Review and testing of gwasurvivr: an R package for genome-wide survival analysis

Nadia   Ivy   Gabriel   Piotr

2023-10-05

Introduction

Methods

Datasets

Results

Michigan Imputation Server pre-phases typed genotypes using HAPI-UR, SHAPEIT, or EAGLE (default is EAGLE2), imputes using Minimac3 imputation engine and outputs Blocked GNU Zip Format VCF files (`.vcf.gz`). These `vcf.gz` files are used as input for `gwasurvivr`. Datafile retreived as a survival dataframe can be represented in `R` as below.

ID_1	ID_2	event	time	age	DrugTxYes	sex	group
1	SAMP1	0	12.00	33.93	0	male	control
2	SAMP2	1	7.61	58.71	1	male	experimental
3	SAMP3	0	12.00	39.38	0	female	control
4	SAMP4	0	4.30	38.85	0	male	control
5	SAMP5	0	12.00	43.58	0	male	experimental
6	SAMP6	1	2.60	57.74	0	male	control

Now we can run `michiganCoxSurv` wrapper for Cox regression model.

```r
#decoding sex into binary format
pheno.file$SexFemale <- ifelse(pheno.file$sex=="female", 1L, 0L)

michiganCoxSurv(vcf.file=vcf.file,
                covariate.file=pheno.file,
                id.column="ID_2",
                time.to.event="time",
                event="event",
                covariates=c("age", "SexFemale", "DrugTxYes"),
                inter.term=NULL,
                print.covs="only",
                out.file="michigan_only",
                r2.filter=0.3,
                maf.filter=0.005,
```

```
            chunk.size=100,
            verbose=TRUE,
            clusterObj=NULL)
```

Accessed results of the performed regression are showcased below.

RSID	rs34919020	rs8005305	rs757545375
TYPED	FALSE	FALSE	FALSE
CHR	14	14	14
POS	19459185	20095842	20097287
REF	C	G	A
ALT	T	T	G
AF	0.301263	0.514583	0.519787
MAF	0.301263	0.485417	0.480213
SAMP_FREQ_ALT	0.3428	0.5022	0.5110
SAMP_MAF	0.3428	0.4978	0.4890
R2	0.551952	0.479015	0.480693
ER2	NA	NA	NA
PVALUE	0.2934544	0.3238959	0.2862329
HR	1.5085220	0.7233560	0.7046073
HR_lowerCI	0.7005469	0.3801063	0.3702421
HR_upperCI	3.248374	1.376573	1.340937
Z	1.0505737	-0.9864835	-1.0664221
COEF	0.4111304	-0.3238538	-0.3501147
SE.COEF	0.3913389	0.3282911	0.3283078
N	100	100	100
N.EVENT	42	42	42

Conclusions