

SNVBox Demo of SNVGet

Prior to running the demo, make sure that you have downloaded and installed SNVBox and SNVGet and that they are working correctly.

We will use a set of mutations from 11 breast tumors published by Sjoblom *et al.* that are pre-formatted in transcript coordinates and made available for download at http://www.karchinlab.org/CHASM_example.tmps.

We retrieve 5 features from SNVBox by building a featurelist file with the following 5 lines:

```
ACharge  
HMMRelEntropy  
PredRSAE  
REGCompILVM  
UniprotDOM_PostModEnz
```

We run SNVGet to retrieve feature values:

```
./snvGetTranscript -f Featurelist -o output.arff CHASM_example.tmps
```

The output.arff file appears as follows (first 20 mutations shown):

```
@ relation headerfile  
  
@attribute UID string  
@attribute ID string  
@attribute ACharge numeric  
@attribute HMMRelEntropy numeric  
@attribute PredRSAE numeric  
@attribute REGCompILVM numeric  
@attribute UniprotDOM_PostModEnz numeric  
  
@data  
  
1 NM_000546.5_R273L 1 1.9317 0.205 0.1333 0  
2 NM_000546.5_R248Q 1 2.3396 0.271 0.4000 0  
3 NM_000546.5_V157F 0 2.1042 0.014 0.2000 0  
4 NM_000546.5_D281H -2 4.0044 0.325 0.0667 0  
5 NM_000546.5_R175H 0 2.2785 0.48 0.2000 0  
6 NM_000546.5_R248W 1 2.3396 0.271 0.4000 0  
7 NM_006218.2_H1047R 0 0.5779 0.173 0.1333 1  
8 NM_000546.5_Y163C 0 2.1497 0.068 0.2667 0  
9 NM_000552.3_Y1570C 0 0.3072 0.597 0.1333 0  
10 NM_002838.3_G863R -1 2.3134 0.016 0.3333 1  
11 NM_005012.2_E562D 0 0.6411 0.697 0.3333 1  
12 NM_017677.3_W127R -1 0.8694 0.17 0.2667 1  
13 NM_024715.3_S248P 0 0.6186 0.773 0.2000 1  
14 NM_014974.2_V1264M 0 0.4848 0.036 0.3333 0  
15 NM_031431.3_R620C 1 1.9215 0.174 0.4000 0  
16 NM_000350.2_T224M 0 0.1070 0.345 0.1333 0  
17 NM_014568.1_L692F 0 1.4232 0.045 0.2000 0  
18 NM_172238.3_V214F 0 2.0093 0.095 0.3333 0  
19 NM_000499.3_R477W 1 0.5196 0.48 0.4667 0  
20 NM_020734.2_N63K -1 0.1321 0.583 0.3333 0
```