

rf_feature_selection

May 25, 2020

```
[4]: import numpy as np
import pandas as pd
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.feature_selection import SelectFromModel
from sklearn.metrics import accuracy_score
```

```
[7]: data = pd.read_csv("data/combined_expression.csv")
```

```
[90]: data.head()
```

```
[90]:
```

	CELL_LINE_NAME	classification	TSPAN6	TNMD	DPM1	SCYL3	\
0	1240121	5	6.419526	3.182094	9.320548	3.759654	
1	1240122	6	7.646494	2.626819	10.153853	3.564755	
2	1240123	5	8.319417	3.111183	9.643558	4.757258	
3	1240124	1	9.006994	3.028173	9.686700	4.280504	
4	1240127	6	7.985676	2.694729	10.676134	4.159685	

	C1orf112	FGR	CFH	FUCA2	...	COL15A1	C6orf10	TMEM225	\
0	3.802619	3.215753	4.698729	7.873672	...	3.245454	2.953508	3.543429	
1	3.942749	3.290760	3.551675	8.252413	...	2.786709	3.077382	3.728232	
2	3.919757	3.602185	3.329644	9.076950	...	3.459089	3.085394	3.462811	
3	3.147646	3.188881	3.293807	8.678790	...	2.835403	2.960303	3.415083	
4	3.804637	3.481942	3.111261	7.555407	...	2.896523	2.849899	3.480114	

	NOTCH4	PBX2	AGER	RNF5	AGPAT1	DFNB59	PRRT1
0	3.352022	4.672310	3.641128	3.135310	3.737072	3.450927	3.168800
1	3.208882	4.586840	3.395654	3.586800	3.519128	3.115323	3.051645
2	3.339030	4.614897	3.395845	3.419193	3.971646	3.729310	3.320022
3	3.290171	4.770123	3.400821	3.383734	3.798107	2.822404	3.297547
4	3.226128	5.832710	3.612179	3.347095	4.457963	5.198524	4.553586

[5 rows x 16383 columns]

```
[50]: X = data.drop(columns=['CELL_LINE_NAME', 'classification'])
y = df['classification']
feat_labels = list(X.columns)
```

```
[54]: # 20% test, 80% train
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
↳random_state=0)

[55]: # create and train the classifier
clf = RandomForestClassifier(n_estimators=X.shape[1], random_state=0, n_jobs=-1)
clf.fit(X_train, y_train)

[55]: RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight=None,
                             criterion='gini', max_depth=None, max_features='auto',
                             max_leaf_nodes=None, max_samples=None,
                             min_impurity_decrease=0.0, min_impurity_split=None,
                             min_samples_leaf=1, min_samples_split=2,
                             min_weight_fraction_leaf=0.0, n_estimators=16381,
                             n_jobs=-1, oob_score=False, random_state=0, verbose=0,
                             warm_start=False)

[79]: # associating each feature with its relative importance
feat_importances = list(zip(feat_labels, clf.feature_importances_))

[84]: # sorting the feature importances
def sort_tuple(tup):
    tup.sort(key = lambda x: x[1], reverse=True)
    return tup
sort_tuple(feat_importances)

[84]: [('CTTN', 0.0008402505637710411),
      ('TM4SF1', 0.0008176207269100246),
      ('NCKAP1', 0.0008161439524089536),
      ('GMFG', 0.0008012264232279094),
      ('SLC35D2', 0.0007651422978405773),
      ('NCKAP1L', 0.0007616336177860741),
      ('ARHGAP15', 0.0007256156159650742),
      ('SDC4', 0.0007216151979827462),
      ('RBF0X2', 0.0007118744433809878),
      ('NQ01', 0.0007071261685252473),
      ('RASAL3', 0.0006983618567021172),
      ('FAM78A', 0.000665932873791482),
      ('FAM177A1', 0.0006599917249745302),
      ('CD53', 0.0006533147429125048),
      ('LCP1', 0.0006511143494124967),
      ('P2RY8', 0.0006489220312550476),
      ('CD276', 0.0006474840607009987),
      ('GNA11', 0.0006327049398194747),
      ('TJP1', 0.0006291339175535226),
      ('ARHGAP30', 0.0006280449824593603),
      ('X06.Sep', 0.0006267780621863761),
```

('IKZF1', 0.0006136084864062751),
 ('PPIC', 0.0006081513987886942),
 ('PIK3CG', 0.000603897159381982),
 ('PRSS23', 0.0006028137245271975),
 ('CORO1A', 0.000587971416072799),
 ('PLK2', 0.0005751974467096185),
 ('SH3D19', 0.0005679992529686441),
 ('ACAP1', 0.0005541411451649759),
 ('LMNA', 0.0005535747667629884),
 ('TNFRSF12A', 0.0005507206907177382),
 ('RCSL1', 0.0005506130358150885),
 ('SASH3', 0.000549740742338346),
 ('MGAT4B', 0.0005448828924354285),
 ('CD48', 0.0005409328431463166),
 ('TRAF3IP3', 0.0005404919103149723),
 ('PTPRC', 0.0005270611532532388),
 ('WASL', 0.0005264283099742402),
 ('PARVA', 0.0005007259748937981),
 ('FLI1', 0.0004993218371614253),
 ('SOWAHC', 0.0004964777711131362),
 ('ERRFI1', 0.0004953129131293057),
 ('YAP1', 0.0004883442564452064),
 ('FAM129B', 0.00048535463575138493),
 ('MYO1C', 0.0004814824059117838),
 ('DOCK2', 0.00048073905377174605),
 ('KBTBD8', 0.00047734086233041414),
 ('MYOF', 0.00047726224467183384),
 ('PERP', 0.0004769027049376092),
 ('CYR61', 0.0004749456362721449),
 ('ARHGEF6', 0.0004733197168375773),
 ('SAFB', 0.00047319910713331307),
 ('MET', 0.0004703250852673371),
 ('GNG12', 0.00046771807435088325),
 ('PTPN7', 0.000467513281319396),
 ('TNFAIP8L2', 0.0004641049310355502),
 ('LCP2', 0.00046363103631413),
 ('LRMP', 0.0004527177233925971),
 ('COQ3', 0.0004428216424613369),
 ('EPHA2', 0.00044063011529793425),
 ('HMHA1', 0.0004395531577353068),
 ('ANKLE1', 0.00043903772619672686),
 ('SPATS2L', 0.00043084179670269975),
 ('ITGAV', 0.00043041086528182877),
 ('C16orf54', 0.0004296015259324859),
 ('INPP5D', 0.000425127578591063),
 ('BCL2L1', 0.0004235029930197001),
 ('ARHGAP9', 0.00042042821024834774),

('MYO1G', 0.0004184342088036042),
 ('MANBAL', 0.0004173927714132732),
 ('HVCN1', 0.00041671839432217875),
 ('MFNG', 0.0004148230863140541),
 ('STAP1', 0.0004116164347680046),
 ('ITGAL', 0.0004056383029696845),
 ('RHOH', 0.0004036211807803685),
 ('SLC38A5', 0.0004033765803546492),
 ('SQSTM1', 0.0004012904922702691),
 ('RASSF5', 0.0004002744126567501),
 ('BCAR1', 0.0003979695742384693),
 ('KLHL6', 0.0003949992412481647),
 ('DIRC2', 0.0003920978465878302),
 ('RNF138', 0.00039179036631381033),
 ('ADAM9', 0.0003888068570346445),
 ('ANXA2', 0.00038531921311951104),
 ('IL21R', 0.0003762675058406869),
 ('IL13RA1', 0.0003757052915969676),
 ('CSK', 0.0003733705021549535),
 ('CTNNA1', 0.00037054098583364403),
 ('TWF1', 0.00036971445999238277),
 ('PLEK', 0.00036863229059823566),
 ('S100A10', 0.00036607276406373215),
 ('SNX7', 0.00036178533588237675),
 ('HCLS1', 0.0003608904137497028),
 ('EVI2B', 0.0003603204374624112),
 ('NLRC3', 0.0003585949567699461),
 ('RASGRP2', 0.00035831655717147954),
 ('LYL1', 0.0003582219374179714),
 ('CDC42BPB', 0.00035668820623767984),
 ('CBY1', 0.00035578487247871514),
 ('TTC8', 0.0003488515319651457),
 ('TNFAIP1', 0.0003471992964513038),
 ('SEPLG', 0.00034403712759678454),
 ('HSH2D', 0.0003435722075642829),
 ('S1PR4', 0.0003403291020389781),
 ('DOCK8', 0.0003356841322843797),
 ('MOB3A', 0.00033333381826355533),
 ('MICALL2', 0.00033176073757997866),
 ('ELF1', 0.0003316207605568021),
 ('CFL2', 0.0003302003447002835),
 ('TMEM106B', 0.00032947864465589227),
 ('LAX1', 0.0003292754746071671),
 ('IGLL1', 0.0003280424591468571),
 ('LAMB2', 0.00032765609016409706),
 ('CSNK1G2', 0.0003261394913018154),
 ('N4BP2L1', 0.00031494874854769676),

('KDELR3', 0.0003133159914217308),
 ('PON2', 0.00031322070212235533),
 ('LAPTM5', 0.00031045941123435236),
 ('MDM4', 0.0003083809230606501),
 ('ZFHX3', 0.00030829397209082885),
 ('HMGB2', 0.00030717719112093186),
 ('ARHGAP29', 0.00030668208746492014),
 ('KIF3B', 0.0003009676573106955),
 ('CD1A', 0.00029973305697079765),
 ('KIAA0922', 0.0002984143270009911),
 ('ELM01', 0.0002970425676954874),
 ('ICAM2', 0.00029575859346039413),
 ('MAP4K1', 0.0002937573314171688),
 ('CD84', 0.00029201450945872694),
 ('ELOVL1', 0.0002906525637650435),
 ('FAF1', 0.0002868051910482566),
 ('PTK2', 0.00028519689175465117),
 ('NR2F2', 0.000284750670111472),
 ('STARD9', 0.0002843843120909784),
 ('CXorf21', 0.0002842730124881648),
 ('CD52', 0.0002822377803008001),
 ('MZB1', 0.00028214426984977633),
 ('CTSA', 0.00028056017667879416),
 ('ATP9A', 0.00028012986262041125),
 ('IL2RG', 0.00027972595745420195),
 ('LIMD2', 0.0002794762908177337),
 ('TSEN34', 0.000278529569905827),
 ('PLS3', 0.0002769992056235048),
 ('CCNB1IP1', 0.00027368659876399534),
 ('CDC14B', 0.0002716224515711235),
 ('FMNL1', 0.000270964383719717),
 ('PARD6B', 0.000269767704117897),
 ('ITGB5', 0.0002688370621545715),
 ('MCPH1', 0.00026866533185906667),
 ('TESPA1', 0.00026714190515679103),
 ('CHMP3', 0.00026596064284309194),
 ('KDM2B', 0.00026474504125478126),
 ('LAIR1', 0.000264556746658893),
 ('TTBK2', 0.0002637840381618141),
 ('SDSL', 0.0002631983218702143),
 ('SH2D3C', 0.0002628784073222012),
 ('CMTM4', 0.000262589356921995),
 ('TBC1D10C', 0.00026194524678010034),
 ('TGOLN2', 0.0002610293799753902),
 ('ITGA4', 0.00026091305870607476),
 ('BCL9L', 0.000260210030219973),
 ('TMTC3', 0.0002599352565362309),

('PPP1R16B', 0.00025955487587991553),
 ('DHX30', 0.00025602158139490904),
 ('CEACAM21', 0.00025462004288654447),
 ('APLP2', 0.0002540918936184018),
 ('GRAMD3', 0.00025366618108632816),
 ('FERMT3', 0.0002521600844905618),
 ('CD37', 0.00025067814438706795),
 ('CTTNBP2NL', 0.0002500648993648106),
 ('TAGAP', 0.00024921799653476284),
 ('ITGA2', 0.0002483817361838878),
 ('PTPN22', 0.000247594270335144),
 ('CAP2', 0.00024722698815196127),
 ('SAMSN1', 0.0002471485318457088),
 ('KRT8', 0.00024673321152957127),
 ('ATG4C', 0.00024634354652292907),
 ('TSPAN6', 0.0002458290842403086),
 ('NOL3', 0.0002454857619521968),
 ('RIOK1', 0.00024351210223366108),
 ('PXDC1', 0.00024297527262775545),
 ('LGALS3', 0.00024217837177068795),
 ('ANKRD46', 0.00024082308624484943),
 ('PTGR1', 0.0002396073891338345),
 ('ILF3', 0.00023957724164478297),
 ('BTK', 0.00023775931838087642),
 ('CCDC84', 0.00023580948482479787),
 ('ASAP2', 0.0002333697611237202),
 ('SLC35F5', 0.00023203761017588176),
 ('BAG3', 0.0002307127672547607),
 ('PTPN6', 0.00023048953256304467),
 ('PPAP2C', 0.0002295656508616471),
 ('TOM1L1', 0.00022934235977538188),
 ('ZNF445', 0.00022875522681496777),
 ('PRKCB', 0.00022825162473477463),
 ('KIAA1598', 0.0002273926976293211),
 ('S100A13', 0.0002268963809517779),
 ('ARHGDIB', 0.00022536344132541742),
 ('BCAR3', 0.00022535871746228573),
 ('IL16', 0.0002249548169137202),
 ('STOML2', 0.00022495296400269128),
 ('FBX07', 0.00022411064219555164),
 ('CYFIP2', 0.00022363364485987147),
 ('GFI1', 0.0002233428058051727),
 ('NCF4', 0.00022325332948025466),
 ('KIAA1191', 0.00022275234952662666),
 ('CECR1', 0.00022194858594710995),
 ('FUCA2', 0.00022167488867452255),
 ('CSF2RB', 0.00022072730282576716),

('SF3A3', 0.0002203492260810328),
 ('FASTKD1', 0.00021958925646034568),
 ('WASF1', 0.00021940332993063843),
 ('CAPN2', 0.0002190102423989505),
 ('ASPH', 0.00021872654880711011),
 ('EMP1', 0.00021783070249915652),
 ('TBC1D4', 0.00021717070124975372),
 ('ZRANB2', 0.00021690343266851113),
 ('P2RY14', 0.00021623205453408126),
 ('HMGXB4', 0.0002160950500054304),
 ('SBDS', 0.0002156218822426089),
 ('ITGA3', 0.00021319337698795233),
 ('QRSL1', 0.00021219898867662134),
 ('ERC2', 0.00021176917078594314),
 ('DAZAP1', 0.00021170561760114128),
 ('HAUS6', 0.00021165528678355696),
 ('NTN4', 0.00021162356582409423),
 ('CCDC69', 0.00021055146075571866),
 ('NKG7', 0.00021029446317473162),
 ('FAM114A1', 0.0002091639179633095),
 ('PTPN21', 0.00020909356856884607),
 ('DCBLD2', 0.0002089172811284392),
 ('KLHL8', 0.00020848671485390802),
 ('HSF2', 0.0002082635666804067),
 ('TGFB2', 0.00020648151578749703),
 ('S100A6', 0.00020565157963978704),
 ('CXorf65', 0.00020527224052592578),
 ('ACVR2A', 0.00020389761994239676),
 ('PTTG1IP', 0.00020361832175373016),
 ('SMARCA4', 0.00020350200819220163),
 ('KHDRBS1', 0.00020297799201286824),
 ('RIN2', 0.00020273761499660755),
 ('BCL2', 0.00020224184189811264),
 ('LAPTM4A', 0.00020179165173845465),
 ('MORN2', 0.0002013901661450819),
 ('RAC1', 0.00020110826244605596),
 ('TPGS2', 0.00020110401115343583),
 ('BBC3', 0.00019958785477424394),
 ('POLRMT', 0.00019896773636197967),
 ('NCEH1', 0.00019892546065683776),
 ('WDR27', 0.000198871766033847),
 ('SLA', 0.0001984492228466343),
 ('TGM2', 0.00019828846299006187),
 ('CD300A', 0.00019827718131737378),
 ('DAG1', 0.0001976844975130656),
 ('EPS8', 0.00019767664464022676),
 ('CD38', 0.00019756770652741305),

('NAV2', 0.00019734417718386623),
 ('MYO1E', 0.00019613815693932612),
 ('C19orf33', 0.00019578704017027035),
 ('MDN1', 0.0001951062395293256),
 ('ZNF22', 0.00019490508858533388),
 ('LAT2', 0.00019402276140987563),
 ('APBB1IP', 0.00019177204903503612),
 ('BOD1', 0.00019104242876060652),
 ('GULP1', 0.00019033198155406526),
 ('CD63', 0.00018982613192953566),
 ('TPM1', 0.00018967531431362953),
 ('WTAP', 0.00018818843249155507),
 ('PHLDA2', 0.00018786853731874258),
 ('GPX8', 0.00018786510082179303),
 ('FAM117A', 0.00018749723876529547),
 ('GPRC5A', 0.00018700446132205182),
 ('GGCX', 0.0001866316603965172),
 ('SGMS2', 0.00018651388075809893),
 ('ADAT2', 0.00018632090838444798),
 ('NEDD4', 0.00018629324544970565),
 ('ME2', 0.0001857411039404022),
 ('PIK3AP1', 0.0001852674148669145),
 ('DCK', 0.0001851356686274468),
 ('SUGP2', 0.0001850740758595044),
 ('GIPC1', 0.00018322508408075649),
 ('MIR22HG', 0.00018288322806925606),
 ('ING1', 0.00018286193760634275),
 ('SMIM8', 0.00018199816114176963),
 ('UACA', 0.000181733440314187),
 ('SLC39A8', 0.00018040903310789213),
 ('S100A16', 0.00018015881919574897),
 ('RSL24D1', 0.00017948023832275496),
 ('NSUN4', 0.0001793104499257302),
 ('ACTN4', 0.00017922186375730323),
 ('ALDH3B1', 0.00017919789638220284),
 ('SLC26A10', 0.00017888588762488198),
 ('INPP5E', 0.00017854823133782989),
 ('AHNAK2', 0.00017794082136178658),
 ('PVRL2', 0.0001777780738465244),
 ('TLCD1', 0.00017772893663816665),
 ('ZAR1', 0.00017770672472354249),
 ('NPTX1', 0.00017759937394956458),
 ('C9orf123', 0.00017722535861825693),
 ('NISCH', 0.0001765883056989772),
 ('LAPTM4B', 0.00017647441612530832),
 ('RAB33A', 0.0001763671927149965),
 ('CCDC28B', 0.00017576215740731306),

('SH2D4A', 0.00017513600255185773),
 ('MYCBP2', 0.00017495579260796135),
 ('TUFT1', 0.00017482231975942674),
 ('PNISR', 0.00017471870576308348),
 ('ZNF197', 0.00017430839768914287),
 ('GTF2F1', 0.0001741534462376313),
 ('OR2D2', 0.00017408008338192865),
 ('C18orf21', 0.00017404799893055336),
 ('DCP2', 0.00017398601321400243),
 ('LSM6', 0.00017381687358875757),
 ('RANBP3', 0.00017365345039565557),
 ('LRRC8E', 0.00017293375443016364),
 ('BLM', 0.00017277010120460274),
 ('SEC31B', 0.00017213033103097696),
 ('KLF4', 0.00017186925817145045),
 ('HCST', 0.00017157412966907303),
 ('ZNF250', 0.0001706705870355665),
 ('KDM4D', 0.00017036783109997502),
 ('SLC30A1', 0.00017022541635690593),
 ('CD151', 0.00016913738673469156),
 ('TFB1M', 0.0001684779548939315),
 ('MARK1', 0.00016838488130823407),
 ('MID1', 0.0001668517858966839),
 ('ARSD', 0.00016640447757417158),
 ('TAF4B', 0.0001663843366155421),
 ('X.25', 0.00016615799167463458),
 ('SBF2', 0.00016590266778214002),
 ('JARID2', 0.00016585840094735576),
 ('LSP1', 0.00016563909481702842),
 ('TIFA', 0.0001653939665409199),
 ('ARHGEF12', 0.00016483578177116227),
 ('RPS6KA3', 0.0001645804960791602),
 ('LY9', 0.0001644262485205431),
 ('AFAP1', 0.0001636161268657191),
 ('EXOSC9', 0.00016352994494533868),
 ('GDF15', 0.0001632106617013881),
 ('TMC8', 0.00016309040432810963),
 ('MPST', 0.00016286735088392456),
 ('FAM189B', 0.0001627069484512911),
 ('EVA1A', 0.0001625465688056152),
 ('HABP4', 0.00016210227174144826),
 ('MX2', 0.00016162942313484148),
 ('FMNL3', 0.00016160688605458467),
 ('ST8SIA4', 0.00016102273302944875),
 ('SPTBN1', 0.0001608438343058988),
 ('MYB', 0.00016083930761407055),
 ('ALOX5AP', 0.0001607440560143867),

('SASH1', 0.00016062207815418754),
 ('ASB9', 0.0001606171978336714),
 ('ZNF330', 0.0001605602929547609),
 ('LIPC', 0.0001603813198779827),
 ('MYO1F', 0.00016026553856602602),
 ('ZMAT2', 0.00016018551076352286),
 ('NFATC3', 0.00016011082649859432),
 ('NAP1L5', 0.00015942308850476272),
 ('CELF2', 0.00015928720608911483),
 ('ZNRD1', 0.00015922492414373332),
 ('POLR2B', 0.00015902408181479513),
 ('C18orf25', 0.00015837812216203684),
 ('EFCAB2', 0.00015833939575360834),
 ('SLC2A8', 0.00015801921272631865),
 ('LY86', 0.00015791805400823894),
 ('LRRC41', 0.00015762393747492434),
 ('C6orf203', 0.00015756073186979502),
 ('BIN2', 0.00015737210482272014),
 ('RNF166', 0.0001570591431603726),
 ('DOK3', 0.00015705791175015527),
 ('LBX1', 0.00015664547119715442),
 ('PLGRKT', 0.00015612391762243673),
 ('EDN1', 0.00015583914608958808),
 ('SERINC3', 0.00015545172321842124),
 ('KCNAB3', 0.00015527752656987865),
 ('TAF5', 0.00015502572266073633),
 ('SNX9', 0.00015479681040111426),
 ('TFAP2A', 0.00015471344440846837),
 ('TNS3', 0.00015449051279236835),
 ('ADA', 0.00015425407016498509),
 ('LTBP4', 0.00015394870960691733),
 ('UNC50', 0.00015361973614129776),
 ('ATM', 0.00015344584928313833),
 ('MRPL4', 0.00015325350553497928),
 ('CLDN12', 0.00015310744631658236),
 ('VAV1', 0.00015273440152804754),
 ('RHOG', 0.00015268422599923668),
 ('CLCN4', 0.00015246765797463936),
 ('DSTN', 0.00015234523384660922),
 ('MUT', 0.00015232700264163906),
 ('PARK2', 0.00015200138976833384),
 ('LTBP3', 0.00015154245513031878),
 ('WIPF1', 0.0001515132370665496),
 ('BCL11A', 0.00015145806924107724),
 ('CRK', 0.0001509042760722372),
 ('TRIM47', 0.00015072908101993178),
 ('ANKRD11', 0.00015060575552772697),

('PCDH1', 0.00015058248128064705),
 ('ZUFSP', 0.00015049685182108683),
 ('ASPHD2', 0.0001502692303239284),
 ('CWF19L1', 0.00015026363262253542),
 ('EPB41L1', 0.00014979929731687508),
 ('HCAR1', 0.00014973154141954023),
 ('KCNAB2', 0.00014963818116063885),
 ('EWSR1', 0.0001494245972396352),
 ('CLPP', 0.00014913526077893154),
 ('PPL', 0.00014901577686624067),
 ('BTBD16', 0.00014886880488648916),
 ('SMCR7L', 0.00014874053364214535),
 ('SAFB2', 0.00014867650965646084),
 ('ZADH2', 0.00014865587375147987),
 ('LSMD1', 0.00014834095525107207),
 ('SCIN', 0.0001481848136168423),
 ('CASP3', 0.00014809592517344732),
 ('UBXN7', 0.00014798034547782956),
 ('HDAC7', 0.00014793170863541206),
 ('TMEM17', 0.00014790296469553038),
 ('HNRNPA1L2', 0.0001478605583646581),
 ('DEF6', 0.0001477917602729084),
 ('TMEM82', 0.00014684961107656752),
 ('HAUS1', 0.00014669117876143736),
 ('ZC2HC1C', 0.00014669017435488133),
 ('RABL6', 0.00014661357376567053),
 ('CISH', 0.00014646936259192178),
 ('TMEM127', 0.00014642051275708928),
 ('CREB1', 0.00014634302692256664),
 ('CASP8AP2', 0.00014586437276071766),
 ('TFRC', 0.00014580757731537402),
 ('ITGB3BP', 0.00014564825795189773),
 ('XP07', 0.000145426908225521),
 ('KRI1', 0.00014531594953409144),
 ('TMTC2', 0.00014530230129661137),
 ('DEPDC5', 0.00014478883187747052),
 ('CCND1', 0.0001444940437515363),
 ('NUP188', 0.00014421773856794677),
 ('ZC3HAV1', 0.00014421296503751445),
 ('INTU', 0.00014418952251952127),
 ('CYB5R2', 0.00014412228837668348),
 ('ANXA4', 0.0001441212739843431),
 ('SAMD4A', 0.00014402411545159945),
 ('KIAA1671', 0.0001438364757626321),
 ('GIMAP2', 0.00014379859894408554),
 ('ATP5A1', 0.0001437008478120661),
 ('X02.Mar.1', 0.00014360370243245895),

('APTX', 0.00014323747597296292),
 ('ANKRD1', 0.00014312266509550252),
 ('C2orf44', 0.00014308983573293676),
 ('ACSL1', 0.00014307976000332224),
 ('ESD', 0.00014307191352142884),
 ('B4GALT4', 0.00014291260835257998),
 ('SREK1', 0.00014289893499474626),
 ('PPWD1', 0.00014273113561106776),
 ('KXD1', 0.00014266516563113402),
 ('AKNA', 0.00014259768872916824),
 ('RPL14', 0.00014255490688884115),
 ('SNRPA1', 0.00014235952549820801),
 ('WLS', 0.00014195697815496503),
 ('GIPR', 0.00014187921213282883),
 ('C9orf152', 0.00014182919556503552),
 ('IFI35', 0.00014151650564014795),
 ('PLD6', 0.0001414582797348543),
 ('UBE3D', 0.00014072929734979802),
 ('ALDH1B1', 0.000140490815317756),
 ('KRT15', 0.0001404636269374821),
 ('DRAM1', 0.00014045627628390893),
 ('FNDC3B', 0.0001399753538313623),
 ('EVL', 0.00013967387317525868),
 ('C17orf76.AS1', 0.00013957380030948832),
 ('FLT3LG', 0.00013946955492528276),
 ('DYNC1LI2', 0.0001394471874857889),
 ('POU2AF1', 0.0001393622637671563),
 ('RAB32', 0.00013928980628822717),
 ('OCSTAMP', 0.00013928649960938592),
 ('DYM', 0.00013904908061600517),
 ('CRYGB', 0.0001388571505872654),
 ('TRIM32', 0.00013877527187085403),
 ('MLL', 0.00013828225385789604),
 ('X01.Sep', 0.00013821152583003447),
 ('YARS', 0.00013821110535954583),
 ('RPAP1', 0.00013765496970724088),
 ('X.14', 0.00013746043176864883),
 ('MID2', 0.00013729899611939313),
 ('LRP11', 0.00013719290194599972),
 ('OSTF1', 0.0001371608869203493),
 ('CAMK2N1', 0.00013702813017368634),
 ('SCRN1', 0.00013693436156682622),
 ('SLC39A13', 0.00013677006560890522),
 ('PHLDA1', 0.0001367604377055156),
 ('RLTPR', 0.00013674393289329732),
 ('PTK2B', 0.0001367262911470528),
 ('WDR18', 0.00013597165595045433),

('PSIP1', 0.00013570936283705665),
 ('C1GALT1', 0.0001356640949685747),
 ('RPL35', 0.00013554732488694032),
 ('GPRC5C', 0.0001355242008078803),
 ('AVPI1', 0.00013517991333084555),
 ('SSBP3', 0.000135127162069331),
 ('LGMN', 0.00013506345723986402),
 ('SNX20', 0.00013477362109696586),
 ('FCHSD2', 0.0001347411103813533),
 ('MAT2B', 0.0001342458147215494),
 ('ATG14', 0.0001341988200873553),
 ('TUBE1', 0.0001341828158254179),
 ('GLUD2', 0.00013404678881089813),
 ('CYSLTR1', 0.00013404231886663786),
 ('KCTD14', 0.0001339094437194585),
 ('PAX9', 0.00013390934347692326),
 ('GBA', 0.00013376811332225434),
 ('ATP5G2', 0.00013357250630428462),
 ('HOXD10', 0.00013352658375401357),
 ('NPEPPS', 0.00013352342121808537),
 ('BTBD2', 0.00013311490734325802),
 ('NGLY1', 0.00013308297835368826),
 ('ABCE1', 0.00013304844722647426),
 ('SLC44A3', 0.00013290847812357472),
 ('CNN3', 0.00013273650512926974),
 ('HEATR5A', 0.00013272632450586335),
 ('INTS10', 0.00013270747944213132),
 ('METAP2', 0.0001324350606413095),
 ('UPP1', 0.00013241570155764971),
 ('ERBB2IP', 0.00013235823688519162),
 ('TEX10', 0.00013214401193550784),
 ('MACF1', 0.00013204571346789264),
 ('TRAF3IP2', 0.00013171835327787797),
 ('KLK6', 0.00013158002612555388),
 ('COX7B', 0.0001310237287202848),
 ('SLC2A1', 0.00013100911157849847),
 ('C4orf19', 0.00013066567244761275),
 ('SET', 0.00013009309431943593),
 ('RHPN2', 0.00012996526419852415),
 ('RAB3B', 0.00012994156942709866),
 ('KBTBD6', 0.00012993154022314932),
 ('SSBP4', 0.00012982258428659405),
 ('SMARCC1', 0.00012964094082532112),
 ('TMEM184B', 0.000129546327600181),
 ('WDFY4', 0.0001295227970947177),
 ('FOXC1', 0.0001292790903663395),
 ('ASB15', 0.00012925014202807953),

('KIAA1199', 0.00012915614573735283),
 ('EPDR1', 0.00012897322408585938),
 ('ENC1', 0.0001288952295785192),
 ('C7orf60', 0.0001288864794106853),
 ('MED28', 0.00012866165419391865),
 ('MTMR9', 0.00012860679324395552),
 ('ADM', 0.00012847369253133565),
 ('PDCD11', 0.00012822404752916647),
 ('ATP6V0E1', 0.00012820436641660957),
 ('CROCC', 0.0001280920791222301),
 ('CXXC1', 0.00012808464051514143),
 ('DUSP3', 0.00012808397340859583),
 ('CCL20', 0.00012803231348135598),
 ('IP05', 0.0001279606096743532),
 ('HSPB8', 0.00012785666558797786),
 ('ITGB7', 0.00012779592055925967),
 ('C11orf57', 0.00012753897544991305),
 ('FAM216B', 0.00012731675028773126),
 ('EIF3D', 0.000127025982086883),
 ('PLA2G16', 0.0001268930302842378),
 ('NR2E1', 0.00012687514152626264),
 ('SRF', 0.00012673405907089874),
 ('SPEN', 0.0001267201729658735),
 ('CASP10', 0.00012671575523145383),
 ('FAM174A', 0.00012670617728512468),
 ('OSBPL3', 0.0001266652000589079),
 ('ERAP1', 0.00012663718022214092),
 ('KHSRP', 0.00012658353814998247),
 ('DNM3', 0.00012657498493282707),
 ('C1QBP', 0.0001265268852767523),
 ('LYSMD1', 0.00012650827308661742),
 ('CTNND1', 0.00012650010377986137),
 ('N4BP2', 0.00012637686128520192),
 ('CCR7', 0.00012634568496842777),
 ('GUCY1B3', 0.00012624939266339064),
 ('MMP21', 0.0001261291649745235),
 ('ATAD2B', 0.0001257623042604906),
 ('ANKS3', 0.00012570023903921233),
 ('RIBC2', 0.00012550514311128567),
 ('YIPF5', 0.00012533336639449614),
 ('NPHP4', 0.00012524052303398628),
 ('ADNP2', 0.0001250715215043914),
 ('SH3BP4', 0.00012496290687881713),
 ('PIK3R5', 0.00012473641049194704),
 ('TCL1A', 0.00012467246161205852),
 ('KDM1A', 0.00012446302652310932),
 ('MAP3K1', 0.0001243593519899237),

('ENDOD1', 0.0001240569937516735),
 ('MRPL50', 0.00012403465411686622),
 ('OR10K1', 0.00012402942418193068),
 ('EFCAB7', 0.00012398634815570585),
 ('ZCCHC14', 0.00012397332054998883),
 ('SLC31A1', 0.0001237580596054543),
 ('BAIAP2', 0.00012372622161735405),
 ('AMOTL2', 0.0001237008563748732),
 ('ADI1', 0.0001236451125663478),
 ('PIK3CD', 0.00012362686699422275),
 ('GNS', 0.00012345852760769845),
 ('RBMS2', 0.00012345850801587024),
 ('PCGF2', 0.00012335625831706815),
 ('SHPRH', 0.00012326107765595848),
 ('PAK4', 0.00012324180754292227),
 ('HEXB', 0.00012317013155469696),
 ('DOCK5', 0.00012306636437902366),
 ('PIGL', 0.00012298797002008373),
 ('ACN9', 0.00012275370650278438),
 ('LIPH', 0.00012260082230577645),
 ('IGHMBP2', 0.00012249026326569828),
 ('RSL1D1', 0.00012244832169586154),
 ('ICAM3', 0.00012243544852144205),
 ('ACTL8', 0.00012236473806677702),
 ('WRNIP1', 0.00012225200130893825),
 ('CRLF3', 0.00012208841927339976),
 ('CCL17', 0.00012205356353809379),
 ('LPP', 0.00012200037412808894),
 ('GGA1', 0.00012199063099195293),
 ('CAV2', 0.00012195250129211553),
 ('RND3', 0.00012190621638775726),
 ('TNFAIP8', 0.00012183963446843707),
 ('CCDC68', 0.00012183941217432787),
 ('ARL6', 0.0001218020292225387),
 ('XPC', 0.00012179406205537766),
 ('CFLAR', 0.00012176668112087709),
 ('LASP1', 0.000121738548185859),
 ('CTGF', 0.00012162692331751296),
 ('KRT34', 0.00012152912009433607),
 ('NOP58', 0.00012149983283534033),
 ('BMP4', 0.00012149822934929255),
 ('C6orf123', 0.0001214337531542468),
 ('ARPC1B', 0.00012136642533373477),
 ('RBP2', 0.00012133562658076044),
 ('POLR3B', 0.00012132964654439954),
 ('BFSP1', 0.0001212043426080878),
 ('CTDSPL', 0.00012116134181251845),

('NKTR', 0.00012090847167945205),
 ('C9orf72', 0.00012072911304268138),
 ('PCBD2', 0.00012064759883994669),
 ('WAS', 0.000120599768586783),
 ('CYP2R1', 0.00012057512785101896),
 ('MEGF11', 0.00012053619963246655),
 ('TRPM8', 0.00012016200444015791),
 ('PCYOX1', 0.00012010938564215425),
 ('PMM1', 0.00011996866100383064),
 ('RGS18', 0.0001199645010709473),
 ('ABHD2', 0.00011989613406362646),
 ('HRCT1', 0.00011984801259864708),
 ('RAB38', 0.0001198178237538417),
 ('NMRAL1', 0.00011964154415303878),
 ('IL1B', 0.0001195747483849211),
 ('ZNF589', 0.00011957429379141637),
 ('RRP1B', 0.00011944470117135958),
 ('GOPC', 0.00011942478317602036),
 ('UQCRH', 0.00011936602099209373),
 ('VAPB', 0.00011936304897736307),
 ('BRI3', 0.00011936137133550061),
 ('CUEDC1', 0.00011908297527707417),
 ('KRT19', 0.00011892027468430329),
 ('PARP1', 0.00011858434874496635),
 ('TRAPPC10', 0.00011850415551145051),
 ('KRT80', 0.00011843601702508718),
 ('PIKFYVE', 0.0001182689688685499),
 ('TMEM51', 0.00011825759411243328),
 ('X.50', 0.00011823435928355975),
 ('ZNF555', 0.0001180270188572001),
 ('BBS7', 0.00011796892880787796),
 ('MGST1', 0.00011785650577265342),
 ('RPL9', 0.0001177880556545619),
 ('MEI1', 0.00011771646723117521),
 ('ATP8B1', 0.00011746570767008947),
 ('ABHD11', 0.00011742836025872087),
 ('CD46', 0.0001173763366581872),
 ('SCML4', 0.00011721773116455194),
 ('MAN2A1', 0.00011710705321654359),
 ('LSM7', 0.00011707200173729526),
 ('ABHD5', 0.00011704553155720936),
 ('PTPRR', 0.00011702249151824114),
 ('GALC', 0.00011693470561925388),
 ('MAGEE2', 0.00011690860140145824),
 ('MLPH', 0.00011689382251414523),
 ('MTA1', 0.00011677388630397428),
 ('TMEM99', 0.00011671454812640526),

('SF3A1', 0.00011662816934594906),
 ('PPDPF', 0.00011656766327363516),
 ('BCL7A', 0.00011649172891757447),
 ('RPRD1A', 0.00011633446531576028),
 ('GNG4', 0.00011615998156696541),
 ('AGPAT9', 0.00011613583063623618),
 ('CNTRL', 0.00011610955444370057),
 ('CLASP2', 0.00011598480158638157),
 ('C19orf54', 0.00011598109000112507),
 ('FAM46C', 0.00011593161107702728),
 ('NUP210', 0.00011591646974223723),
 ('RTP2', 0.00011588735207124466),
 ('GTPBP1', 0.00011583246909511282),
 ('THAP10', 0.0001158001306878675),
 ('ANXA1', 0.00011573510985981667),
 ('SERBP1', 0.00011570498130644588),
 ('AEN', 0.00011561410452645814),
 ('SKP2', 0.00011540423853864352),
 ('ZNF671', 0.00011537730725361691),
 ('CHD4', 0.00011534302201365356),
 ('PICALM', 0.00011523237184867309),
 ('ALG1L', 0.00011521040156695685),
 ('ENAH', 0.00011517129432703514),
 ('APBB2', 0.00011515599373138206),
 ('GNB3', 0.00011510555450417863),
 ('PEX11G', 0.00011500380588735572),
 ('TMEM151B', 0.0001149758477657973),
 ('PRR14L', 0.00011497303557730877),
 ('EIF2B5', 0.00011463270902072979),
 ('GRAP', 0.00011460651825770968),
 ('CAMSAP2', 0.00011452912622086475),
 ('RNASEH2B', 0.00011443349191496305),
 ('ELAVL2', 0.00011440923400122807),
 ('CHST11', 0.00011440888928633352),
 ('KIAA1551', 0.00011435365481813104),
 ('COBL', 0.00011432285585762194),
 ('GPR17', 0.00011429233211520934),
 ('PLCB3', 0.00011426280383522246),
 ('C14orf93', 0.00011423276326034433),
 ('MGAT5', 0.00011410820304461554),
 ('KIR3DL1', 0.00011407830499010771),
 ('TFPI', 0.00011404156308186282),
 ('TNFSF11', 0.00011398656157297198),
 ('SMU1', 0.00011392370168759248),
 ('H1FX', 0.0001136096877798406),
 ('TUSC1', 0.00011354002491026577),
 ('RFXAP', 0.00011350025506326198),

('ZNF107', 0.00011343030673251011),
 ('GFPT2', 0.00011338990168359622),
 ('MFSD5', 0.00011328386840862007),
 ('XAF1', 0.00011327311976204652),
 ('BEND4', 0.00011314440231383566),
 ('ARHGAP5', 0.00011310837424554454),
 ('SNAPC3', 0.00011292492585471644),
 ('ERBB3', 0.00011290833364303866),
 ('POU6F1', 0.00011286248344722757),
 ('PVR', 0.00011278884504792778),
 ('NAA15', 0.00011276856656106587),
 ('ANXA11', 0.0001127344970780901),
 ('VMP1', 0.00011263963402784415),
 ('RPE65', 0.00011262225686473857),
 ('GALNT11', 0.00011257258461112762),
 ('EVI2A', 0.00011255714788631409),
 ('LARP6', 0.00011251865901548839),
 ('ATP5I', 0.00011251787340746946),
 ('ITM2A', 0.00011248133454743226),
 ('OR10A7', 0.00011247388866610998),
 ('CARM1', 0.00011246438165729981),
 ('ARHGAP28', 0.00011244891492111743),
 ('WWTR1', 0.00011243409081442315),
 ('EPB41', 0.00011240533053619167),
 ('FRYL', 0.00011240162644795762),
 ('FKBP8', 0.00011238425409065097),
 ('SYNE4', 0.00011229213706689652),
 ('SKAP1', 0.00011216917621912346),
 ('CTNNBIP1', 0.00011209738867083127),
 ('RBM17', 0.00011206215781678064),
 ('S100A11', 0.0001120420513332606),
 ('MUC15', 0.00011201052195578578),
 ('HTRA1', 0.00011199333747469589),
 ('P4HA2', 0.00011192435475930044),
 ('IQCG', 0.00011180855350219889),
 ('CAMSAP1', 0.00011179378784275736),
 ('IKZF5', 0.00011175857765403449),
 ('FHL2', 0.0001117431109566433),
 ('BLMH', 0.0001116012370995695),
 ('DPY19L3', 0.00011159692514000392),
 ('RAB13', 0.00011158227297685637),
 ('VPS41', 0.00011154577259976589),
 ('FRAT2', 0.00011154385735711122),
 ('HDHD2', 0.00011149722032060953),
 ('ZBTB20', 0.00011149247445087527),
 ('TM4SF4', 0.00011148210120677045),
 ('MCCC1', 0.00011141991332532933),

('GPR75.ASB3', 0.00011127734024855738),
 ('SMAD1', 0.00011127487646346602),
 ('CCDC90B', 0.00011117642714221538),
 ('RARA', 0.0001111678303781745),
 ('PAPD4', 0.00011113785682538296),
 ('FAM127B', 0.00011106670550086356),
 ('CD69', 0.00011100597094480035),
 ('TMEM56', 0.00011096447467713389),
 ('SAT2', 0.000110789784284026),
 ('SLC43A1', 0.00011074107812047477),
 ('TFDP3', 0.0001107384040886407),
 ('RPL22', 0.00011061994767761544),
 ('TEC', 0.00011057478488656847),
 ('TRIM62', 0.00011055733067174073),
 ('ZSCAN2', 0.00011041771271576988),
 ('SLC6A15', 0.00011039635463847679),
 ('C4orf27', 0.00011035449793640854),
 ('TGFBI', 0.00011029198416399506),
 ('PHF15', 0.00011004971872685587),
 ('NPW', 0.00011003747244167662),
 ('RBM15', 0.00010984503506320652),
 ('ZNF79', 0.00010984330554389048),
 ('RCOR2', 0.00010983971871731245),
 ('PPP6C', 0.0001097143631274259),
 ('ECHDC2', 0.00010970965665812953),
 ('SUMF1', 0.0001095722418428369),
 ('C12orf29', 0.00010953880754384743),
 ('ZBTB39', 0.00010949035351453622),
 ('SIDT1', 0.0001094001787255519),
 ('UQCR10', 0.00010939180547414589),
 ('TOP2B', 0.00010930361828747904),
 ('NSMF', 0.00010910899311846455),
 ('LYPD1', 0.00010909776949377458),
 ('GJB5', 0.00010906667205986142),
 ('C5orf15', 0.00010900674600676006),
 ('MAP3K15', 0.00010895175611637288),
 ('HTATIP2', 0.00010890881057244286),
 ('MS4A14', 0.00010887154435584562),
 ('SNRPD2', 0.00010887059863767743),
 ('VRK1', 0.00010886532063314632),
 ('ST3GAL3', 0.00010886126580360578),
 ('SCARB2', 0.00010883436511705235),
 ('GPR113', 0.00010881585075286191),
 ('DYRK1B', 0.00010872024920634399),
 ('LRRC40', 0.00010845182430356704),
 ('LRFN5', 0.00010828058448548861),
 ('SUCLG1', 0.00010824850707385943),

('UGP2', 0.00010821718355374024),
 ('MED27', 0.00010820955781989824),
 ('MEF2C', 0.00010818530793015991),
 ('HMG20A', 0.00010790919659909924),
 ('NUDT4', 0.0001078526612694671),
 ('PTMS', 0.0001075731531194638),
 ('LTBR', 0.00010755453346923461),
 ('RFC1', 0.000107512366959247),
 ('CLIP1', 0.00010740409034250282),
 ('SH3KBP1', 0.00010738316296734861),
 ('MED30', 0.00010737213186645761),
 ('ICK', 0.00010732526481394212),
 ('SLC16A10', 0.00010714741347913862),
 ('ZNF28', 0.00010710400645098433),
 ('LIF', 0.00010706723353247943),
 ('NFIB', 0.00010705061772758125),
 ('SLAIN1', 0.00010698134194910164),
 ('DCAF4L2', 0.00010692249763091712),
 ('MRPL2', 0.00010674473256373113),
 ('NPTN', 0.00010673788285093256),
 ('SMAD2', 0.00010673784891257665),
 ('ZDHHC4', 0.00010655757453685207),
 ('PUM1', 0.00010642927939192631),
 ('PPP1R8', 0.00010640351312454171),
 ('GIMAP4', 0.00010636792746760418),
 ('FBX038', 0.00010631006214653777),
 ('ISG20', 0.00010624856152362197),
 ('TRAPPC2L', 0.00010623128935751227),
 ('MLC1', 0.00010622711896422615),
 ('AGPAT2', 0.00010621909617352277),
 ('IGF2.AS', 0.00010621219051427197),
 ('OSGIN1', 0.00010620922397040009),
 ('FBXL2', 0.0001060641684009232),
 ('FIP1L1', 0.00010602088249234575),
 ('SYTL2', 0.00010601021416752407),
 ('DFFB', 0.00010599264379060954),
 ('KIAA1967', 0.00010597475104346802),
 ('CDC42SE1', 0.0001059714420395247),
 ('TRIT1', 0.00010588841551799454),
 ('PPP2R3A', 0.00010587854382400501),
 ('STAC3', 0.00010585219361415649),
 ('CHAMP1', 0.00010577590484247487),
 ('ZFAND4', 0.00010573708037211556),
 ('FBXL13', 0.00010562859748055484),
 ('WDR75', 0.00010561990037243382),
 ('PHACTR2', 0.0001055839513810187),
 ('CD70', 0.00010554413227233732),

('SYNJ1', 0.00010552999911829764),
('VIP', 0.00010552803367676713),
('NEK7', 0.00010542722256796734),
('C9orf9', 0.00010527718413514347),
('SMAD9', 0.00010526532686076545),
('UBR1', 0.00010519155341626905),
('ABHD12', 0.00010518016330648017),
('CD79B', 0.00010516376465101892),
('CALB2', 0.000105139456015031),
('SIK2', 0.0001051055771479859),
('NHP2L1', 0.00010506612394526101),
('C14orf79', 0.0001050230496122317),
('JUN', 0.00010498627841141874),
('MLF2', 0.00010498016454222681),
('CYP4Z1', 0.0001049733477086255),
('RBMX', 0.00010495984978566082),
('MAPKAPK3', 0.00010486568186372841),
('OR1I1', 0.00010484010017171282),
('TMPRSS15', 0.00010475687766843887),
('PAPSS1', 0.00010461811837308674),
('WDR37', 0.00010456212716946877),
('C16orf87', 0.00010452222823165324),
('FAM212A', 0.00010446753364629972),
('SKIL', 0.00010440995558067784),
('GTPBP4', 0.00010438977129989565),
('KIAA1468', 0.000104321041499251),
('NLRP1', 0.0001043109656952559),
('MRPS26', 0.00010429985860612822),
('PNO1', 0.00010419419458917214),
('GLUD1', 0.00010417831895476096),
('SFMBT2', 0.0001041345284978642),
('LONP1', 0.00010410693970906983),
('WBSR28', 0.00010408838126522272),
('PPA2', 0.0001040818512374933),
('GRB14', 0.00010405598454915496),
('AGRN', 0.00010402583245687129),
('CYC1', 0.00010399251504255047),
('ZBTB24', 0.00010396285354342392),
('GTF3C4', 0.00010395741670951613),
('ZNF837', 0.00010393867886479913),
('PAK1IP1', 0.00010393070622197144),
('HERC5', 0.00010393022389198331),
('PHF21A', 0.00010391778795320858),
('CLHC1', 0.00010378814462831482),
('AIM1L', 0.0001037518284058673),
('GTF2E1', 0.00010374220038502846),
('ZNF267', 0.0001037079489614743),

('TTC39B', 0.00010369815956417242),
 ('HKDC1', 0.0001036679335330172),
 ('CYB5B', 0.00010360648666985341),
 ('MFAP1', 0.00010356858099463029),
 ('MT01', 0.00010348688541644995),
 ('CAST', 0.00010347678129715028),
 ('AQP11', 0.00010344901131608555),
 ('NDUFS7', 0.00010340056326445456),
 ('COX4I2', 0.00010336562457461265),
 ('PLEKHA1', 0.00010335541899482498),
 ('MYL6B', 0.00010331052472957153),
 ('ZNF266', 0.00010326463251337943),
 ('SPAG9', 0.00010317486854548657),
 ('CERS5', 0.00010308734104546983),
 ('REV3L', 0.00010306129011470458),
 ('FAXC', 0.0001030525466670221),
 ('NACAD', 0.0001029404399470263),
 ('CENPJ', 0.0001028519518906932),
 ('MARVELD2', 0.00010283161771277776),
 ('OSCP1', 0.00010278019571202039),
 ('TSPYL1', 0.00010272338313656459),
 ('RNASE6', 0.00010271013649625021),
 ('PIGA', 0.0001026891538739377),
 ('SLC25A18', 0.00010266574968790338),
 ('C1orf159', 0.00010265795542153722),
 ('TRMT10C', 0.00010261392295209455),
 ('SPG7', 0.0001022421226374167),
 ('TRABD', 0.00010217026077828),
 ('LHFPL2', 0.00010215341163605212),
 ('TRAPPC8', 0.00010209866267409131),
 ('EPM2A', 0.00010209467919622587),
 ('LMAN2', 0.00010200821630505888),
 ('C17orf97', 0.00010199912334850358),
 ('LINC00315', 0.00010199343774981952),
 ('SLC12A3', 0.00010199194024083494),
 ('PPARA', 0.00010198668260210426),
 ('TMEM161A', 0.0001018344290788641),
 ('GGNBP2', 0.00010182624853877242),
 ('CMTM8', 0.00010179936358407899),
 ('GID4', 0.00010171930272906663),
 ('THAP9', 0.00010171897651270741),
 ('USP42', 0.0001016824853948489),
 ('SNIP1', 0.0001016776894714302),
 ('RPS15A', 0.00010164942529390884),
 ('CBFA2T3', 0.0001016266108622534),
 ('MOB3B', 0.00010151572820857508),
 ('NCOR2', 0.00010147786339564384),

```
( 'AGAP2', 0.00010146597061535894),
( 'MAP4K3', 0.00010146101215379553),
( 'PTPLAD1', 0.00010130131759459086),
( 'ALDH8A1', 0.0001012822160318171),
( 'ATP2B3', 0.00010122307663956416),
( 'CSTL1', 0.00010121413312946474),
( 'RHOC', 0.00010120735274661704),
( 'RNLS', 0.00010115225846280895),
( 'RSF1', 0.00010106017299545783),
( 'RUSC1', 0.00010094160150174401),
( 'SNX24', 0.00010081028307566847),
( 'FATE1', 0.00010074208423758049),
( 'TMEM136', 0.00010066406952137548),
( 'KCND2', 0.00010060499433829198),
( 'SLC7A7', 0.00010059482366605378),
( 'TRAF7', 0.00010055319719941096),
( 'TAB2', 0.00010054724394675346),
( 'CCDC25', 0.00010046392845210267),
( 'DNALI1', 0.00010045268828118123),
( 'INSL5', 0.00010039814467487944),
( 'CCDC66', 0.00010037321392971355),
( 'ATPBD4', 0.00010037177800990573),
( 'MANSC1', 0.00010030572578477986),
( 'ERI3', 0.00010030194270695263),
( 'VPRBP', 0.00010027111504671298),
( 'AASDH', 0.00010025882669965391),
( 'MYC', 0.0001002444069595565),
( 'CST3', 0.00010022054617121901),
( 'TMEM139', 0.00010020858939594932),
( 'OR56B4', 0.00010017661815077853),
( 'NDUFV1', 0.00010009959573706836),
( 'C2CD2', 0.00010001739626859894),
( 'IL3RA', 0.00010000219394167435),
( 'EPB41L5', 9.998416238125454e-05),
( 'MAU2', 9.997281547341168e-05),
( 'FRG1', 9.993872928755831e-05),
( 'FBXL12', 9.993659015740556e-05),
( 'OR51G2', 9.990882087453549e-05),
( 'SEC23A', 9.987882497584519e-05),
...]
```

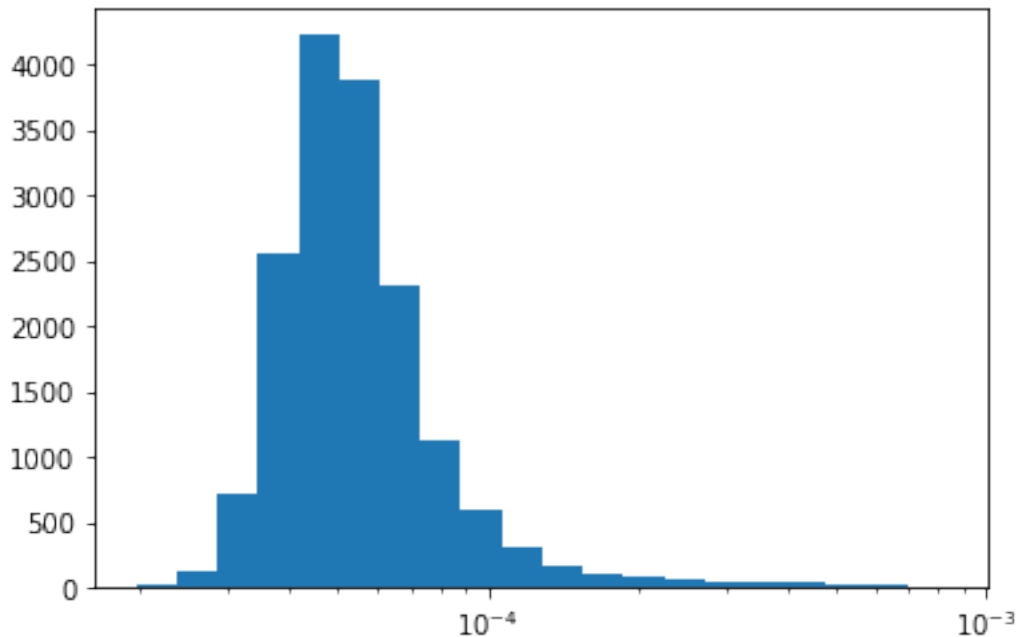
```
[130]: # plotting the feature importances
import matplotlib.pyplot as plt
indices = np.arange(len(feat_importances))
x, y = zip(*feat_importances)

def plot_loghist(x, bins):
```

```

hist, bins = np.histogram(x, bins=bins)
logbins = np.logspace(np.log10(bins[0]), np.log10(bins[-1]), len(bins))
plt.hist(x, bins=logbins)
plt.xscale('log')
plot_loghist(y, 20)

```



```

[80]: # creating selector object that uses random forest classifier to identify
      ↪ features
      # importance of more than  $e-04$  will be used to select features
      sfm = SelectFromModel(clf, threshold=1e-4)
      sfm.fit(X_train, y_train)

```

```

[80]: SelectFromModel(estimator=RandomForestClassifier(bootstrap=True, ccp_alpha=0.0,
class_weight=None,
criterion='gini',
max_depth=None,
max_features='auto',
max_leaf_nodes=None,
max_samples=None,
min_impurity_decrease=0.0,
min_impurity_split=None,
min_samples_leaf=1,
min_samples_split=2,
min_weight_fraction_leaf=0.0,
n_estimators=16381, n_jobs=-1,

```



```

                                oob_score=False,
                                random_state=0, verbose=0,
                                warm_start=False),
    max_features=None, norm_order=1, prefit=False,
    threshold=0.0001)

```

```

[85]: # transforming data to create new dataset containing only the most important
      ↪ features
      X_important_train = sfm.transform(X_train)
      X_important_test = sfm.transform(X_test)

```

```

[86]: # creating and training new random forest classifier for most important features
      clf_important = RandomForestClassifier(n_estimators=X.shape[1], random_state=0,
      ↪ n_jobs=-1)
      clf_important.fit(X_important_train, y_train)

```

```

[86]: RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight=None,
                             criterion='gini', max_depth=None, max_features='auto',
                             max_leaf_nodes=None, max_samples=None,
                             min_impurity_decrease=0.0, min_impurity_split=None,
                             min_samples_leaf=1, min_samples_split=2,
                             min_weight_fraction_leaf=0.0, n_estimators=16381,
                             n_jobs=-1, oob_score=False, random_state=0, verbose=0,
                             warm_start=False)

```

```

[87]: # apply the full-featured classifier to test data
      y_pred = clf.predict(X_test)
      accuracy_score(y_test, y_pred)

```

```

[87]: 0.5045871559633027

```

```

[88]: # apply important features classifier to test data
      y_important_pred = clf_important.predict(X_important_test)
      accuracy_score(y_test, y_important_pred)

```

```

[88]: 0.48623853211009177

```

```

[102]: s_feats = pd.DataFrame(selected_feats)
      s_feats.to_csv('selected_genes.csv', index=False)

```

```

[ ]:

```

```

[ ]:

```