Pile|Line

Example #2.

Using Nsmc to compare variants between Genomic Position files.

PileLine

Step 1.

Download and uncompress the Genomic Position example files:

Sample Data

Here you can find some files to test PileLine and PileLine GUI. All files are compressed. You need to uncompress them before use.

File	Type	Description
hg18.zip **(929MB)**	Genome (.fa and .fai)	Human Genome 18 in FASTA (.fa) and indexed FASTA (.fai). The sequence names are without the 'chr' prefix.
Control1Files.zip **(36MB)** Case1Files.zip **(36MB)** Control2Files.zip **(38MB)** Case2Files.zip **(38MB)**	Genotype data (.pileup)	Several genotyped samples. each .zip file contains 2 files: a complete and a variants-only pileup. **Note: there is only data for the chromosome 10**
Hg18_hgnc_ensembl_genes.bed.zip **(365KB)**	Gene annotation (.bed file)	Ensembl Genes for the Human Genome 18
DbSNP_36.3.bed.bgz.zip **(150MB)**	SNP annotation (custom in compressed (bgz) BED format)	SNPs from dbSNP version 36.3 for the Human Genome 18

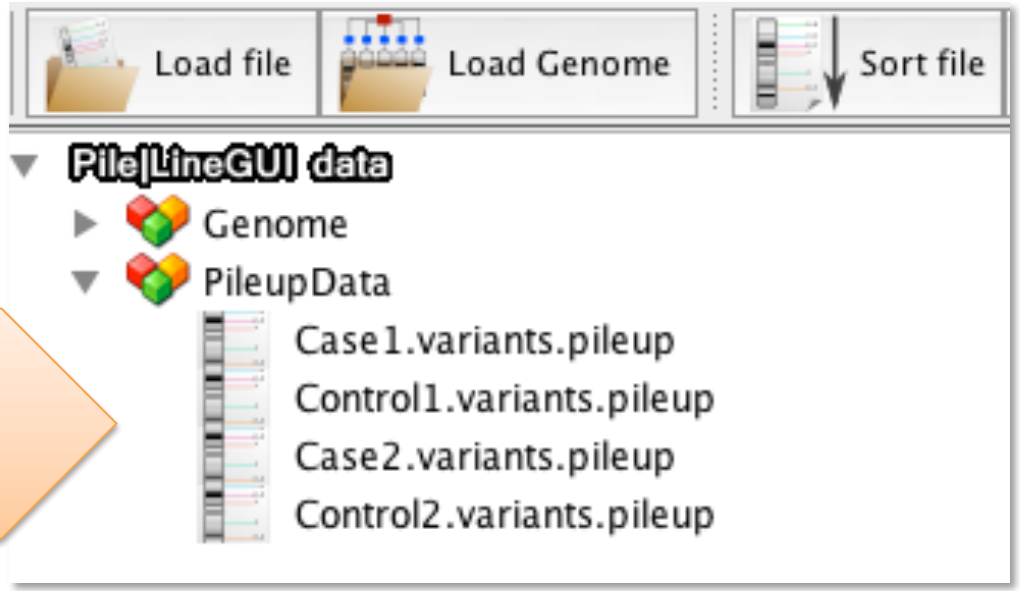
Pile|Line

Step 2.

Load uncompressed Genomic Position files:

- Case1.variants.pileup
- Control1. variants.pileup
- Case2.variants.pileup
- Control2.variants.pileup

Find the loaded files in the PileLine Data Repository



Pile|Line

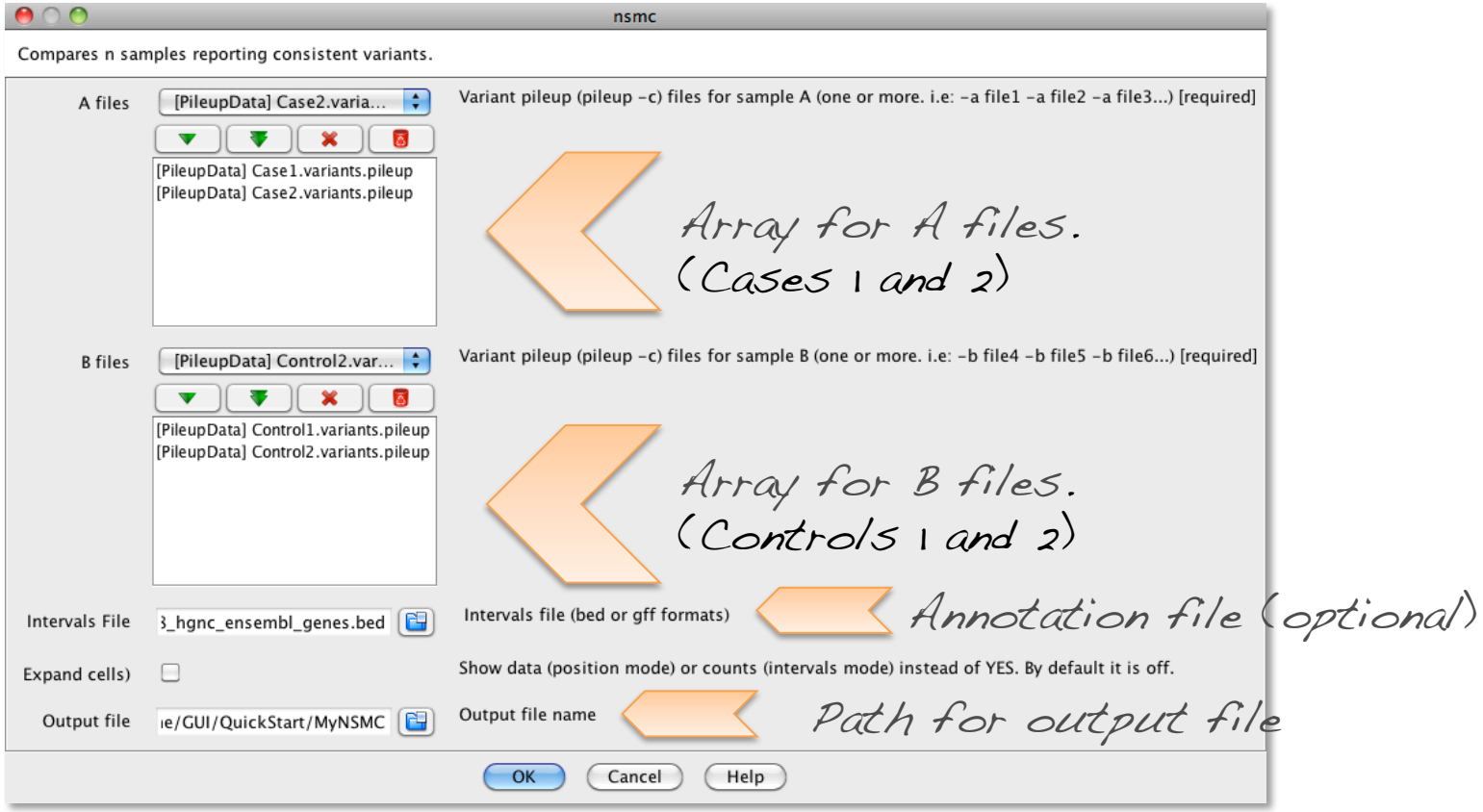
Step 3.

a) Click on nsmc button in the upper toolbar.
This will display the nsmc parameters menu.

Pile|Line

Step 4.

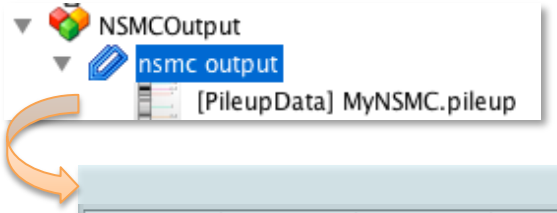
a) To report variants in all Cases vs all Controls select the parameters indicated below.



Pile|Line

Step 5.

Check results output in the Data Repository!!



Case 1 | Case 2 | Control 1 | Control 2

seq	inverval st...	interval end	annotation data 0	Case1.vari...	Case2.vari...	Control1.v...	Control2.v...	no. sample...	p-val	FDR
10	116334	122205	ENSG00000217176|IL9RP2	YES	YES	YES	YES	4	1.000000...	1.0
10	170643	290577	ENSG00000015171|ZMYND11	YES	YES	NO	NO	2	0.166666...	1.0
10	685888	699940	ENSG00000180525|C10orf108	NO	NO	NO	NO	0	1.000000...	1.0
10	311432	725518	ENSG00000151240|DIP2C	NO	NO	NO	NO	0	1.000000...	1.0
10	845484	967440	ENSG00000107929|LARP5	NO	NO	NO	NO	0	1.000000...	1.0
10	1024349	1053688	ENSG00000107937|GTPBP4	NO	NO	NO	NO	0	1.000000...	1.0
10	1054847	1061799	ENSG00000148377|IDI2	NO	NO	NO	NO	0	1.000000...	1.0
10	1075966	1085061	ENSG00000067064|IDI1	NO	NO	NO	NO	0	1.000000...	1.0
10	1085478	1168237	ENSG00000047056|WDR37	NO	NO	NO	NO	0	1.000000...	1.0
10	1218073	1769670	ENSG00000185736|ADARB2	NO	NO	NO	NO	0	1.000000...	1.0
10	3099740	3169762	ENSG00000067057|PFKP	NO	NO	NO	NO	0	1.000000...	1.0
10	3169922	3205003	ENSG00000107959|PITRM1	NO	NO	NO	NO	0	1.000000...	1.0
10	3811260	3817455	ENSG00000067082|KLF6	NO	NO	NO	NO	0	1.000000...	1.0
10	4858402	4880249	ENSG00000165568|AKR1CL2	NO	NO	NO	NO	0	1.000000...	1.0
10	4995566	5010445	ENSG00000187134|AKR1C1	NO	NO	NO	NO	0	1.000000...	1.0
10	5021966	5050207	ENSG00000151632|AKR1C2	NO	NO	NO	NO	0	1.000000...	1.0
10	5126583	5139878	ENSG00000196139|AKR1C3	NO	NO	NO	NO	0	1.000000...	1.0
10	5186335	5217144	ENSG00000196326|AKR1CL1	NO	NO	NO	NO	0	1.000000...	1.0
10	5227426	5250912	ENSG00000198610|AKR1C4	NO	NO	NO	NO	0	1.000000...	1.0
10	5396976	5406169	ENSG00000178473|UCN3	NO	NO	NO	NO	0	1.000000...	1.0

Variants coordinates and annotation

Presence of variants in files

samples with variant

p-value and FDR

PileLine

For further info see PileLineGUI Documentation at:

http://sing.ei.uvigo.es/pileline/pilelineguihelp/