyphen_prediction': 'benign',
    'gene_symbol': 'AGT',
    'protein_end': 268,
    'cds_end': 803,
    'sift_score': 1},
   {'distance': 650,
    'strand': -1,
    'impact': 'MODIFIER',
    'gene_id': 'ENSG00000244137',
    'variant_allele': 'C',
    'transcript_id': 'ENST00000412344',
    'gene_symbol_source': 'Clone_based_ensembl_gene',
    'biotype': 'lncRNA',
    'gene_symbol': 'AL512328.1',
    'consequence_terms': ['downstream_gene_variant']}],
  'colocated_variants': [{'start': 230710048,
    'strand': 1,
    'phenotype_or_disease': 1,
    'seq_region_name': '1',
    'end': 230710048,
    'id': 'CM920010',
    'allele_string': 'HGMD_MUTATION'},
   {'seq_region_name': '1',
    'frequencies': {'C': {'gnomad_amr': 0.7196,
      'eas': 0.8532,
      'eur': 0.4115,
      'gnomad': 0.5481,
      'aa': 0.8268,
      'gnomad_eas': 0.8388,
      'ea': 0.4258,
      'afr': 0.9032,
      'gnomad_nfe': 0.4197,
      'sas': 0.636,
      'gnomad_afr': 0.8451,
      'gnomad_fin': 0.441,
      'gnomad_sas': 0.6202,
```

```
      'gnomad_oth': 0.5002,
      'gnomad_asj': 0.44,
      'amr': 0.6354}},
    'start': 230710048,
    'strand': 1,
    'var_synonyms': 'ClinVar::RCV000835695,RCV000405686,VCV0000180
68,RCV000019691,RCV000019692,RCV000019693,RCV000242838--OMIM::1061
50.0001--PharmGKB::PA166153539--Uniprot::VAR_007096',
    'allele_string': 'A/G',
    'clin_sig': ['benign', 'risk_factor'],
    'end': 230710048,
    'clin_sig_allele': 'G:benign;G:risk_factor',
    'phenotype_or_disease': 1,
    'pubmed': [19131662,
     19263529,
     19330901,
     19559392,
     21919968,
     27616475,
     25741868,
     21467728,
     1394429,
     7649545,
     7883995,
     8348146,
     8513325,
     8518804,
     9259580,
     9421481,
     9831339,
     16059745,
     18069999,
     18248681,
     18279468,
     18513389,
     18603647,
     18637188,
     18653189,
     18698212,
     18953568,
     19108684,
     19770777,
     19932491,
     20029521,
     20047954,
     20061926,
     20185782,
     20486282,
     20570668,
     20577119,
     20592051,
     20811292,
     20981351,
     21056700,
     21058046,
     21127830,
     21146954,
     21261619,
     21304999,
     21306748,
```

```
21438754,
21444836,
21515823,
21533139,
21540342,
21573014,
21681796,
21894447,
21988197,
22099458,
22100073,
22531885,
22569109,
22817530,
22858200,
23021345,
23036011,
23132613,
23133444,
23205182,
23251296,
23287839,
23333443,
23354977,
23497168,
23681449,
23716723,
24452035,
24622918,
24722536,
24737640,
25474356,
25512783,
25683681,
25723521,
26102248,
26283679,
26318936,
26335431,
26509357,
26588355,
26621708,
26627480,
26818744,
26819062,
26824906,
26933222,
27068935,
27274104,
27342049,
27348238,
27380726,
27454254,
27480094,
27584680,
27940662,
28361007,
28488548,
28605058,
28666769,
```

```
            28690685,
            28770234,
            28828324,
            28881807,
            28903744,
            29057680,
            29520984,
            2957843],
           'minor_allele': 'A',
           'minor_allele_freq': 0.2949,
           'id': 'rs699'}],
         'end': 230710048,
         'id': 'AGT:c.803T>C',
```

In [9]:
```
a = [1,2,3, {'aa': 'asdasD', '1': {'BBB': [1,2,3,2,1,1]}},]
```

In [10]:
```
a
```

Out[10]:
```
[1, 2, 3, {'aa': 'asdasD', '1': {'BBB': [1, 2, 3, 2, 1, 1]}}]
```

Serialization

In [11]:
```
import json
```

In [12]:
```
json.dumps(a)
```

Out[12]:
```
'[1, 2, 3, {"aa": "asdasD", "1": {"BBB": [1, 2, 3, 2, 1, 1]}}]'
```

In [13]:
```
type(json.dumps(a))
```

Out[13]:
```
str
```

In [14]:
```
with open('data.json', 'w') as f:
    json.dump(a, f)
```

In [15]:
```
!cat data.json
```

```
[1, 2, 3, {"aa": "asdasD", "1": {"BBB": [1, 2, 3, 2, 1, 1]}}]
```

In [16]:
```
with open('data.json') as f:
    b = json.load(f)
```

In [18]:
```
b
```

Out[18]:
```
[1, 2, 3, {'aa': 'asdasD', '1': {'BBB': [1, 2, 3, 2, 1, 1]}}]
```

In [19]:
```
type(b)
```

Out[19]:
```
list
```

In [20]:
```
a_str = json.dumps(a)
```

In [21]:
```
a_str
```

Out[21]:
```
'[1, 2, 3, {"aa": "asdasD", "1": {"BBB": [1, 2, 3, 2, 1, 1]}}]'
```

In [22]:
```
b = json.loads(a_str)
```

In [25]:
```
b
```

Out[25]:
```
[1, 2, 3, {'aa': 'asdasD', '1': {'BBB': [1, 2, 3, 2, 1, 1]}}]
```

In [29]:
```
s = {4,5,6,7}
```

In [30]:
```
json.dumps(s)
```

```
---------------------------------------------------------------------
---------
TypeError                                 Traceback (most recent c
all last)
<ipython-input-30-eb8174a6fbaa> in <module>()
----> 1 json.dumps(s)

~/anaconda3/lib/python3.7/json/__init__.py in dumps(obj, skipkeys,
ensure_ascii, check_circular, allow_nan, cls, indent, separators,
default, sort_keys, **kw)
    229         cls is None and indent is None and separators is N
one and
    230         default is None and not sort_keys and not kw):
--> 231         return _default_encoder.encode(obj)
    232     if cls is None:
    233         cls = JSONEncoder

~/anaconda3/lib/python3.7/json/encoder.py in encode(self, o)
    197         # exceptions aren't as detailed.  The list call sh
ould be roughly
    198         # equivalent to the PySequence_Fast that ''.join()
would do.
--> 199         chunks = self.iterencode(o, _one_shot=True)
    200         if not isinstance(chunks, (list, tuple)):
    201             chunks = list(chunks)

~/anaconda3/lib/python3.7/json/encoder.py in iterencode(self, o, _
one_shot)
    255                     self.key_separator, self.item_separator, s
elf.sort_keys,
    256                     self.skipkeys, _one_shot)
--> 257         return _iterencode(o, 0)
    258
    259 def _make_iterencode(markers, _default, _encoder, _indent,
_floatstr,

~/anaconda3/lib/python3.7/json/encoder.py in default(self, o)
    177
    178         """
--> 179         raise TypeError(f'Object of type {o.__class__.__na
me__} '
    180                         f'is not JSON serializable')
    181

TypeError: Object of type set is not JSON serializable
```

In [28]:
```
import pickle
```

In [31]:
```
pickle.dumps(s)
```

Out[31]:
```
b'\x80\x03cbuiltins\nset\nq\x00]q\x01(K\x04K\x05K\x06K\x07e\x85q\x
02Rq\x03.'
```

In [32]:
```
def f():
    print ("hello")
    return 42
```

In [34]:
```
pickle.dumps(f)
```

Out[34]:
```
b'\x80\x03c__main__\nf\nq\x00.'
```

In [36]:
```
a.append(f)
```

In [37]:
```
a
```

Out[37]:
```
[1,
 2,
 3,
 {'aa': 'asdasD', '1': {'BBB': [1, 2, 3, 2, 1, 1]}},
 <function __main__.f()>]
```

In [38]:
```
pickle.dumps(a)
```

Out[38]:
```
b'\x80\x03]q\x00(K\x01K\x02K\x03}q\x01(X\x02\x00\x00\x00aaq\x02X\x
06\x00\x00\x00asdasDq\x03X\x01\x00\x00\x001q\x04}q\x05X\x03\x00\x0
0\x00BBBq\x06]q\x07(K\x01K\x02K\x03K\x02K\x01K\x01esuc__main__\nf\
nq\x08e.'
```

In [42]:
```
with open('data.pickle', 'wb') as f2:
    pickle.dump(a[:-1], f2)
```

In [ ]:
```

```

In [43]:
```
url = "https://rest.ensembl.org/vep/human/hgvs/AGT:c.803T>C?"

r = requests.get(url, headers={ "Content-Type" : "application/jso
n"})
```

In [44]:
```
r.ok
```

Out[44]:
```
True
```

In [45]:
```
data = r.json()
```

In [46]:
```
data
```

Out[46]:

```
[{'end': 230710048,
  'strand': -1,
  'assembly_name': 'GRCh38',
  'colocated_variants': [{'end': 230710048,
    'seq_region_name': '1',
    'allele_string': 'HGMD_MUTATION',
    'start': 230710048,
    'strand': 1,
    'id': 'CM920010',
    'phenotype_or_disease': 1},
   {'strand': 1,
    'frequencies': {'C': {'gnomad_oth': 0.5002,
      'gnomad_asj': 0.44,
      'gnomad': 0.5481,
      'gnomad_afr': 0.8451,
      'gnomad_amr': 0.7196,
      'aa': 0.8268,
      'gnomad_eas': 0.8388,
      'amr': 0.6354,
      'afr': 0.9032,
      'eur': 0.4115,
      'sas': 0.636,
      'gnomad_sas': 0.6202,
      'gnomad_fin': 0.441,
      'gnomad_nfe': 0.4197,
      'eas': 0.8532,
      'ea': 0.4258}},
    'allele_string': 'A/G',
    'minor_allele': 'A',
    'end': 230710048,
    'minor_allele_freq': 0.2949,
    'pubmed': [19131662,
     19263529,
     19330901,
     19559392,
     21919968,
     27616475,
     25741868,
     21467728,
     1394429,
     7649545,
     7883995,
     8348146,
     8513325,
     8518804,
     9259580,
     9421481,
     9831339,
     16059745,
     18069999,
     18248681,
     18279468,
     18513389,
     18603647,
     18637188,
     18653189,
     18698212,
     18953568,
     19108684,
```

```
19770777,
19932491,
20029521,
20047954,
20061926,
20185782,
20486282,
20570668,
20577119,
20592051,
20811292,
20981351,
21056700,
21058046,
21127830,
21146954,
21261619,
21304999,
21306748,
21438754,
21444836,
21515823,
21533139,
21540342,
21573014,
21681796,
21894447,
21988197,
22099458,
22100073,
22531885,
22569109,
22817530,
22858200,
23021345,
23036011,
23132613,
23133444,
23205182,
23251296,
23287839,
23333443,
23354977,
23497168,
23681449,
23716723,
24452035,
24622918,
24722536,
24737640,
25474356,
25512783,
25683681,
25723521,
26102248,
26283679,
26318936,
26335431,
26509357,
26588355,
```

```
      26621708,
      26627480,
      26818744,
      26819062,
      26824906,
      26933222,
      27068935,
      27274104,
      27342049,
      27348238,
      27380726,
      27454254,
      27480094,
      27584680,
      27940662,
      28361007,
      28488548,
      28605058,
      28666769,
      28690685,
      28770234,
      28828324,
      28881807,
      28903744,
      29057680,
      29520984,
      2957843],
     'id': 'rs699',
     'clin_sig_allele': 'G:benign;G:risk_factor',
     'phenotype_or_disease': 1,
     'seq_region_name': '1',
     'start': 230710048,
     'clin_sig': ['benign', 'risk_factor'],
     'var_synonyms': 'ClinVar::RCV000835695,RCV000405686,VCV0000180
68,RCV000019691,RCV000019692,RCV000019693,RCV000242838--OMIM::1061
50.0001--PharmGKB::PA166153539--Uniprot::VAR_007096'}],
  'seq_region_name': '1',
  'id': 'AGT:c.803T>C',
  'transcript_consequences': [{'polyphen_prediction': 'benign',
    'sift_score': 1,
    'variant_allele': 'C',
    'sift_prediction': 'tolerated',
    'amino_acids': 'M/T',
    'biotype': 'protein_coding',
    'impact': 'MODERATE',
    'gene_id': 'ENSG00000135744',
    'transcript_id': 'ENST00000366667',
    'hgnc_id': 'HGNC:333',
    'protein_end': 268,
    'gene_symbol_source': 'HGNC',
    'strand': -1,
    'cdna_start': 843,
    'cds_start': 803,
    'cdna_end': 843,
    'gene_symbol': 'AGT',
    'codons': 'aTg/aCg',
    'polyphen_score': 0,
    'consequence_terms': ['missense_variant'],
    'protein_start': 268,
    'cds_end': 803},
```

```
  {'distance': 650,
   'consequence_terms': ['downstream_gene_variant'],
   'transcript_id': 'ENST00000412344',
   'gene_id': 'ENSG00000244137',
   'gene_symbol': 'AL512328.1',
   'impact': 'MODIFIER',
   'biotype': 'lncRNA',
   'variant_allele': 'C',
   'gene_symbol_source': 'Clone_based_ensembl_gene',
   'strand': -1}],
 'input': 'AGT:c.803T>C',
 'allele_string': 'T/C',
 'start': 230710048,
```

In [55]:
```
url = "https://rest.ensembl.org/vep/human/hgvs/AGT:c.803T>C?"

r = requests.get(url, headers={ "Content-Type" : "application/jso
n"})
```

In [51]:
```
r.ok
```

Out[51]:
```
False
```

In [52]:
```
r.json()
```

Out[52]:
```
{'error': "Unable to parse HGVS notation 'AGTAAA:c.803T>C': Could
not get a Transcript object for 'AGTAAA'"}
```

In [53]:
```
r.text
```

Out[53]:
```
'{"error":"Unable to parse HGVS notation \'AGTAAA:c.803T>C\': Coul
d not get a Transcript object for \'AGTAAA\'"}'
```

In [54]:
```
json.loads(r.text) # r.json()
```

Out[54]:
```
{'error': "Unable to parse HGVS notation 'AGTAAA:c.803T>C': Could
not get a Transcript object for 'AGTAAA'"}
```

In [ ]:
```

```

In [56]:
```
import requests, sys

url = "https://rest.ensembl.org/vep/human/hgvs"
headers={ "Content-Type" : "application/json", "Accept" : "applicat
ion/json"}

data={ "hgvs_notations" : ["AGT:c.803T>C", "9:g.22125504G>C" ] }

r = requests.post(url, headers=headers, data=json.dumps(data))
```

In [57]:
```
r.ok
```

Out[57]:
```
True
```

In [58]:
```
d = r.json()
```

In [60]:
```
len(d)
```

Out[60]:
```
2
```

In [66]:
```
d[0]['transcript_consequences'][0]['gene_symbol']
```

Out[66]:
```
'AGT'
```

In [67]:
```
for x in d:
    for y in x['transcript_consequences']:
        print (y['gene_symbol'])
```

```
AGT
AL512328.1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
CDKN2B-AS1
```

In [ ]:
```

```

In [62]:
```
d[1]
```

Out[62]:

```
{'id': '9:g.22125504G>C',
 'end': 22125504,
 'allele_string': 'G/C',
 'most_severe_consequence': 'intron_variant',
 'transcript_consequences': [{'consequence_terms': ['downstream_ge
ne_variant'],
   'biotype': 'lncRNA',
   'strand': 1,
   'impact': 'MODIFIER',
   'gene_symbol_source': 'HGNC',
   'transcript_id': 'ENST00000421632',
   'hgnc_id': 'HGNC:34341',
   'gene_id': 'ENSG00000240498',
   'variant_allele': 'C',
   'gene_symbol': 'CDKN2B-AS1',
   'distance': 4815},
  {'hgnc_id': 'HGNC:34341',
   'transcript_id': 'ENST00000422420',
   'distance': 4407,
   'gene_symbol': 'CDKN2B-AS1',
   'variant_allele': 'C',
   'gene_id': 'ENSG00000240498',
   'impact': 'MODIFIER',
   'strand': 1,
   'biotype': 'lncRNA',
   'consequence_terms': ['downstream_gene_variant'],
   'gene_symbol_source': 'HGNC'},
  {'transcript_id': 'ENST00000428597',
   'hgnc_id': 'HGNC:34341',
   'gene_symbol': 'CDKN2B-AS1',
   'distance': 4407,
   'gene_id': 'ENSG00000240498',
   'variant_allele': 'C',
   'impact': 'MODIFIER',
   'biotype': 'lncRNA',
   'consequence_terms': ['downstream_gene_variant'],
   'strand': 1,
   'gene_symbol_source': 'HGNC'},
  {'impact': 'MODIFIER',
   'biotype': 'lncRNA',
   'consequence_terms': ['downstream_gene_variant'],
   'strand': 1,
   'gene_symbol_source': 'HGNC',
   'transcript_id': 'ENST00000577551',
   'hgnc_id': 'HGNC:34341',
   'gene_symbol': 'CDKN2B-AS1',
   'distance': 4932,
   'gene_id': 'ENSG00000240498',
   'variant_allele': 'C'},
  {'consequence_terms': ['downstream_gene_variant'],
   'biotype': 'lncRNA',
   'strand': 1,
   'impact': 'MODIFIER',
   'gene_symbol_source': 'HGNC',
   'transcript_id': 'ENST00000580576',
   'hgnc_id': 'HGNC:34341',
   'gene_id': 'ENSG00000240498',
   'variant_allele': 'C',
   'gene_symbol': 'CDKN2B-AS1',
```

```
 'distance': 4867},
{'distance': 4932,
 'gene_symbol': 'CDKN2B-AS1',
 'variant_allele': 'C',
 'gene_id': 'ENSG00000240498',
 'hgnc_id': 'HGNC:34341',
 'transcript_id': 'ENST00000581051',
 'gene_symbol_source': 'HGNC',
 'impact': 'MODIFIER',
 'strand': 1,
 'consequence_terms': ['downstream_gene_variant'],
 'biotype': 'lncRNA'},
{'impact': 'MODIFIER',
 'strand': 1,
 'biotype': 'lncRNA',
 'consequence_terms': ['downstream_gene_variant'],
 'gene_symbol_source': 'HGNC',
 'hgnc_id': 'HGNC:34341',
 'transcript_id': 'ENST00000582072',
 'distance': 4932,
 'gene_symbol': 'CDKN2B-AS1',
 'variant_allele': 'C',
 'gene_id': 'ENSG00000240498'},
{'distance': 4932,
 'gene_symbol': 'CDKN2B-AS1',
 'variant_allele': 'C',
 'gene_id': 'ENSG00000240498',
 'hgnc_id': 'HGNC:34341',
 'transcript_id': 'ENST00000584020',
 'gene_symbol_source': 'HGNC',
 'impact': 'MODIFIER',
 'strand': 1,
 'biotype': 'lncRNA',
 'consequence_terms': ['downstream_gene_variant']},
{'transcript_id': 'ENST00000584637',
 'hgnc_id': 'HGNC:34341',
 'gene_id': 'ENSG00000240498',
 'variant_allele': 'C',
 'gene_symbol': 'CDKN2B-AS1',
 'distance': 4932,
 'biotype': 'lncRNA',
 'consequence_terms': ['downstream_gene_variant'],
 'strand': 1,
 'impact': 'MODIFIER',
 'gene_symbol_source': 'HGNC'},
{'hgnc_id': 'HGNC:34341',
 'transcript_id': 'ENST00000584816',
 'distance': 4932,
 'gene_symbol': 'CDKN2B-AS1',
 'variant_allele': 'C',
 'gene_id': 'ENSG00000240498',
 'impact': 'MODIFIER',
 'strand': 1,
 'consequence_terms': ['downstream_gene_variant'],
 'biotype': 'lncRNA',
 'gene_symbol_source': 'HGNC'},
{'gene_symbol_source': 'HGNC',
 'impact': 'MODIFIER',
 'consequence_terms': ['downstream_gene_variant'],
 'biotype': 'lncRNA',
```

```
   'strand': 1,
   'gene_symbol': 'CDKN2B-AS1',
   'distance': 4960,
   'gene_id': 'ENSG00000240498',
   'variant_allele': 'C',
   'transcript_id': 'ENST00000585267',
   'hgnc_id': 'HGNC:34341'},
  {'gene_symbol_source': 'HGNC',
   'biotype': 'lncRNA',
   'consequence_terms': ['downstream_gene_variant'],
   'strand': 1,
   'impact': 'MODIFIER',
   'gene_id': 'ENSG00000240498',
   'variant_allele': 'C',
   'gene_symbol': 'CDKN2B-AS1',
   'distance': 4791,
   'transcript_id': 'ENST00000643286',
   'hgnc_id': 'HGNC:34341'},
  {'transcript_id': 'ENST00000644233',
   'hgnc_id': 'HGNC:34341',
   'gene_id': 'ENSG00000240498',
   'variant_allele': 'C',
   'gene_symbol': 'CDKN2B-AS1',
   'distance': 4793,
   'biotype': 'lncRNA',
   'consequence_terms': ['downstream_gene_variant'],
   'strand': 1,
   'impact': 'MODIFIER',
   'gene_symbol_source': 'HGNC'},
  {'consequence_terms': ['downstream_gene_variant'],
   'biotype': 'lncRNA',
   'strand': 1,
   'impact': 'MODIFIER',
   'gene_symbol_source': 'HGNC',
   'transcript_id': 'ENST00000645223',
   'hgnc_id': 'HGNC:34341',
   'gene_id': 'ENSG00000240498',
   'variant_allele': 'C',
   'gene_symbol': 'CDKN2B-AS1',
   'distance': 4795},
  {'gene_symbol': 'CDKN2B-AS1',
   'variant_allele': 'C',
   'gene_id': 'ENSG00000240498',
   'hgnc_id': 'HGNC:34341',
   'transcript_id': 'ENST00000650946',
   'gene_symbol_source': 'HGNC',
   'impact': 'MODIFIER',
   'strand': 1,
   'consequence_terms': ['intron_variant', 'non_coding_transcript_
variant'],
   'biotype': 'lncRNA'},
  {'consequence_terms': ['downstream_gene_variant'],
   'biotype': 'lncRNA',
   'strand': 1,
   'impact': 'MODIFIER',
   'gene_symbol_source': 'HGNC',
   'transcript_id': 'ENST00000658981',
   'hgnc_id': 'HGNC:34341',
   'gene_id': 'ENSG00000240498',
   'variant_allele': 'C',
```

```
    'gene_symbol': 'CDKN2B-AS1',
    'distance': 4794}],
  'input': '9:g.22125504G>C',
  'colocated_variants': [{'var_synonyms': 'ClinVar::RCV001003460,VC
V000812642--PharmGKB::PA166157726',
    'minor_allele_freq': 0.4181,
    'strand': 1,
    'minor_allele': 'C',
    'seq_region_name': '9',
    'clin_sig_allele': 'C:risk_factor',
    'phenotype_or_disease': 1,
    'id': 'rs1333049',
    'pubmed': [21894447,
     18224312,
     22400124,
     22403240,
     22856518,
     24728607,
     22029572,
     21860704,
     18780302,
     29340220,
     27424552,
     26252781,
     17634449,
     18852197,
     18979498,
     19164808,
     19207022,
     19750184,
     19924713,
     19955471,
     20017983,
     20098575,
     20159871,
     20549515,
     20981302,
     21242481,
     21369780,
     21606135,
     21698238,
     21804106,
     22042884,
     22144573,
     22623978,
     26958643,
     26950853,
     28209224,
     20502693,
     19956433,
     21297524,
     22429504,
     22848412,
     26483964,
     22199011,
     19474294,
     19885677,
     24098343,
     18533027,
     22621687,
```

```
21149552,
29331485,
23963167,
27153677,
24676469,
18469204,
27892471,
23142796,
27386823,
26677855,
20435227,
25617895,
29143599,
18654002,
22295058,
28400043,
26999117,
21424681,
27015805,
19501493,
24573017,
28813480,
25717410,
26729200,
27249003,
26789557,
26982883,
28260796,
26775120,
30072947,
22322877,
30764545,
29309886,
29141072,
24906238,
26866580,
20175863,
30457165,
28138111,
19463184,
24926413,
27736948,
27004807,
22505696,
26559855,
19578366,
18362232,
19173706,
19214202,
20335276,
23870195,
23587283,
31115525,
18675980,
19135198,
24942486,
27721851,
18704761,
21971053,
21400687,
```

```
   26071660,
   27340317,
   30482443,
   21270277,
   21470412,
   2496741],
  'clin_sig': ['risk_factor'],
  'frequencies': {'C': {'eas': 0.5367,
    'eur': 0.4722,
    'sas': 0.4908,
    'afr': 0.2133,
    'amr': 0.4553}},
  'allele_string': 'G/C',
  'end': 22125504,
  'start': 22125504}],
 'seq_region_name': '9',
 'start': 22125504,
```

In [ ]:

Classes

In [70]:
```
a = [1,2,3]
```

In [71]:
```
a.append(4)
```

In [198]:
```
class Gene:

    reference = 'GRCh38' # Attribute (variable in a class)

    def __init__(self, name): # Initialization
        self.name = name

    def f(self,): # Method (function in a class)
        print (self.reference + ' ' + self.name)
        return 42

    def g(self,):
        self.f()

    def __len__(self,):
        return len(self.name)

    def __str__(self,):
        return 'Mitsos'

    def __iter__(self,):

        def mitsos():
            for x in range(10):
                yield x

        return mitsos()

    def __call__(self, ):
        return 42

    def __add__(self, b):
        return self.name + ' MITSOS ' + b.name

    @staticmethod  # <-- Decorator
    def check_if_ensembl_is_online():
        return 42
```

In [195]:
```
a = Gene('hba1')
```

In [196]:
```
a.f()
```

```
GRCh38 hba1
```

Out[196]:
```
42
```

In [197]:
```
#Gene.f()
```

```
---------------------------------------------------------------------
---------
TypeError                                 Traceback (most recent c
all last)
<ipython-input-197-f202ef4aff22> in <module>()
----> 1 Gene.f()

TypeError: f() missing 1 required positional argument: 'self'
```

In [190]:
```
a.check_if_ensembl_is_online()
```

Out[190]:
```
42
```

In [191]:
```
Gene.check_if_ensembl_is_online()
```

Out[191]:
```
42
```

In [ ]:
```

```

In [176]:
```
class Cell(Gene):

    def t(self,):
        print ('hello')

class Tissue(Cell):
    pass
```

In [159]:
```
c = Cell('hba1')
```

In [160]:
```
c.f()
```

```
GRCh38
```

Out[160]:
```
42
```

In [161]:
```
c.t()
```

```
hello
```

In [ ]:
```

```

In [ ]:
```

```

In [152]:
```
a = Gene()
```

```
---------------------------------------------------------------------
---------
TypeError                                 Traceback (most recent c
all last)
<ipython-input-152-15d23d67f116> in <module>()
----> 1 a = Gene()

TypeError: __init__() missing 1 required positional argument: 'nam
e'
```

In [ ]:
```

```

In [154]:
```
a = Gene('hba1')
```

In [155]:
```
a()
```

Out[155]:
```
42
```

In [ ]:
```

```

In [122]:
```
len(a)
```

Out[122]:
```
4
```

In [134]:
```
a = Gene('aaa')
b = Gene('ccc')

a+b
```

Out[134]:
```
'aaa MITSOS ccc'
```

In [130]:
```
print(a)
```

```
Mitsos
```

In [ ]:

In [100]:
```
print (a)
```

```
Mitsos
```

In [114]:
```
for x in a:
    print (x)
```

```
0
1
2
3
4
5
6
7
8
9
```

In [ ]:

In [97]:
```
len(a)
```

Out[97]:
```
100
```

In [ ]:
```
len()
```

In [87]:
```
a.f()
```

```
GRCh38
```

Out[87]:
```
42
```

In [91]:
```
a.g()
```

```
GRCh38
```

In [131]:
```
a = [4,5,6]
```

In [132]:
```
len(a)
```

Out[132]:
```
3
```

In [133]:
```
a.__len__()
```

Out[133]:
```
3
```

In [136]:
```
dir(a)
```

Out[136]:
```
['__add__',
 '__class__',
 '__delattr__',
 '__dict__',
 '__dir__',
 '__doc__',
 '__eq__',
 '__format__',
 '__ge__',
 '__getattribute__',
 '__gt__',
 '__hash__',
 '__init__',
 '__init_subclass__',
 '__iter__',
 '__le__',
 '__len__',
 '__lt__',
 '__module__',
 '__ne__',
 '__new__',
 '__reduce__',
 '__reduce_ex__',
 '__repr__',
 '__setattr__',
 '__sizeof__',
 '__str__',
 '__subclasshook__',
 '__weakref__',
 'f',
 'g',
 'name',
 'reference']
```

In [137]:
```
dir(53)
```

Out[137]:
```
['__abs__',
 '__add__',
 '__and__',
 '__bool__',
 '__ceil__',
 '__class__',
 '__delattr__',
 '__dir__',
 '__divmod__',
 '__doc__',
 '__eq__',
 '__float__',
 '__floor__',
 '__floordiv__',
 '__format__',
 '__ge__',
 '__getattribute__',
 '__getnewargs__',
 '__gt__',
 '__hash__',
 '__index__',
 '__init__',
 '__init_subclass__',
 '__int__',
 '__invert__',
 '__le__',
 '__lshift__',
 '__lt__',
 '__mod__',
 '__mul__',
 '__ne__',
 '__neg__',
 '__new__',
 '__or__',
 '__pos__',
 '__pow__',
 '__radd__',
 '__rand__',
 '__rdivmod__',
 '__reduce__',
 '__reduce_ex__',
 '__repr__',
 '__rfloordiv__',
 '__rlshift__',
 '__rmod__',
 '__rmul__',
 '__ror__',
 '__round__',
 '__rpow__',
 '__rrshift__',
 '__rshift__',
 '__rsub__',
 '__rtruediv__',
 '__rxor__',
 '__setattr__',
 '__sizeof__',
 '__str__',
 '__sub__',
 '__subclasshook__',
```

```
 '__truediv__',
 '__trunc__',
 '__xor__',
 'bit_length',
 'conjugate',
 'denominator',
 'from_bytes',
 'imag',
 'numerator',
 'real',
 'to_bytes']
```

In [138]:
```
53.__add__(20)
```

```
  File "<ipython-input-138-4d35a4f9b6d0>", line 1
    53.__add__(20)
              ^
SyntaxError: invalid syntax
```

In [143]:
```
getattr(53, '__add__')(20)
```

Out[143]: 73

In [144]:
```
53..__add__(10)
```

Out[144]: 63.0

In [145]:
```
(53).__add__(10)
```

Out[145]: 63

In [147]:
```
def f(x):
    return x+42
```

In [149]:
```
f.__call__(1)
```

Out[149]: 43

In [164]:
```
f(1)
```

Out[164]: 43

pass

In [169]:
```
x = 10

if x == 10:
    #print ('is 10')
    pass
else:
    print ('It is not 10')
```

In [172]:
```
x=10
y=51


if x==10 and y==50:
    pass
else:
    print ('hello')
```

```
hello
```

In [174]:
```
for x in range(1000):
    pass
    #print(x)

print ('hello')
```

```
hello
```

In [175]:
```
pass
```

In [ ]:
```
! pip install biopython
```

In [274]:
```
from Bio import Entrez

Entrez.email = 'your_email@provider.com'

def get_pmid(pmid):

    pmids = [pmid]
    handle = Entrez.efetch(db="pubmed", id=','.join(map(str, pmid
s)),
                           rettype="xml", retmode="text")
    records = Entrez.read(handle)
    abstracts = [pubmed_article['MedlineCitation']['Article']['Abst
ract']['AbstractText'][0]
                 for pubmed_article in records['PubmedArticle']]


    abstract_dict = dict(zip(pmids, abstracts))
    return abstract_dict
```

In [272]:
```
abstract_dict
```

Out[272]:
```
{17284678: 'Eimeria tenella is an intracellular protozoan parasite
that infects the intestinal tracts of domestic fowl and causes coc
cidiosis, a serious and sometimes lethal enteritis. Eimeria falls
in the same phylum (Apicomplexa) as several human and animal paras
ites such as Cryptosporidium, Toxoplasma, and the malaria parasit
e, Plasmodium. Here we report the sequencing and analysis of the f
irst chromosome of E. tenella, a chromosome believed to carry loci
associated with drug resistance and known to differ between virule
nt and attenuated strains of the parasite. The chromosome--which a
ppears to be representative of the genome--is gene-dense and rich
in simple-sequence repeats, many of which appear to give rise to r
epetitive amino acid tracts in the predicted proteins. Most striki
ng is the segmentation of the chromosome into repeat-rich regions
peppered with transposon-like elements and telomere-like repeats,
alternating with repeat-free regions. Predicted genes differ in ch
aracter between the two types of segment, and the repeat-rich regi
ons appear to be associated with strain-to-strain variation.',
 9997: 'Electron paramagnetic resonance and magnetic susceptibilit
y studies of Chromatium flavocytochrome C552 and its diheme flavin
-free subunit at temperatures below 45 degrees K are reported. The
results show that in the intact protein and the subunit the two lo
w-spin (S = 1/2) heme irons are distinguishable, giving rise to se
parate EPR signals. In the intact protein only, one of the heme ir
ons exists in two different low spin environments in the pH range
5.5 to 10.5, while the other remains in a constant environment. Fa
ctors influencing the variable heme iron environment also influenc
e flavin reactivity, indicating the existence of a mechanism for h
eme-flavin interaction.'}
```

In [324]:
```
import requests

class Mutation():
    def __init__(self, name):
        self.name = name
        self.get_info_from_ensmble()

    def get_info_from_ensmble(self, ):
        server = "https://rest.ensembl.org/vep/human/hgvs/{}?".form
at(self.name)
        r = requests.get(server, headers={ "Content-Type" : "applic
ation/json"})
        if not r.ok:
            raise Exception('oops')
        self.data = r.json()

    def get_consequenmce(self):
        return [x['most_severe_consequence'] for x in self.data]

    def __len__(self,):
        #return m.data[0]['end'] - m.data[0]['start'] + 1
        return len(self.data[0]['transcript_consequences'][0]['vari
ant_allele'])


    def __iter__(self,):

        def g():
            for x in self.data[0]['transcript_consequences']:
                yield x

        return g()

    def __str__(self,):
        return self.name

    def get_pubmed_ids(self,):
        self.pubmed_ids = []
        for x in self.data:
            for y in x['colocated_variants']:
                if 'pubmed' in y:
                    self.pubmed_ids.extend(y['pubmed'])

    def get_abstracts(self,):
        self.get_pubmed_ids()
        return Mutation.get_pmid(self.pubmed_ids)


    @staticmethod
    def get_pmid(pmids):
        '''
        pmids : list of pmids
        '''

        handle = Entrez.efetch(db="pubmed", id=','.join(map(str, pm
ids)),
                               rettype="xml", retmode="text")
        records = Entrez.read(handle)
        ret = {}
```

```
        for pubmed_article in records['PubmedArticle']:

            if not 'MedlineCitation' in pubmed_article:
                continue
            MedlineCitation = pubmed_article['MedlineCitation']

            pmid = str(MedlineCitation['PMID'])

            if not 'Article' in MedlineCitation:
                continue
            Article = MedlineCitation['Article']

            if not 'Abstract' in Article:
                continue
            Abstract = Article['Abstract']

            if not 'AbstractText' in Abstract:
                continue

            AbstractText = Abstract['AbstractText']
            if not AbstractText:
                continue

            ret[pmid] = AbstractText[0]

        return ret
```

In [325]:
```
m = Mutation('AGT:c.803T>C')
abstracts = m.get_abstracts()
```

In [327]:
```
abstracts.keys()
```

Out[327]:
```
dict_keys(['19131662', '19263529', '19330901', '19559392', '219199
68', '27616475', '25741868', '21467728', '1394429', '7649545', '78
83995', '8513325', '9259580', '9421481', '9831339', '16059745', '1
8069999', '18248681', '18279468', '18513389', '18603647', '1863718
8', '18653189', '18698212', '18953568', '19108684', '19770777', '1
9932491', '20029521', '20047954', '20061926', '20185782', '2048628
2', '20570668', '20577119', '20592051', '20811292', '20981351', '2
1056700', '21058046', '21127830', '21146954', '21261619', '2130499
9', '21306748', '21438754', '21444836', '21515823', '21533139', '2
1540342', '21573014', '21681796', '21894447', '21988197', '2209945
8', '22100073', '22531885', '22569109', '22817530', '22858200', '2
3021345', '23036011', '23132613', '23133444', '23205182', '2325129
6', '23287839', '23333443', '23354977', '23497168', '23681449', '2
3716723', '24452035', '24622918', '24722536', '24737640', '2547435
6', '25512783', '25683681', '25723521', '26102248', '26283679', '2
6318936', '26335431', '26509357', '26588355', '26621708', '2662748
0', '26818744', '26819062', '26824906', '27068935', '27274104', '2
7342049', '27348238', '27380726', '27480094', '27584680', '2794066
2', '28361007', '28488548', '28605058', '28666769', '28690685', '2
8770234', '28828324', '28881807', '28903744', '29057680', '2952098
4', '2957843'])
```

In [328]:
```
abstracts['26627480']
```

Out[328]:
```
StringElement('The risk of an individual to develop an acute kidne
y injury (AKI), or its severity, cannot be reliably predicted by c
ommon clinical risk factors. Whether genetic risk factors have an
explanatory role poses an interesting question, however. Thus, we
conducted a systematic literature review regarding genetic predisp
osition to AKI or outcome of AKI patients.', attributes={'Label':
'BACKGROUND', 'NlmCategory': 'BACKGROUND'})
```

In [329]:
```
m.pubmed_ids
```

Out[329]:
```
[19131662,
 19263529,
 19330901,
 19559392,
 21919968,
 27616475,
 25741868,
 21467728,
 1394429,
 7649545,
 7883995,
 8348146,
 8513325,
 8518804,
 9259580,
 9421481,
 9831339,
 16059745,
 18069999,
 18248681,
 18279468,
 18513389,
 18603647,
 18637188,
 18653189,
 18698212,
 18953568,
 19108684,
 19770777,
 19932491,
 20029521,
 20047954,
 20061926,
 20185782,
 20486282,
 20570668,
 20577119,
 20592051,
 20811292,
 20981351,
 21056700,
 21058046,
 21127830,
 21146954,
 21261619,
 21304999,
 21306748,
 21438754,
 21444836,
 21515823,
 21533139,
 21540342,
 21573014,
 21681796,
 21894447,
 21988197,
 22099458,
 22100073,
 22531885,
```

```
22569109,
22817530,
22858200,
23021345,
23036011,
23132613,
23133444,
23205182,
23251296,
23287839,
23333443,
23354977,
23497168,
23681449,
23716723,
24452035,
24622918,
24722536,
24737640,
25474356,
25512783,
25683681,
25723521,
26102248,
26283679,
26318936,
26335431,
26509357,
26588355,
26621708,
26627480,
26818744,
26819062,
26824906,
26933222,
27068935,
27274104,
27342049,
27348238,
27380726,
27454254,
27480094,
27584680,
27940662,
28361007,
28488548,
28605058,
28666769,
28690685,
28770234,
28828324,
28881807,
28903744,
29057680,
29520984,
29578431
```

In [330]:
```
m.get_consequenmce()
```

Out[330]:
```
['missense_variant']
```

In [331]:
```
len(m)
```

Out[331]:
```
1
```

In [332]:
```
print (json.dumps(m.data[0], indent=4))
```

```
{
    "end": 230710048,
    "assembly_name": "GRCh38",
    "strand": -1,
    "colocated_variants": [
        {
            "end": 230710048,
            "allele_string": "HGMD_MUTATION",
            "start": 230710048,
            "seq_region_name": "1",
            "strand": 1,
            "phenotype_or_disease": 1,
            "id": "CM920010"
        },
        {
            "phenotype_or_disease": 1,
            "clin_sig_allele": "G:benign;G:risk_factor",
            "id": "rs699",
            "pubmed": [
                19131662,
                19263529,
                19330901,
                19559392,
                21919968,
                27616475,
                25741868,
                21467728,
                1394429,
                7649545,
                7883995,
                8348146,
                8513325,
                8518804,
                9259580,
                9421481,
                9831339,
                16059745,
                18069999,
                18248681,
                18279468,
                18513389,
                18603647,
                18637188,
                18653189,
                18698212,
                18953568,
                19108684,
                19770777,
                19932491,
                20029521,
                20047954,
                20061926,
                20185782,
                20486282,
                20570668,
                20577119,
                20592051,
                20811292,
                20981351,
```

```
21056700,
21058046,
21127830,
21146954,
21261619,
21304999,
21306748,
21438754,
21444836,
21515823,
21533139,
21540342,
21573014,
21681796,
21894447,
21988197,
22099458,
22100073,
22531885,
22569109,
22817530,
22858200,
23021345,
23036011,
23132613,
23133444,
23205182,
23251296,
23287839,
23333443,
23354977,
23497168,
23681449,
23716723,
24452035,
24622918,
24722536,
24737640,
25474356,
25512783,
25683681,
25723521,
26102248,
26283679,
26318936,
26335431,
26509357,
26588355,
26621708,
26627480,
26818744,
26819062,
26824906,
26933222,
27068935,
27274104,
27342049,
27348238,
27380726,
27454254,
```

```
                    27480094,
                    27584680,
                    27940662,
                    28361007,
                    28488548,
                    28605058,
                    28666769,
                    28690685,
                    28770234,
                    28828324,
                    28881807,
                    28903744,
                    29057680,
                    29520984,
                    2957843
                ],
                "var_synonyms": "ClinVar::RCV000835695,RCV000405686,VC
V000018068,RCV000019691,RCV000019692,RCV000019693,RCV000242838--OM
IM::106150.0001--PharmGKB::PA166153539--Uniprot::VAR_007096",
                "clin_sig": [
                    "benign",
                    "risk_factor"
                ],
                "start": 230710048,
                "seq_region_name": "1",
                "frequencies": {
                    "C": {
                        "gnomad_amr": 0.7196,
                        "gnomad_afr": 0.8451,
                        "gnomad": 0.5481,
                        "gnomad_asj": 0.44,
                        "gnomad_oth": 0.5002,
                        "ea": 0.4258,
                        "eas": 0.8532,
                        "gnomad_nfe": 0.4197,
                        "gnomad_fin": 0.441,
                        "gnomad_sas": 0.6202,
                        "sas": 0.636,
                        "eur": 0.4115,
                        "afr": 0.9032,
                        "amr": 0.6354,
                        "gnomad_eas": 0.8388,
                        "aa": 0.8268
                    }
                },
                "strand": 1,
                "minor_allele_freq": 0.2949,
                "end": 230710048,
                "minor_allele": "A",
                "allele_string": "A/G"
            }
        ],
        "seq_region_name": "1",
        "id": "AGT:c.803T>C",
        "transcript_consequences": [
            {
                "sift_prediction": "tolerated",
                "amino_acids": "M/T",
                "polyphen_prediction": "benign",
                "variant_allele": "C",
```

```
            "sift_score": 1,
            "biotype": "protein_coding",
            "impact": "MODERATE",
            "gene_id": "ENSG00000135744",
            "hgnc_id": "HGNC:333",
            "transcript_id": "ENST00000366667",
            "cdna_start": 843,
            "protein_end": 268,
            "strand": -1,
            "gene_symbol_source": "HGNC",
            "cds_start": 803,
            "polyphen_score": 0,
            "cdna_end": 843,
            "gene_symbol": "AGT",
            "codons": "aTg/aCg",
            "protein_start": 268,
            "cds_end": 803,
            "consequence_terms": [
                "missense_variant"
            ]
        },
        {
            "gene_id": "ENSG00000244137",
            "gene_symbol": "AL512328.1",
            "distance": 650,
            "consequence_terms": [
                "downstream_gene_variant"
            ],
            "transcript_id": "ENST00000412344",
            "variant_allele": "C",
            "strand": -1,
            "gene_symbol_source": "Clone_based_ensembl_gene",
            "impact": "MODIFIER",
            "biotype": "lncRNA"
        }
    ],
    "allele_string": "T/C",
    "input": "AGT:c.803T>C",
    "start": 230710048,
    "most_severe_consequence": "missense_variant"
```

In [333]:
```
m.data[0]['colocated_variants'][1]['pubmed']
```

Out[333]:
```
[19131662,
 19263529,
 19330901,
 19559392,
 21919968,
 27616475,
 25741868,
 21467728,
 1394429,
 7649545,
 7883995,
 8348146,
 8513325,
 8518804,
 9259580,
 9421481,
 9831339,
 16059745,
 18069999,
 18248681,
 18279468,
 18513389,
 18603647,
 18637188,
 18653189,
 18698212,
 18953568,
 19108684,
 19770777,
 19932491,
 20029521,
 20047954,
 20061926,
 20185782,
 20486282,
 20570668,
 20577119,
 20592051,
 20811292,
 20981351,
 21056700,
 21058046,
 21127830,
 21146954,
 21261619,
 21304999,
 21306748,
 21438754,
 21444836,
 21515823,
 21533139,
 21540342,
 21573014,
 21681796,
 21894447,
 21988197,
 22099458,
 22100073,
 22531885,
```

```
22569109,
22817530,
22858200,
23021345,
23036011,
23132613,
23133444,
23205182,
23251296,
23287839,
23333443,
23354977,
23497168,
23681449,
23716723,
24452035,
24622918,
24722536,
24737640,
25474356,
25512783,
25683681,
25723521,
26102248,
26283679,
26318936,
26335431,
26509357,
26588355,
26621708,
26627480,
26818744,
26819062,
26824906,
26933222,
27068935,
27274104,
27342049,
27348238,
27380726,
27454254,
27480094,
27584680,
27940662,
28361007,
28488548,
28605058,
28666769,
28690685,
28770234,
28828324,
28881807,
28903744,
29057680,
29520984,
29578431
```

In [ ]:

In [335]:
```
m2 = Mutation('ENST00000003084:c.1431_1432insTTC')
```

In [336]:
```
print (m2)
```

```
ENST00000003084:c.1431_1432insTTC
```

In [337]:
```
len(m2)
```

Out[337]:
```
3
```

In [338]:
```
m2.data[0]['transcript_consequences'][3]
```

Out[338]:
```
{'consequence_terms': ['upstream_gene_variant'],
 'gene_symbol': 'CFTR',
 'gene_symbol_source': 'HGNC',
 'biotype': 'processed_transcript',
 'gene_id': 'ENSG00000001626',
 'variant_allele': 'TTC',
 'transcript_id': 'ENST00000472848',
 'hgnc_id': 'HGNC:1884',
 'distance': 135,
 'strand': 1,
 'impact': 'MODIFIER'}
```

In [268]:
```
for x in m2:
    print (x['transcript_id'])
    #print (x)
```

```
ENST00000003084
ENST00000426809
ENST00000441019
ENST00000472848
ENST00000647978
ENST00000648260
ENST00000649406
ENST00000649781
```

In [ ]:
```

```