CHASM (Cancer-specific High-throughput Annotation of Somatic Mutations) Demo

Prior to running the demo, make sure that you have downloaded and installed CHASM correctly.

Next, download a set of mutations from 11 breast tumors published by Sjoblom *et al.*

http://www.karchinlab.org/CHASM_example.tmps.

The mutations are pre-formatted using transcript coordinates.

This command line will train a classifier using a passenger mutation rate table representative of breast tumors:

```
cd CHASM
./BuildClassifier -m ClassfierPack/contexts/BRCA.context -f
ClassifierPack/features/Features.list -o ExampleClassifier -i CHASM_example.tmps -s 1234
```

The following command line will score the mutations in the supplied dataset:

```
./RunChasm ExampleClassifier CHASM_example.tmps
```

The top scoring results for the example dataset are shown in Table 2. An extra column with gene symbol has been added for illustrative purposes.

In this example, 27 out of 878 somatic missense mutations detected in the 11 breast tumors received significant CHASM scores at an FDR < 25%. The choice of FDR cutoff represents a tradeoff between sensitivity and specificity and should be selected as appropriate the application. Some actual driver mutations are likely to be missed by choosing a conservative FDR cutoff, therefore a less conservative cutoff may be desired if false positive predictions are acceptable.

The top scoring mutations in the example dataset include known driver mutations in TP53, PIK3CA as well as mutations in NOTCH1 and HNF1A, two genes known to contribute to tumorigenesis. The top candidate driver mutations also include mutations in genes of unknown significance to cancer, including MTMR3, MTMR8, UBR4 and PRDM13.

Combining CHASM prioritization with information about the functional role and previous cancer association of the mutated gene may highlight particularly interesting predicted driver mutations.

Table 1 CHASM scores for top 20 predicted driver mutations

HUGO Symbol	Mutation ID	Mutation	CHASM score	p-value	BHFDR
TP53	3	NM_000546.5_V157F	0.002	0.000	0.05
TP53	8	NM_000546.5_Y163C	0.004	0.000	0.05
TP53	1	NM_000546.5_R273L	0.006	0.000	0.05
TP53	6	NM_000546.5_R248W	0.014	0.000	0.05
PIK3CA	7	NM_006218.2_H1047R	0.016	0.000	0.05
TP53	4	NM_000546.5_D281H	0.032	0.000	0.05
TP53	5	NM_000546.5_R175H	0.034	0.000	0.05
TP53	2	NM_000546.5_R248Q	0.036	0.000	0.05
NOTCH1	401	NM_017617.3_G1216D	0.144	0.000	0.05
MTMR3	32	NM_021090.3_V221L	0.192	0.000	0.05
HNF1A	406	NM_000545.5_K273E	0.202	0.000	0.05
NOTCH1	165	NM_017617.3_D620Y	0.242	0.000	0.05
TXNDC15	13	NM_024715.3_S248P	0.246	0.001	0.05
DNAJC10	24	NM_018981.1_D727H	0.248	0.001	0.05
UBR4	374	NM_020765.2_R1394H	0.266	0.001	0.10
PRDM13	809	NM_021620.3_T54N	0.330	0.003	0.20
SUV39H2	870	NM_024670.3_D323H	0.332	0.004	0.20
MTMR8	12	NM_017677.3_W127R	0.342	0.004	0.20
NXN	31	NM_022463.4_H195Y	0.342	0.004	0.20