Joint/conditional analysis and fine-mapping

The INTERVAL data is used as reference panel. The logic of this specific directory is a simple solution of the dilemma that the reference data, possibly like others, uses reference sequence ID (rsid) whenever possible. However, during meta-analysis the practice of using rsid is undesirable so SNPID, i.e., chr:pos_A1_A2, (A1<=A2) is necessary.

After a rather long and laborious process involving many software, it turned out a simple way out is to obtain sentinels using SNPID but return to rsid at this stage and forward. The implementation here reflects this. The file INTERVAL.rsid contains SNPID-rsid mapping and could be generated from programs such as `qctool/bgenix/plink`.

A note on regions is ready. It is attractive to use the last genomic region from iterative merging for analysis and perhaps a flanking version. This is more appropriate than genomewide hard and fast 10MB windows or approximately independent LD blocks. For the latter, we found that the boundaries from the distributed 1000Genomes project were often inappropriate and one may not attempt to compute them for specific reference panel. Nevertheless, the iterative procedure actually just does empirically. Again the HLA region is condensed.

The last point regards software `finemap`, which uses summary statistics associated with the reference panel rather than that from meta-analysis.

Main analyses

File specification	Function
NLRP2.sh	the exclusion list
cs.sh	Credible sets
ma.sh	INF1 sumstats
INTERVAL-ma.sh	INTERVAL sumstats
prune.sh	pruning
slct.sh	GCTA --cojo-slct analysis
finemap.sh	`finemap` analysis
jam.sh	`JAM` analysis
coloc.sb	coloc analysis -- clumsy verion
coloc.R	coloc analysis via pQTLtools
fastenloc.sb	fastenloc analysis
garfield.sh	GARFIELD analysis
hyprcoloc.sh	hyprcoloc analysis
st.sh	batch command file
cs/, finemap/, jam/, prune/, work/	working directories

Miscellaneous analyses and utilities

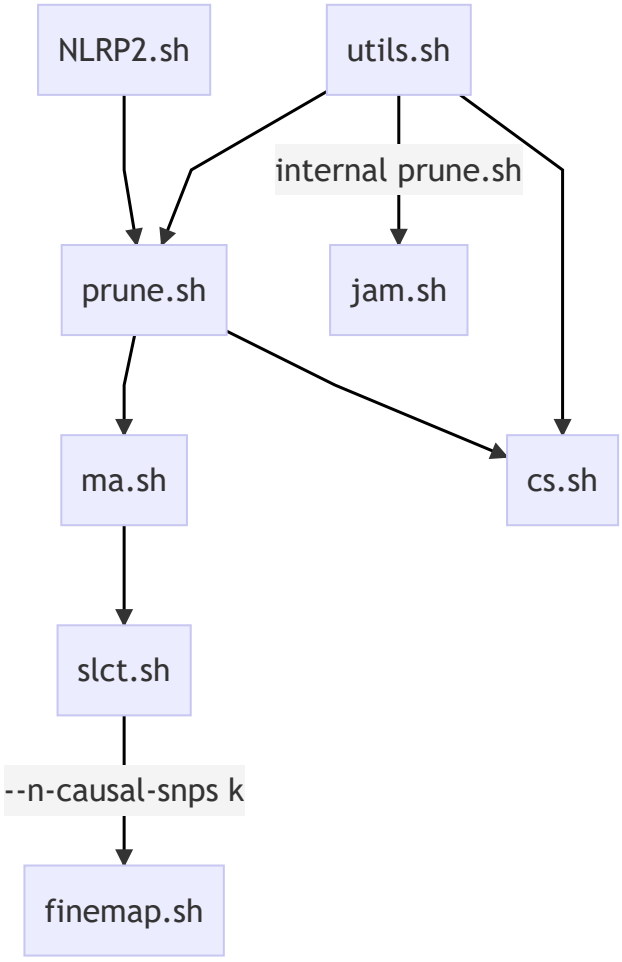
FIle	Description
chembl.sh	toy CHEMBL queries
CRP.sh	an inflammation score
cvd1.sh	SCALLOP-CVDI supplementary tables
efo.R	experimental factor ontology
eQTL.R	cis-pQTL eQTL lookup (PhenoScanner)
GTEx.R	cis-pQTL GTEx eQTL lookup (PhenoScanner)
GTEx.sh	cis-pQTL GTEx eQTL lookup (GTEx v8)
gdigt.R	GDI and gene-trials
gwasvcf.sh	GWAS to VCF conversion
gwas2vcf.sb	gwas2vcf + VCF operations
HGI.sh	HGI data analysis
latex.tex	LaTeX formulas embedded in Markdown
latex.docx	LaTeX formulas embedded in MicroSoft Word
LTBR.sh	LTBR LocusZoom plots
magma.sh	MAGMA for IL.12B
neale.sh	signal selection for Neale UKB data (HbA1c)
pleiotropy.R	horizontal and vertical pleiotropy
ppi.R	PPI using EpiGraphDB
pqtlGWAS.R	pQTL-GWAS lookup
pQTL.R	pQTL lookup
pqtlMR.sh	pQTL MR analysis
rentrez.sh	reuse of rentrez
rGREAT.R	GREAT analysis
stringdb.sh	STRINGdb
tables.R	code to create Excel Tables
uniprot.R	UniProt IDs to others
utils.sh	utilties
vep.sh	VEP annotation
wgcna.sh	experiment on modules

Stacked associaiton plots

1. IL.18-rs385076.sh
2. rs12075.sh
3. TNFB-rs2364485.sh
4. TNFB-rs2364485-MR.sh (a two-sample MR)
5. OPG-TRANCE.sh

Steps

`st.sh` conceptually executes the following elements,



Note that the `GCTA` .ma, jma.cojo, .ldr.cojo become -rsid.ma, -rsid.jma.cojo, -rsid.ldr.cojo, respectively; the same are true for files related to `finemap`.

Date last changed: **18/1/2021**