BTRY4830-Project

Vedha Mahesh

2023-05-05

libs = c('tidyverse', 'data.table', 'MASS',  
 'ggsci', 'ggthemes', 'ggpubr', 'KEGGREST',  
 'rentrez', 'tidyimpute','ggplot','ggfortify')  
suppressMessages(  
 suppressWarnings(sapply(libs, require, character.only = TRUE))  
)

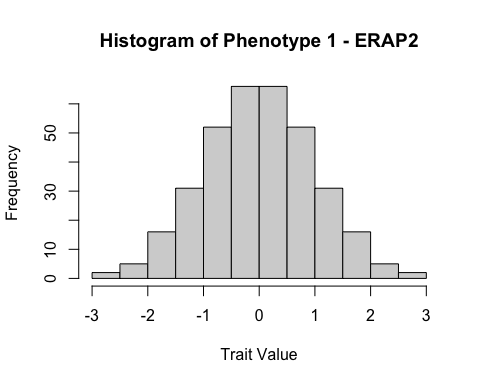
## tidyverse data.table MASS ggsci ggthemes ggpubr KEGGREST   
## TRUE TRUE TRUE FALSE TRUE FALSE FALSE   
## rentrez tidyimpute ggplot ggfortify   
## FALSE FALSE FALSE TRUE

rm(libs)  
theme\_set(theme\_few(14))

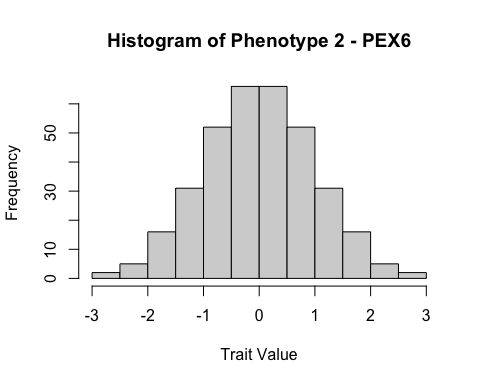
### 1.

#### a.

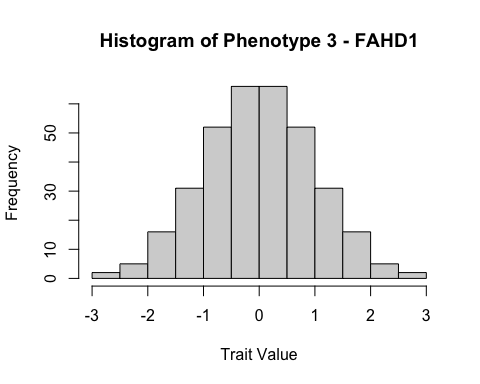
pheno\_import <- read.csv("projectphenotypes.csv")  
  
pheno1 <- pheno\_import$ENSG00000164308.12  
  
hist(pheno1, main = "Histogram of Phenotype 1 - ERAP2", xlab = "Trait Value", ylab = "Frequency")



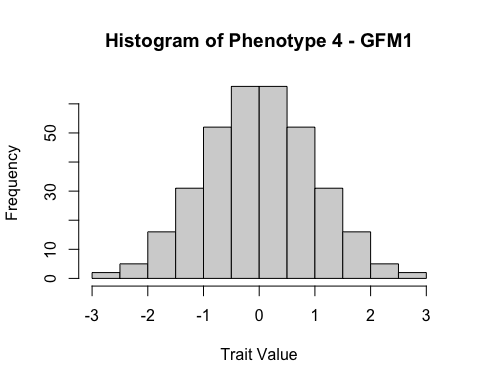
pheno\_import <- read.csv("projectphenotypes.csv")  
  
pheno2 <- pheno\_import$ENSG00000124587.9  
  
hist(pheno2, main = "Histogram of Phenotype 2 - PEX6", xlab = "Trait Value", ylab = "Frequency")



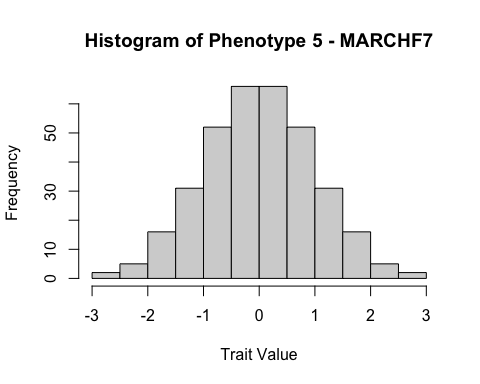
pheno\_import <- read.csv("projectphenotypes.csv")  
  
pheno3 <- pheno\_import$ENSG00000180185.7  
  
hist(pheno3, main = "Histogram of Phenotype 3 - FAHD1", xlab = "Trait Value", ylab = "Frequency")



pheno\_import <- read.csv("projectphenotypes.csv")  
  
pheno4 <- pheno\_import$ENSG00000168827.9  
  
hist(pheno4, main = "Histogram of Phenotype 4 - GFM1", xlab = "Trait Value", ylab = "Frequency")



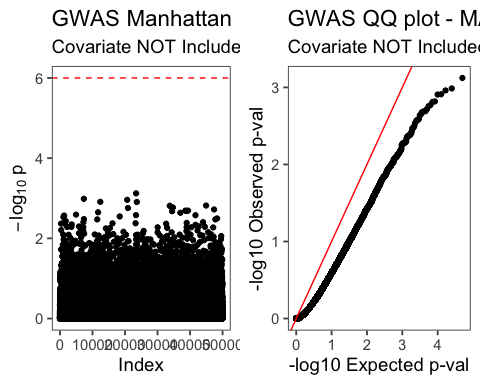
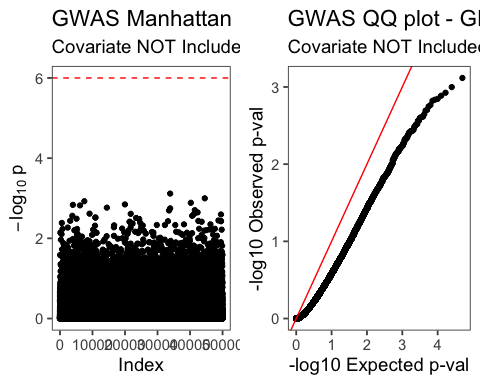
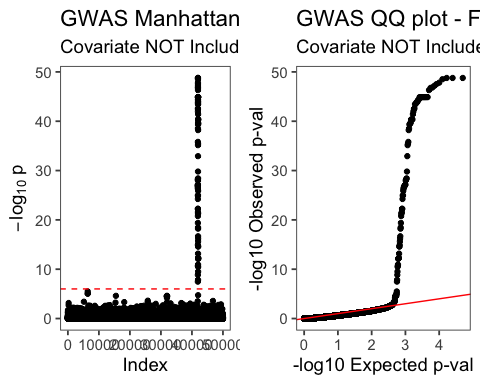
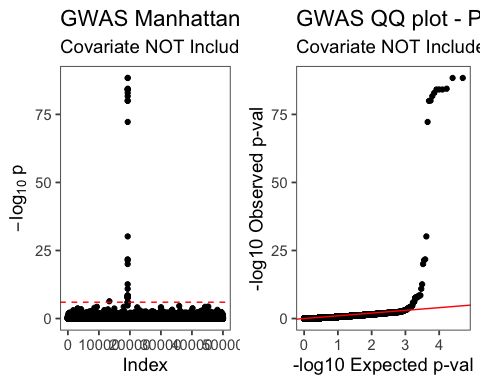
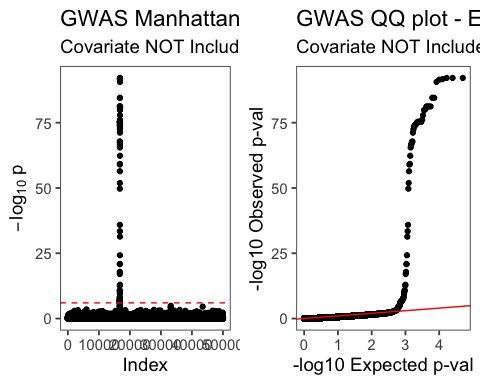
pheno\_import <- read.csv("projectphenotypes.csv")  
  
pheno5 <- pheno\_import$ENSG00000136536.9  
  
hist(pheno5, main = "Histogram of Phenotype 5 - MARCHF7", xlab = "Trait Value", ylab = "Frequency")



geno\_import <- read.csv("projectgenotypes.csv")  
  
N <- ncol(geno\_import) - 1  
  
cat("Number of SNPs (N):", N, "\n")

## Number of SNPs (N): 50000

gen\_import <- read.csv("projectgenotypes.csv")  
gen\_import<- gen\_import[,-1]  
colnames(gen\_import) <- NULL   
rownames(gen\_import) <- NULL



# Convert genotypes to Xa & Xd  
# # Details can be found in lab 6  
gen\_import <- read.csv("projectgenotypes.csv")  
gen\_import<- gen\_import[,-1]  
colnames(gen\_import) <- NULL   
rownames(gen\_import) <- NULL   
  
xa\_mat <- as.matrix(gen\_import )  
xd\_mat <- ((2\*xa\_mat) - 1)  
  
  
# Load in phenotype data and run GWAS  
sim\_pheno\_mx <- read.csv("projectphenotypes.csv",   
 header = TRUE, row.names = 1)  
  
# Read-in our covariate data  
xc\_mat = read.csv('projectcovars.csv')  
Xc <- model.matrix(~Population + Sex, data=xc\_mat)  
xc\_mat1 <- Xc[,-1]  
  
xc.pca<- prcomp(xa\_mat %\*% t(xa\_mat))  
  
xc.pca$sdev

## [1] 6.427237e+03 3.116876e+03 1.626246e+03 1.393769e+03 1.348115e+03  
## [6] 1.262807e+03 1.209274e+03 1.203364e+03 1.196482e+03 1.189354e+03  
## [11] 1.183723e+03 1.176994e+03 1.169424e+03 1.163703e+03 1.159626e+03  
## [16] 1.154370e+03 1.152850e+03 1.148883e+03 1.141628e+03 1.138455e+03  
## [21] 1.135365e+03 1.131426e+03 1.129800e+03 1.124041e+03 1.121787e+03  
## [26] 1.119182e+03 1.117536e+03 1.113179e+03 1.111885e+03 1.109252e+03  
## [31] 1.104816e+03 1.103658e+03 1.101757e+03 1.100419e+03 1.098486e+03  
## [36] 1.096698e+03 1.095543e+03 1.095326e+03 1.089676e+03 1.088675e+03  
## [41] 1.087383e+03 1.086510e+03 1.082802e+03 1.081972e+03 1.079094e+03  
## [46] 1.077650e+03 1.076481e+03 1.073658e+03 1.073091e+03 1.070557e+03  
## [51] 1.067799e+03 1.067384e+03 1.064559e+03 1.063028e+03 1.061835e+03  
## [56] 1.059029e+03 1.057953e+03 1.055338e+03 1.054601e+03 1.052785e+03  
## [61] 1.052125e+03 1.050047e+03 1.049038e+03 1.048526e+03 1.046296e+03  
## [66] 1.043911e+03 1.042524e+03 1.041937e+03 1.038166e+03 1.036866e+03  
## [71] 1.035134e+03 1.033292e+03 1.031230e+03 1.030714e+03 1.029552e+03  
## [76] 1.027296e+03 1.026493e+03 1.026187e+03 1.023989e+03 1.022830e+03  
## [81] 1.020409e+03 1.018483e+03 1.018153e+03 1.017024e+03 1.015774e+03  
## [86] 1.015339e+03 1.013081e+03 1.011066e+03 1.010345e+03 1.009144e+03  
## [91] 1.007144e+03 1.004131e+03 1.003394e+03 1.002108e+03 1.001168e+03  
## [96] 1.000130e+03 9.990847e+02 9.977815e+02 9.966950e+02 9.953574e+02  
## [101] 9.934285e+02 9.928147e+02 9.921263e+02 9.889116e+02 9.878715e+02  
## [106] 9.870330e+02 9.858830e+02 9.844095e+02 9.828766e+02 9.817144e+02  
## [111] 9.815441e+02 9.801736e+02 9.794226e+02 9.773364e+02 9.762718e+02  
## [116] 9.747273e+02 9.743429e+02 9.733624e+02 9.721300e+02 9.694334e+02  
## [121] 9.690543e+02 9.665108e+02 9.659888e+02 9.651578e+02 9.646666e+02  
## [126] 9.628692e+02 9.617965e+02 9.606342e+02 9.589537e+02 9.579263e+02  
## [131] 9.576399e+02 9.565771e+02 9.551647e+02 9.534546e+02 9.527577e+02  
## [136] 9.513468e+02 9.510565e+02 9.497074e+02 9.475445e+02 9.472500e+02  
## [141] 9.452895e+02 9.447093e+02 9.439998e+02 9.431473e+02 9.419851e+02  
## [146] 9.413799e+02 9.396646e+02 9.373141e+02 9.370022e+02 9.351853e+02  
## [151] 9.336451e+02 9.327803e+02 9.318289e+02 9.309163e+02 9.301936e+02  
## [156] 9.289218e+02 9.275203e+02 9.272192e+02 9.248979e+02 9.237018e+02  
## [161] 9.223170e+02 9.212970e+02 9.207484e+02 9.190583e+02 9.175964e+02  
## [166] 9.168758e+02 9.155734e+02 9.144915e+02 9.139177e+02 9.115838e+02  
## [171] 9.100090e+02 9.089549e+02 9.082701e+02 9.075609e+02 9.070621e+02  
## [176] 9.061834e+02 9.037305e+02 9.031297e+02 9.026705e+02 9.015705e+02  
## [181] 9.006897e+02 8.995660e+02 8.988972e+02 8.969267e+02 8.967292e+02  
## [186] 8.947837e+02 8.945184e+02 8.930036e+02 8.923525e+02 8.912859e+02  
## [191] 8.905865e+02 8.886448e+02 8.882295e+02 8.867257e+02 8.855461e+02  
## [196] 8.846877e+02 8.830573e+02 8.825466e+02 8.818479e+02 8.814751e+02  
## [201] 8.794259e+02 8.777186e+02 8.767972e+02 8.758468e+02 8.757262e+02  
## [206] 8.744545e+02 8.736686e+02 8.721873e+02 8.717075e+02 8.697694e+02  
## [211] 8.692546e+02 8.677861e+02 8.661287e+02 8.652436e+02 8.642717e+02  
## [216] 8.633542e+02 8.617654e+02 8.611184e+02 8.597113e+02 8.594240e+02  
## [221] 8.587240e+02 8.581504e+02 8.569184e+02 8.564457e+02 8.553471e+02  
## [226] 8.539759e+02 8.532587e+02 8.521061e+02 8.513172e+02 8.493478e+02  
## [231] 8.485337e+02 8.477899e+02 8.466858e+02 8.455183e+02 8.446625e+02  
## [236] 8.424588e+02 8.415318e+02 8.411388e+02 8.405069e+02 8.389686e+02  
## [241] 8.382553e+02 8.380126e+02 8.361499e+02 8.351868e+02 8.348635e+02  
## [246] 8.329343e+02 8.322722e+02 8.314363e+02 8.306184e+02 8.290522e+02  
## [251] 8.274878e+02 8.267802e+02 8.260722e+02 8.250299e+02 8.235588e+02  
## [256] 8.224138e+02 8.219797e+02 8.203458e+02 8.194079e+02 8.182833e+02  
## [261] 8.171338e+02 8.157413e+02 8.156224e+02 8.139621e+02 8.137759e+02  
## [266] 8.123691e+02 8.112856e+02 8.101449e+02 8.091977e+02 8.087009e+02  
## [271] 8.077159e+02 8.062131e+02 8.055174e+02 8.031145e+02 8.023779e+02  
## [276] 8.010343e+02 8.007215e+02 7.984976e+02 7.981396e+02 7.966775e+02  
## [281] 7.958301e+02 7.954928e+02 7.941076e+02 7.935662e+02 7.911468e+02  
## [286] 7.908311e+02 7.903589e+02 7.891410e+02 7.869310e+02 7.858087e+02  
## [291] 7.855454e+02 7.849577e+02 7.820512e+02 7.811003e+02 7.802921e+02  
## [296] 7.796162e+02 7.790589e+02 7.770834e+02 7.754008e+02 7.753526e+02  
## [301] 7.728763e+02 7.726947e+02 7.705890e+02 7.701995e+02 7.693325e+02  
## [306] 7.686392e+02 7.667707e+02 7.665011e+02 7.648718e+02 7.640755e+02  
## [311] 7.615802e+02 7.614540e+02 7.589987e+02 7.568496e+02 7.560320e+02  
## [316] 7.550266e+02 7.538642e+02 7.529631e+02 7.526921e+02 7.492002e+02  
## [321] 7.486189e+02 7.473058e+02 7.456907e+02 7.421292e+02 7.408742e+02  
## [326] 7.401063e+02 7.387272e+02 7.382018e+02 7.359075e+02 7.341505e+02  
## [331] 7.326539e+02 7.291693e+02 7.283385e+02 7.257851e+02 7.224219e+02  
## [336] 7.202882e+02 7.180573e+02 7.160515e+02 7.140979e+02 7.082242e+02  
## [341] 7.067593e+02 6.938449e+02 5.987201e+02 2.584256e-12

(xc.pca$sdev / sum(xc.pca$sdev))\*100

## [1] 1.982205e+00 9.612662e-01 5.015456e-01 4.298480e-01 4.157681e-01  
## [6] 3.894586e-01 3.729485e-01 3.711259e-01 3.690034e-01 3.668051e-01  
## [11] 3.650685e-01 3.629930e-01 3.606586e-01 3.588940e-01 3.576369e-01  
## [16] 3.560158e-01 3.555470e-01 3.543236e-01 3.520862e-01 3.511074e-01  
## [21] 3.501545e-01 3.489397e-01 3.484383e-01 3.466620e-01 3.459669e-01  
## [26] 3.451635e-01 3.446559e-01 3.433122e-01 3.429130e-01 3.421011e-01  
## [31] 3.407328e-01 3.403757e-01 3.397894e-01 3.393770e-01 3.387808e-01  
## [36] 3.382295e-01 3.378732e-01 3.378061e-01 3.360637e-01 3.357549e-01  
## [41] 3.353566e-01 3.350872e-01 3.339438e-01 3.336879e-01 3.328002e-01  
## [46] 3.323548e-01 3.319942e-01 3.311237e-01 3.309487e-01 3.301672e-01  
## [51] 3.293167e-01 3.291886e-01 3.283173e-01 3.278452e-01 3.274773e-01  
## [56] 3.266120e-01 3.262802e-01 3.254735e-01 3.252462e-01 3.246862e-01  
## [61] 3.244827e-01 3.238418e-01 3.235307e-01 3.233727e-01 3.226850e-01  
## [66] 3.219494e-01 3.215217e-01 3.213407e-01 3.201777e-01 3.197768e-01  
## [71] 3.192426e-01 3.186745e-01 3.180386e-01 3.178795e-01 3.175211e-01  
## [76] 3.168254e-01 3.165776e-01 3.164833e-01 3.158052e-01 3.154480e-01  
## [81] 3.147013e-01 3.141074e-01 3.140056e-01 3.136574e-01 3.132717e-01  
## [86] 3.131377e-01 3.124413e-01 3.118198e-01 3.115973e-01 3.112271e-01  
## [91] 3.106101e-01 3.096809e-01 3.094538e-01 3.090570e-01 3.087671e-01  
## [96] 3.084471e-01 3.081247e-01 3.077228e-01 3.073877e-01 3.069752e-01  
## [101] 3.063803e-01 3.061910e-01 3.059787e-01 3.049872e-01 3.046665e-01  
## [106] 3.044079e-01 3.040532e-01 3.035988e-01 3.031260e-01 3.027676e-01  
## [111] 3.027150e-01 3.022924e-01 3.020608e-01 3.014174e-01 3.010890e-01  
## [116] 3.006127e-01 3.004942e-01 3.001918e-01 2.998117e-01 2.989800e-01  
## [121] 2.988631e-01 2.980787e-01 2.979177e-01 2.976614e-01 2.975099e-01  
## [126] 2.969556e-01 2.966247e-01 2.962663e-01 2.957480e-01 2.954312e-01  
## [131] 2.953428e-01 2.950151e-01 2.945795e-01 2.940521e-01 2.938371e-01  
## [136] 2.934020e-01 2.933125e-01 2.928964e-01 2.922293e-01 2.921385e-01  
## [141] 2.915339e-01 2.913549e-01 2.911361e-01 2.908732e-01 2.905148e-01  
## [146] 2.903281e-01 2.897991e-01 2.890742e-01 2.889780e-01 2.884177e-01  
## [151] 2.879427e-01 2.876760e-01 2.873825e-01 2.871011e-01 2.868782e-01  
## [156] 2.864860e-01 2.860537e-01 2.859609e-01 2.852450e-01 2.848761e-01  
## [161] 2.844490e-01 2.841344e-01 2.839653e-01 2.834440e-01 2.829931e-01  
## [166] 2.827709e-01 2.823692e-01 2.820356e-01 2.818586e-01 2.811388e-01  
## [171] 2.806531e-01 2.803280e-01 2.801168e-01 2.798981e-01 2.797443e-01  
## [176] 2.794733e-01 2.787168e-01 2.785315e-01 2.783899e-01 2.780506e-01  
## [181] 2.777790e-01 2.774324e-01 2.772262e-01 2.766185e-01 2.765575e-01  
## [186] 2.759575e-01 2.758757e-01 2.754086e-01 2.752078e-01 2.748788e-01  
## [191] 2.746631e-01 2.740643e-01 2.739362e-01 2.734724e-01 2.731086e-01  
## [196] 2.728439e-01 2.723410e-01 2.721835e-01 2.719680e-01 2.718531e-01  
## [201] 2.712211e-01 2.706945e-01 2.704104e-01 2.701173e-01 2.700801e-01  
## [206] 2.696879e-01 2.694455e-01 2.689887e-01 2.688407e-01 2.682429e-01  
## [211] 2.680842e-01 2.676313e-01 2.671201e-01 2.668472e-01 2.665474e-01  
## [216] 2.662644e-01 2.657745e-01 2.655749e-01 2.651410e-01 2.650523e-01  
## [221] 2.648365e-01 2.646596e-01 2.642796e-01 2.641338e-01 2.637950e-01  
## [226] 2.633721e-01 2.631509e-01 2.627955e-01 2.625522e-01 2.619448e-01  
## [231] 2.616937e-01 2.614643e-01 2.611238e-01 2.607637e-01 2.604998e-01  
## [236] 2.598202e-01 2.595343e-01 2.594131e-01 2.592182e-01 2.587438e-01  
## [241] 2.585238e-01 2.584489e-01 2.578745e-01 2.575774e-01 2.574777e-01  
## [246] 2.568828e-01 2.566786e-01 2.564208e-01 2.561685e-01 2.556855e-01  
## [251] 2.552030e-01 2.549848e-01 2.547664e-01 2.544450e-01 2.539913e-01  
## [256] 2.536382e-01 2.535043e-01 2.530004e-01 2.527111e-01 2.523643e-01  
## [261] 2.520098e-01 2.515803e-01 2.515437e-01 2.510316e-01 2.509742e-01  
## [266] 2.505403e-01 2.502062e-01 2.498543e-01 2.495622e-01 2.494090e-01  
## [271] 2.491052e-01 2.486418e-01 2.484272e-01 2.476861e-01 2.474589e-01  
## [276] 2.470446e-01 2.469481e-01 2.462622e-01 2.461518e-01 2.457009e-01  
## [281] 2.454396e-01 2.453355e-01 2.449083e-01 2.447414e-01 2.439952e-01  
## [286] 2.438978e-01 2.437522e-01 2.433766e-01 2.426950e-01 2.423489e-01  
## [291] 2.422677e-01 2.420864e-01 2.411901e-01 2.408968e-01 2.406475e-01  
## [296] 2.404391e-01 2.402672e-01 2.396579e-01 2.391390e-01 2.391241e-01  
## [301] 2.383604e-01 2.383045e-01 2.376550e-01 2.375349e-01 2.372675e-01  
## [306] 2.370537e-01 2.364774e-01 2.363943e-01 2.358918e-01 2.356462e-01  
## [311] 2.348766e-01 2.348377e-01 2.340805e-01 2.334177e-01 2.331655e-01  
## [316] 2.328555e-01 2.324970e-01 2.322191e-01 2.321355e-01 2.310586e-01  
## [321] 2.308793e-01 2.304743e-01 2.299762e-01 2.288778e-01 2.284908e-01  
## [326] 2.282540e-01 2.278286e-01 2.276666e-01 2.269590e-01 2.264172e-01  
## [331] 2.259556e-01 2.248809e-01 2.246247e-01 2.238372e-01 2.228000e-01  
## [336] 2.221419e-01 2.214539e-01 2.208353e-01 2.202328e-01 2.184213e-01  
## [341] 2.179695e-01 2.139866e-01 1.846495e-01 7.970027e-16

summary(xc.pca)

## Importance of components:  
## PC1 PC2 PC3 PC4 PC5  
## Standard deviation 6427.2371 3.117e+03 1.626e+03 1.394e+03 1.348e+03  
## Proportion of Variance 0.1189 2.796e-02 7.610e-03 5.590e-03 5.230e-03  
## Cumulative Proportion 0.1189 1.469e-01 1.545e-01 1.601e-01 1.653e-01  
## PC6 PC7 PC8 PC9 PC10  
## Standard deviation 1.263e+03 1.209e+03 1.203e+03 1.196e+03 1.189e+03  
## Proportion of Variance 4.590e-03 4.210e-03 4.170e-03 4.120e-03 4.070e-03  
## Cumulative Proportion 1.699e-01 1.741e-01 1.782e-01 1.824e-01 1.864e-01  
## PC11 PC12 PC13 PC14 PC15  
## Standard deviation 1.184e+03 1.177e+03 1.169e+03 1163.7026 1.160e+03  
## Proportion of Variance 4.030e-03 3.990e-03 3.940e-03 0.0039 3.870e-03  
## Cumulative Proportion 1.905e-01 1.945e-01 1.984e-01 0.2023 2.062e-01  
## PC16 PC17 PC18 PC19 PC20  
## Standard deviation 1.154e+03 1.153e+03 1148.8830 1.142e+03 1.138e+03  
## Proportion of Variance 3.840e-03 3.830e-03 0.0038 3.750e-03 3.730e-03  
## Cumulative Proportion 2.100e-01 2.138e-01 0.2176 2.214e-01 2.251e-01  
## PC21 PC22 PC23 PC24 PC25  
## Standard deviation 1.135e+03 1.131e+03 1.130e+03 1.124e+03 1.122e+03  
## Proportion of Variance 3.710e-03 3.680e-03 3.670e-03 3.640e-03 3.620e-03  
## Cumulative Proportion 2.288e-01 2.325e-01 2.362e-01 2.398e-01 2.434e-01  
## PC26 PC27 PC28 PC29 PC30  
## Standard deviation 1.119e+03 1.118e+03 1.113e+03 1.112e+03 1.109e+03  
## Proportion of Variance 3.610e-03 3.590e-03 3.570e-03 3.560e-03 3.540e-03  
## Cumulative Proportion 2.470e-01 2.506e-01 2.542e-01 2.578e-01 2.613e-01  
## PC31 PC32 PC33 PC34 PC35  
## Standard deviation 1.105e+03 1.104e+03 1.102e+03 1.100e+03 1.098e+03  
## Proportion of Variance 3.510e-03 3.510e-03 3.490e-03 3.490e-03 3.470e-03  
## Cumulative Proportion 2.648e-01 2.683e-01 2.718e-01 2.753e-01 2.788e-01  
## PC36 PC37 PC38 PC39 PC40  
## Standard deviation 1.097e+03 1.096e+03 1.095e+03 1.090e+03 1.089e+03  
## Proportion of Variance 3.460e-03 3.450e-03 3.450e-03 3.420e-03 3.410e-03  
## Cumulative Proportion 2.822e-01 2.857e-01 2.891e-01 2.926e-01 2.960e-01  
## PC41 PC42 PC43 PC44 PC45  
## Standard deviation 1087.3834 1086.5098 1.083e+03 1.082e+03 1.079e+03  
## Proportion of Variance 0.0034 0.0034 3.370e-03 3.370e-03 3.350e-03  
## Cumulative Proportion 0.2994 0.3028 3.061e-01 3.095e-01 3.129e-01  
## PC46 PC47 PC48 PC49 PC50  
## Standard deviation 1.078e+03 1.076e+03 1.074e+03 1.073e+03 1070.5568  
## Proportion of Variance 3.340e-03 3.340e-03 3.320e-03 3.310e-03 0.0033  
## Cumulative Proportion 3.162e-01 3.195e-01 3.229e-01 3.262e-01 0.3295  
## PC51 PC52 PC53 PC54 PC55  
## Standard deviation 1.068e+03 1.067e+03 1.065e+03 1.063e+03 1.062e+03  
## Proportion of Variance 3.280e-03 3.280e-03 3.260e-03 3.250e-03 3.250e-03  
## Cumulative Proportion 3.327e-01 3.360e-01 3.393e-01 3.426e-01 3.458e-01  
## PC56 PC57 PC58 PC59 PC60  
## Standard deviation 1.059e+03 1.058e+03 1.055e+03 1054.6006 1.053e+03  
## Proportion of Variance 3.230e-03 3.220e-03 3.210e-03 0.0032 3.190e-03  
## Cumulative Proportion 3.490e-01 3.522e-01 3.554e-01 0.3587 3.618e-01  
## PC61 PC62 PC63 PC64 PC65  
## Standard deviation 1.052e+03 1.050e+03 1.049e+03 1.049e+03 1.046e+03  
## Proportion of Variance 3.190e-03 3.170e-03 3.170e-03 3.160e-03 3.150e-03  
## Cumulative Proportion 3.650e-01 3.682e-01 3.714e-01 3.745e-01 3.777e-01  
## PC66 PC67 PC68 PC69 PC70  
## Standard deviation 1.044e+03 1.043e+03 1.042e+03 1038.1663 1.037e+03  
## Proportion of Variance 3.140e-03 3.130e-03 3.120e-03 0.0031 3.090e-03  
## Cumulative Proportion 3.808e-01 3.840e-01 3.871e-01 0.3902 3.933e-01  
## PC71 PC72 PC73 PC74 PC75  
## Standard deviation 1.035e+03 1.033e+03 1.031e+03 1.031e+03 1.030e+03  
## Proportion of Variance 3.080e-03 3.070e-03 3.060e-03 3.060e-03 3.050e-03  
## Cumulative Proportion 3.963e-01 3.994e-01 4.025e-01 4.055e-01 4.086e-01  
## PC76 PC77 PC78 PC79 PC80  
## Standard deviation 1.027e+03 1.026e+03 1.026e+03 1.024e+03 1.023e+03  
## Proportion of Variance 3.040e-03 3.030e-03 3.030e-03 3.020e-03 3.010e-03  
## Cumulative Proportion 4.116e-01 4.147e-01 4.177e-01 4.207e-01 4.237e-01  
## PC81 PC82 PC83 PC84 PC85  
## Standard deviation 1020.4091 1.018e+03 1.018e+03 1.017e+03 1.016e+03  
## Proportion of Variance 0.0030 2.990e-03 2.980e-03 2.980e-03 2.970e-03  
## Cumulative Proportion 0.4267 4.297e-01 4.327e-01 4.357e-01 4.386e-01  
## PC86 PC87 PC88 PC89 PC90  
## Standard deviation 1.015e+03 1.013e+03 1.011e+03 1.010e+03 1.009e+03  
## Proportion of Variance 2.970e-03 2.950e-03 2.940e-03 2.940e-03 2.930e-03  
## Cumulative Proportion 4.416e-01 4.446e-01 4.475e-01 4.504e-01 4.534e-01  
## PC91 PC92 PC93 PC94 PC95  
## Standard deviation 1.007e+03 1004.1305 1003.3942 1.002e+03 1.001e+03  
## Proportion of Variance 2.920e-03 0.0029 0.0029 2.890e-03 2.880e-03  
## Cumulative Proportion 4.563e-01 0.4592 0.4621 4.650e-01 4.679e-01  
## PC96 PC97 PC98 PC99 PC100  
## Standard deviation 1.000e+03 999.08472 997.78146 996.69503 995.35740  
## Proportion of Variance 2.880e-03 0.00287 0.00287 0.00286 0.00285  
## Cumulative Proportion 4.707e-01 0.47361 0.47648 0.47934 0.48219  
## PC101 PC102 PC103 PC104 PC105  
## Standard deviation 993.42854 992.81470 992.12632 988.91157 987.87151  
## Proportion of Variance 0.00284 0.00284 0.00283 0.00281 0.00281  
## Cumulative Proportion 0.48503 0.48787 0.49070 0.49351 0.49632  
## PC106 PC107 PC108 PC109 PC110  
## Standard deviation 987.0330 985.8830 984.40946 982.87657 981.71439  
## Proportion of Variance 0.0028 0.0028 0.00279 0.00278 0.00277  
## Cumulative Proportion 0.4991 0.5019 0.50471 0.50749 0.51027  
## PC111 PC112 PC113 PC114 PC115  
## Standard deviation 981.54406 980.17362 979.42257 977.33637 976.27183  
## Proportion of Variance 0.00277 0.00277 0.00276 0.00275 0.00274  
## Cumulative Proportion 0.51304 0.51580 0.51856 0.52131 0.52406  
## PC116 PC117 PC118 PC119 PC120  
## Standard deviation 974.72735 974.34291 973.36240 972.13000 969.4334  
## Proportion of Variance 0.00273 0.00273 0.00273 0.00272 0.0027  
## Cumulative Proportion 0.52679 0.52952 0.53225 0.53497 0.5377  
## PC121 PC122 PC123 PC124 PC125  
## Standard deviation 969.0543 966.51082 965.98879 965.15780 964.66663  
## Proportion of Variance 0.0027 0.00269 0.00269 0.00268 0.00268  
## Cumulative Proportion 0.5404 0.54307 0.54575 0.54843 0.55111  
## PC126 PC127 PC128 PC129 PC130  
## Standard deviation 962.86922 961.79648 960.63423 958.95375 957.92629  
## Proportion of Variance 0.00267 0.00266 0.00266 0.00265 0.00264  
## Cumulative Proportion 0.55378 0.55644 0.55910 0.56175 0.56439  
## PC131 PC132 PC133 PC134 PC135  
## Standard deviation 957.63992 956.57709 955.16471 953.45459 952.75773  
## Proportion of Variance 0.00264 0.00263 0.00263 0.00262 0.00261  
## Cumulative Proportion 0.56703 0.56966 0.57229 0.57490 0.57751  
## PC136 PC137 PC138 PC139 PC140 PC141  
## Standard deviation 951.3468 951.0565 949.7074 947.54449 947.25003 945.28954  
## Proportion of Variance 0.0026 0.0026 0.0026 0.00258 0.00258 0.00257  
## Cumulative Proportion 0.5801 0.5827 0.5853 0.58790 0.59048 0.59306  
## PC142 PC143 PC144 PC145 PC146  
## Standard deviation 944.70931 943.99979 943.14729 941.98509 941.37995  
## Proportion of Variance 0.00257 0.00256 0.00256 0.00255 0.00255  
## Cumulative Proportion 0.59563 0.59819 0.60075 0.60330 0.60585  
## PC147 PC148 PC149 PC150 PC151  
## Standard deviation 939.66465 937.31413 937.00220 935.18535 933.64506  
## Proportion of Variance 0.00254 0.00253 0.00253 0.00252 0.00251  
## Cumulative Proportion 0.60840 0.61092 0.61345 0.61597 0.61848  
## PC152 PC153 PC154 PC155 PC156  
## Standard deviation 932.7803 931.8289 930.91633 930.19361 928.92185  
## Proportion of Variance 0.0025 0.0025 0.00249 0.00249 0.00248  
## Cumulative Proportion 0.6210 0.6235 0.62597 0.62847 0.63095  
## PC157 PC158 PC159 PC160 PC161  
## Standard deviation 927.52026 927.21918 924.89786 923.70176 922.31703  
## Proportion of Variance 0.00248 0.00247 0.00246 0.00246 0.00245  
## Cumulative Proportion 0.63342 0.63590 0.63836 0.64082 0.64327  
## PC162 PC163 PC164 PC165 PC166  
## Standard deviation 921.29699 920.74845 919.05833 917.59636 916.87582  
## Proportion of Variance 0.00244 0.00244 0.00243 0.00242 0.00242  
## Cumulative Proportion 0.64571 0.64815 0.65058 0.65300 0.65542  
## PC167 PC168 PC169 PC170 PC171  
## Standard deviation 915.57335 914.49148 913.9177 911.58384 910.00905  
## Proportion of Variance 0.00241 0.00241 0.0024 0.00239 0.00238  
## Cumulative Proportion 0.65783 0.66024 0.6626 0.66504 0.66742  
## PC172 PC173 PC174 PC175 PC176  
## Standard deviation 908.95494 908.27011 907.56094 907.06211 906.18340  
## Proportion of Variance 0.00238 0.00237 0.00237 0.00237 0.00236  
## Cumulative Proportion 0.66980 0.67217 0.67454 0.67691 0.67927  
## PC177 PC178 PC179 PC180 PC181  
## Standard deviation 903.73049 903.12970 902.67053 901.57053 900.68965  
## Proportion of Variance 0.00235 0.00235 0.00235 0.00234 0.00233  
## Cumulative Proportion 0.68163 0.68397 0.68632 0.68866 0.69099  
## PC182 PC183 PC184 PC185 PC186  
## Standard deviation 899.56596 898.89720 896.92670 896.72921 894.7837  
## Proportion of Variance 0.00233 0.00233 0.00232 0.00231 0.0023  
## Cumulative Proportion 0.69332 0.69565 0.69796 0.70028 0.7026  
## PC187 PC188 PC189 PC190 PC191  
## Standard deviation 894.5184 893.0036 892.35254 891.28595 890.58649  
## Proportion of Variance 0.0023 0.0023 0.00229 0.00229 0.00228  
## Cumulative Proportion 0.7049 0.7072 0.70947 0.71176 0.71404  
## PC192 PC193 PC194 PC195 PC196  
## Standard deviation 888.64477 888.22955 886.72565 885.54610 884.68771  
## Proportion of Variance 0.00227 0.00227 0.00226 0.00226 0.00225  
## Cumulative Proportion 0.71631 0.71858 0.72085 0.72310 0.72536  
## PC197 PC198 PC199 PC200 PC201  
## Standard deviation 883.05731 882.54659 881.84786 881.47513 879.42589  
## Proportion of Variance 0.00224 0.00224 0.00224 0.00224 0.00223  
## Cumulative Proportion 0.72760 0.72984 0.73208 0.73432 0.73654  
## PC202 PC203 PC204 PC205 PC206  
## Standard deviation 877.71856 876.79720 875.84678 875.72619 874.4545  
## Proportion of Variance 0.00222 0.00221 0.00221 0.00221 0.0022  
## Cumulative Proportion 0.73876 0.74097 0.74318 0.74539 0.7476  
## PC207 PC208 PC209 PC210 PC211  
## Standard deviation 873.6686 872.18734 871.70746 869.76936 869.25465  
## Proportion of Variance 0.0022 0.00219 0.00219 0.00218 0.00217  
## Cumulative Proportion 0.7498 0.75197 0.75416 0.75634 0.75851  
## PC212 PC213 PC214 PC215 PC216  
## Standard deviation 867.78609 866.12867 865.24357 864.27167 863.35416  
## Proportion of Variance 0.00217 0.00216 0.00215 0.00215 0.00215  
## Cumulative Proportion 0.76068 0.76284 0.76499 0.76714 0.76929  
## PC217 PC218 PC219 PC220 PC221  
## Standard deviation 861.76538 861.11841 859.71126 859.42395 858.72401  
## Proportion of Variance 0.00214 0.00213 0.00213 0.00213 0.00212  
## Cumulative Proportion 0.77143 0.77356 0.77569 0.77781 0.77994  
## PC222 PC223 PC224 PC225 PC226  
## Standard deviation 858.15042 856.91837 856.44571 855.34707 853.9759  
## Proportion of Variance 0.00212 0.00211 0.00211 0.00211 0.0021  
## Cumulative Proportion 0.78206 0.78417 0.78628 0.78839 0.7905  
## PC227 PC228 PC229 PC230 PC231  
## Standard deviation 853.2587 852.10611 851.31720 849.34779 848.53370  
## Proportion of Variance 0.0021 0.00209 0.00209 0.00208 0.00207  
## Cumulative Proportion 0.7926 0.79467 0.79676 0.79883 0.80091  
## PC232 PC233 PC234 PC235 PC236  
## Standard deviation 847.78987 846.68578 845.51827 844.66254 842.45881  
## Proportion of Variance 0.00207 0.00206 0.00206 0.00205 0.00204  
## Cumulative Proportion 0.80297 0.80504 0.80709 0.80915 0.81119  
## PC237 PC238 PC239 PC240 PC241  
## Standard deviation 841.53183 841.13881 840.50695 838.96856 838.25535  
## Proportion of Variance 0.00204 0.00204 0.00203 0.00203 0.00202  
## Cumulative Proportion 0.81323 0.81527 0.81730 0.81932 0.82135  
## PC242 PC243 PC244 PC245 PC246  
## Standard deviation 838.01260 836.14987 835.18682 834.86347 832.9343  
## Proportion of Variance 0.00202 0.00201 0.00201 0.00201 0.0020  
## Cumulative Proportion 0.82337 0.82538 0.82739 0.82939 0.8314  
## PC247 PC248 PC249 PC250 PC251  
## Standard deviation 832.27220 831.43634 830.61835 829.05215 827.48778  
## Proportion of Variance 0.00199 0.00199 0.00199 0.00198 0.00197  
## Cumulative Proportion 0.83338 0.83537 0.83736 0.83934 0.84131  
## PC252 PC253 PC254 PC255 PC256  
## Standard deviation 826.78021 826.07222 825.02992 823.55881 822.41380  
## Proportion of Variance 0.00197 0.00196 0.00196 0.00195 0.00195  
## Cumulative Proportion 0.84328 0.84524 0.84720 0.84915 0.85110  
## PC257 PC258 PC259 PC260 PC261  
## Standard deviation 8.22e+02 820.34579 819.40790 818.28326 817.13380  
## Proportion of Variance 1.94e-03 0.00194 0.00193 0.00193 0.00192  
## Cumulative Proportion 8.53e-01 0.85498 0.85691 0.85884 0.86076  
## PC262 PC263 PC264 PC265 PC266  
## Standard deviation 815.74132 815.62244 813.96213 813.77593 812.3691  
## Proportion of Variance 0.00192 0.00191 0.00191 0.00191 0.0019  
## Cumulative Proportion 0.86268 0.86459 0.86650 0.86840 0.8703  
## PC267 PC268 PC269 PC270 PC271  
## Standard deviation 811.28565 810.14490 809.19772 808.70092 807.71591  
## Proportion of Variance 0.00189 0.00189 0.00188 0.00188 0.00188  
## Cumulative Proportion 0.87220 0.87409 0.87597 0.87785 0.87973  
## PC272 PC273 PC274 PC275 PC276  
## Standard deviation 806.21314 805.51739 803.11445 802.37787 8.01e+02  
## Proportion of Variance 0.00187 0.00187 0.00186 0.00185 1.85e-03  
## Cumulative Proportion 0.88160 0.88347 0.88532 0.88718 8.89e-01  
## PC277 PC278 PC279 PC280 PC281  
## Standard deviation 800.72148 798.49763 798.13957 796.67753 795.83013  
## Proportion of Variance 0.00185 0.00184 0.00183 0.00183 0.00182  
## Cumulative Proportion 0.89087 0.89270 0.89454 0.89637 0.89819  
## PC282 PC283 PC284 PC285 PC286 PC287  
## Standard deviation 795.49283 794.10760 793.56622 791.1468 790.8311 790.3589  
## Proportion of Variance 0.00182 0.00181 0.00181 0.0018 0.0018 0.0018  
## Cumulative Proportion 0.90001 0.90182 0.90364 0.9054 0.9072 0.9090  
## PC288 PC289 PC290 PC291 PC292  
## Standard deviation 789.14105 786.93102 785.80873 785.54538 784.95766  
## Proportion of Variance 0.00179 0.00178 0.00178 0.00178 0.00177  
## Cumulative Proportion 0.91083 0.91261 0.91439 0.91616 0.91794  
## PC293 PC294 PC295 PC296 PC297  
## Standard deviation 782.05124 781.10025 780.29212 779.61622 779.05885  
## Proportion of Variance 0.00176 0.00176 0.00175 0.00175 0.00175  
## Cumulative Proportion 0.91970 0.92145 0.92321 0.92496 0.92670  
## PC298 PC299 PC300 PC301 PC302  
## Standard deviation 777.08340 775.40078 775.35255 772.87626 772.69474  
## Proportion of Variance 0.00174 0.00173 0.00173 0.00172 0.00172  
## Cumulative Proportion 0.92844 0.93017 0.93190 0.93362 0.93534  
## PC303 PC304 PC305 PC306 PC307  
## Standard deviation 770.58903 770.19950 769.3325 768.6392 766.77070  
## Proportion of Variance 0.00171 0.00171 0.0017 0.0017 0.00169  
## Cumulative Proportion 0.93705 0.93875 0.9405 0.9422 0.94385  
## PC308 PC309 PC310 PC311 PC312  
## Standard deviation 766.50114 764.87184 764.07549 761.58017 761.45398  
## Proportion of Variance 0.00169 0.00168 0.00168 0.00167 0.00167  
## Cumulative Proportion 0.94554 0.94723 0.94891 0.95058 0.95224  
## PC313 PC314 PC315 PC316 PC317  
## Standard deviation 758.99870 756.84958 756.03198 755.02663 753.86420  
## Proportion of Variance 0.00166 0.00165 0.00165 0.00164 0.00164  
## Cumulative Proportion 0.95390 0.95555 0.95720 0.95884 0.96047  
## PC318 PC319 PC320 PC321 PC322  
## Standard deviation 752.96305 752.69211 749.20016 748.61892 747.30583  
## Proportion of Variance 0.00163 0.00163 0.00162 0.00161 0.00161  
## Cumulative Proportion 0.96210 0.96373 0.96535 0.96696 0.96857  
## PC323 PC324 PC325 PC326 PC327  
## Standard deviation 745.6907 742.12923 740.87419 740.10633 738.72716  
## Proportion of Variance 0.0016 0.00159 0.00158 0.00158 0.00157  
## Cumulative Proportion 0.9702 0.97176 0.97334 0.97491 0.97648  
## PC328 PC329 PC330 PC331 PC332  
## Standard deviation 738.20175 735.90746 734.15054 732.65389 729.16925  
## Proportion of Variance 0.00157 0.00156 0.00155 0.00154 0.00153  
## Cumulative Proportion 0.97805 0.97961 0.98116 0.98271 0.98424  
## PC333 PC334 PC335 PC336 PC337  
## Standard deviation 728.33855 725.78514 722.4219 720.28823 718.05731  
## Proportion of Variance 0.00153 0.00152 0.0015 0.00149 0.00148  
## Cumulative Proportion 0.98576 0.98728 0.9888 0.99027 0.99176  
## PC338 PC339 PC340 PC341 PC342  
## Standard deviation 716.05153 714.09795 708.22422 706.75925 693.84489  
## Proportion of Variance 0.00148 0.00147 0.00144 0.00144 0.00139  
## Cumulative Proportion 0.99323 0.99470 0.99615 0.99758 0.99897  
## PC343 PC344  
## Standard deviation 598.72011 2.584e-12  
## Proportion of Variance 0.00103 0.000e+00  
## Cumulative Proportion 1.00000 1.000e+00

Xc.just.sex <- model.matrix(~ Sex, data=xc\_mat)  
  
xc.pca.and.sex <- cbind(xc.pca, Xc.just.sex)

## Warning in cbind(xc.pca, Xc.just.sex): number of rows of result is not a  
## multiple of vector length (arg 1)

gene\_info <- read.csv('gene\_info.csv')  
SNP\_info <- read.csv('SNP\_info.csv')  
  
# Define our function to run GWAS and get p-values  
pval\_calculator\_w\_covars <- function(pheno\_input, xa\_input, xd\_input, xz\_input){  
 n\_samples <- length(xa\_input) # calculate your number of samples  
 X\_mx <- cbind(rep(1,length(xa\_input)),xa\_input, xd\_input, xz\_input) #create your X matrix under H1  
   
 MLE\_beta <- ginv(t(X\_mx) %\*% X\_mx) %\*% t(X\_mx) %\*% pheno\_input #calculate your MLE of the betas  
   
 x\_h0 = cbind(rep(1,length(xa\_input)), xz\_input) #calculate your x under H0  
 MLE\_h0 = ginv(t(x\_h0) %\*% x\_h0) %\*% t(x\_h0) %\*% pheno\_input #calculate your MLE under h0  
 y\_hat\_0 = x\_h0 %\*% MLE\_h0 #calculate y\_hat under the null hypothesis  
 y\_hat\_1 = X\_mx%\*% MLE\_beta #calculate y\_hat under H1  
   
 SSE\_theta\_0 = sum((pheno\_input-y\_hat\_0)^2) #calculate SSE under null   
 SSE\_theta\_1 = sum((pheno\_input-y\_hat\_1)^2) #calculate SSE under H1  
   
 df\_M <- 2  
 df\_E <- n\_samples - 4   
   
 numerator <- (SSE\_theta\_0-SSE\_theta\_1) / df\_M #calculate your F statistic  
 denom <- SSE\_theta\_1 / df\_E  
 Fstatistic <-numerator / denom  
   
 # to check if it is correct   
 pval <- pf(Fstatistic, df\_M, df\_E,lower.tail = FALSE) #calculate your p value and return it  
 return(pval)  
}  
  
# Run the functions  
  
results.1 <- lapply(1:ncol(xa\_mat), function(column.counter){  
 data.table(pval\_calculator\_w\_covars(pheno\_input = sim\_pheno\_mx[,1],  
 xa\_input = as.matrix(xa\_mat[, column.counter]),  
 xd\_input = as.matrix(xd\_mat[, column.counter]),  
 xz\_input = as.matrix(xc\_mat1)))  
 }) %>% rbindlist() %>% mutate(p=V1, index = 1:ncol(xa\_mat))  
  
indices1 <- which(p.adjust(results.1$p, method = "bonferroni") <= 0.05)  
hits1 <- sort(SNP\_info$position[indices1])  
length(hits1)

## [1] 71

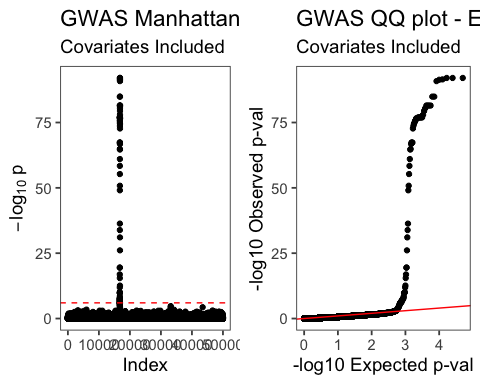
results.2 <- lapply(1:ncol(xa\_mat), function(column.counter){  
 data.table(pval\_calculator\_w\_covars(pheno\_input = sim\_pheno\_mx[,2],  
 xa\_input = as.matrix(xa\_mat[, column.counter]),  
 xd\_input = as.matrix(xd\_mat[, column.counter]),  
 xz\_input = as.matrix(xc\_mat1)))  
 }) %>% rbindlist() %>% mutate(p=V1, index = 1:ncol(xa\_mat))  
  
indices2 <- which(p.adjust(results.2$p, method = "bonferroni") <= 0.05)  
hits2 <- sort(SNP\_info$position[indices2])  
hits2

## [1] 42889467 42893842 42896182 42903013 42904736 42907726 42911587 42916172  
## [9] 42919954 42924756 42938645 42942927 42946612 42949278 42952810 42953822  
## [17] 42954621 42956454 42957216 42962100 42964461 42969161 42972496 42973585  
## [25] 42977844 43108015 98486048

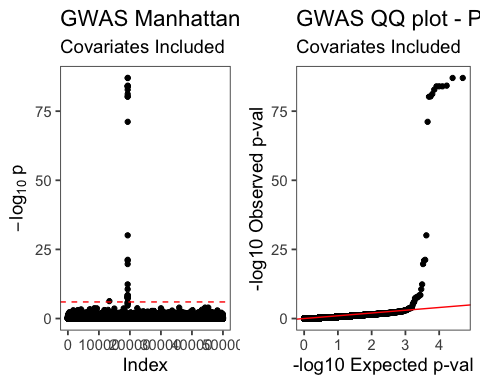
results.3 <- lapply(1:ncol(xa\_mat), function(column.counter){  
 data.table(pval\_calculator\_w\_covars(pheno\_input = sim\_pheno\_mx[,3],  
 xa\_input = as.matrix(xa\_mat[, column.counter]),  
 xd\_input = as.matrix(xd\_mat[, column.counter]),  
 xz\_input = as.matrix(xc\_mat1)))  
 }) %>% rbindlist() %>% mutate(p=V1, index = 1:ncol(xa\_mat))  
  
indices3 <- which(p.adjust(results.3$p, method = "bonferroni") <= 0.05)  
hits3 <- sort(SNP\_info$position[indices3])  
hits3

## [1] 1524250 1604317 1614929 1780619 1785391 1789021 1792207 1794224 1798982  
## [10] 1804934 1806080 1806559 1807707 1808392 1809480 1810194 1810803 1811984  
## [19] 1812168 1813518 1816100 1819604 1820639 1824144 1825190 1826990 1828065  
## [28] 1829958 1831143 1831554 1832328 1832761 1835286 1836231 1837572 1839643  
## [37] 1842605 1842970 1845028 1845980 1846938 1848043 1849190 1850033 1852576  
## [46] 1854320 1854638 1855229 1856231 1857300 1858216 1858666 1859725 1860371  
## [55] 1861583 1862683 1862993 1864524 1864887 1865581 1866730 1867504 1868123  
## [64] 1868857 1869663 1870120 1871282 1872310 1873062 1873933 1874763 1883987  
## [73] 1892929 1895185 1897411 1898338 1899224 1901135 1902367 1904506 1911549  
## [82] 1916400 1919081 1921599 1923514 1926150 1926998 1928485 1929366

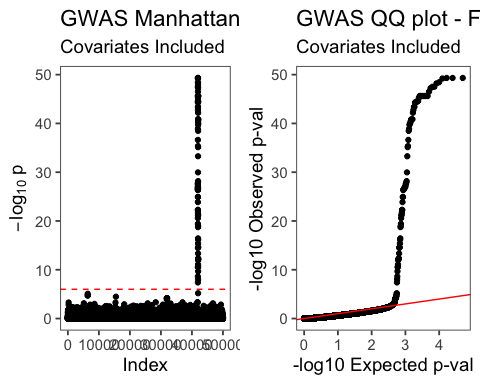
results.4 <- lapply(1:ncol(xa\_mat), function(column.counter){  
 data.table(pval\_calculator\_w\_covars(pheno\_input = sim\_pheno\_mx[,4],  
 xa\_input = as.matrix(xa\_mat[, column.counter]),  
 xd\_input = as.matrix(xd\_mat[, column.counter]),  
 xz\_input = as.matrix(xc\_mat1)))  
 }) %>% rbindlist() %>% mutate(p=V1, index = 1:ncol(xa\_mat))  
  
indices4 <- which(p.adjust(results.4$p, method = "bonferroni") <= 0.05)  
hits4 <- sort(SNP\_info$position[indices4])  
  
  
results.5 <- lapply(1:ncol(xa\_mat), function(column.counter){  
 data.table(pval\_calculator\_w\_covars(pheno\_input = sim\_pheno\_mx[,5],  
 xa\_input = as.matrix(xa\_mat[, column.counter]),  
 xd\_input = as.matrix(xd\_mat[, column.counter]),  
 xz\_input = as.matrix(xc\_mat1)))  
 }) %>% rbindlist() %>% mutate(p=V1, index = 1:ncol(xa\_mat))  
  
indices5 <- which(p.adjust(results.5$p, method = "bonferroni") <= 0.05)  
hits5 <- sort(SNP\_info$position[indices5])  
  
  
# Manhattan Plot  
  
## phenotype 1 - ERAP2  
my.alpha = 0.05/ncol(xa\_mat)  
man1 <- ggplot(results.1, aes(x = index, y = -log10(p))) +  
 geom\_point() +   
 geom\_hline(yintercept = -log10(my.alpha), color = 'red', lty = 2) +  
 labs(x = 'Index', y = expression(-log[10]~p),   
 title = 'GWAS Manhattan Plot- ERAP2', subtitle='Covariates Included')  
  
## phenotype 2 - PEX6  
my.alpha = 0.05/ncol(xa\_mat)  
man2 <- ggplot(results.2, aes(x = index, y = -log10(p))) +  
 geom\_point() +   
 geom\_hline(yintercept = -log10(my.alpha), color = 'red', lty = 2) +  
 labs(x = 'Index', y = expression(-log[10]~p),   
 title = 'GWAS Manhattan Plot - PEX6', subtitle='Covariates Included')  
  
## phenotype 3 - FAHD1  
my.alpha = 0.05/ncol(xa\_mat)  
man3 <- ggplot(results.3, aes(x = index, y = -log10(p))) +  
 geom\_point() +   
 geom\_hline(yintercept = -log10(my.alpha), color = 'red', lty = 2) +  
 labs(x = 'Index', y = expression(-log[10]~p),   
 title = 'GWAS Manhattan Plot - FAHD1', subtitle='Covariates Included')  
  
## phenotype 4 - GFM1  
my.alpha = 0.05/ncol(xa\_mat)  
man4 <- ggplot(results.4, aes(x = index, y = -log10(p))) +  
 geom\_point() +   
 geom\_hline(yintercept = -log10(my.alpha), color = 'red', lty = 2) +  
 labs(x = 'Index', y = expression(-log[10]~p),   
 title = 'GWAS Manhattan Plot - GFM1', subtitle='Covariates Included')  
  
## phenotype 5 - MARCHF7  
my.alpha = 0.05/ncol(xa\_mat)  
man5 <- ggplot(results.5, aes(x = index, y = -log10(p))) +  
 geom\_point() +   
 geom\_hline(yintercept = -log10(my.alpha), color = 'red', lty = 2) +  
 labs(x = 'Index', y = expression(-log[10]~p),   
 title = 'GWAS Manhattan Plot - MARCHF7', subtitle='Covariates Included')  
  
  
# QQ plot  
## phenotype 1 - ERAP2  
observed\_pvals = sort(results.1$p)  
expected\_pvals = qunif(seq(0, 1, length.out = length(observed\_pvals) + 2), min = 0, max = 1) ## Generate expected values. Note that we are using length+2  
expected\_pvals = expected\_pvals[expected\_pvals != 0 & expected\_pvals != 1] ## Remove the two extra values since they are 0 and 1  
  
p\_df = data.frame(observed = -log10(observed\_pvals),  
 expected = -log10(expected\_pvals))  
  
qq <- ggplot(p\_df, aes(x = expected, y = observed)) +  
 geom\_point() +  
 geom\_abline(intercept = 0, slope = 1, color = 'red') +  
 labs(x = '-log10 Expected p-val',  
 y = '-log10 Observed p-val',  
 title = 'GWAS QQ plot - ERAP2',  
 subtitle = 'Covariates Included')  
grid.arrange(man1,qq, ncol=2)



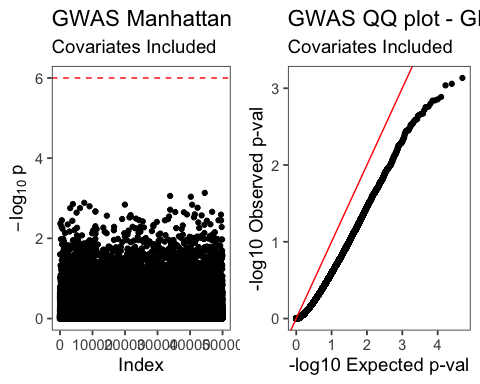
## phenotype 2 - PEX6  
observed\_pvals = sort(results.2$p)  
expected\_pvals = qunif(seq(0, 1, length.out = length(observed\_pvals) + 2), min = 0, max = 1) ## Generate expected values. Note that we are using length+2  
expected\_pvals = expected\_pvals[expected\_pvals != 0 & expected\_pvals != 1] ## Remove the two extra values since they are 0 and 1  
  
p\_df = data.frame(observed = -log10(observed\_pvals),  
 expected = -log10(expected\_pvals))  
  
qq <- ggplot(p\_df, aes(x = expected, y = observed)) +  
 geom\_point() +  
 geom\_abline(intercept = 0, slope = 1, color = 'red') +  
 labs(x = '-log10 Expected p-val',  
 y = '-log10 Observed p-val',  
 title = 'GWAS QQ plot - PEX6',  
 subtitle = 'Covariates Included')  
grid.arrange(man2,qq, ncol=2)



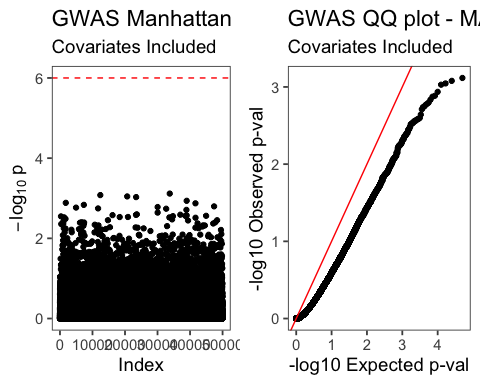
## phenotype 3 - FAHD1  
observed\_pvals = sort(results.3$p)  
expected\_pvals = qunif(seq(0, 1, length.out = length(observed\_pvals) + 2), min = 0, max = 1) ## Generate expected values. Note that we are using length+2  
expected\_pvals = expected\_pvals[expected\_pvals != 0 & expected\_pvals != 1] ## Remove the two extra values since they are 0 and 1  
  
p\_df = data.frame(observed = -log10(observed\_pvals),  
 expected = -log10(expected\_pvals))  
  
qq <- ggplot(p\_df, aes(x = expected, y = observed)) +  
 geom\_point() +  
 geom\_abline(intercept = 0, slope = 1, color = 'red') +  
 labs(x = '-log10 Expected p-val',  
 y = '-log10 Observed p-val',  
 title = 'GWAS QQ plot - FAHD1',  
 subtitle = 'Covariates Included')  
grid.arrange(man3,qq, ncol=2)



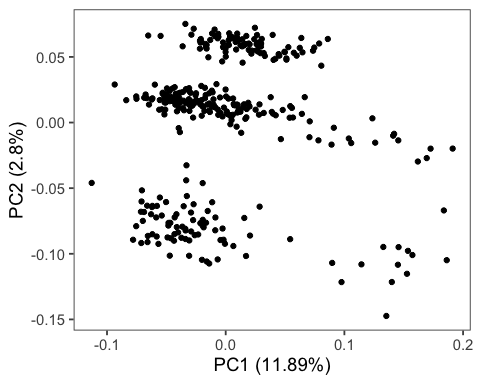
## phenotype 4 - GFM1  
observed\_pvals = sort(results.4$p)  
expected\_pvals = qunif(seq(0, 1, length.out = length(observed\_pvals) + 2), min = 0, max = 1) ## Generate expected values. Note that we are using length+2  
expected\_pvals = expected\_pvals[expected\_pvals != 0 & expected\_pvals != 1] ## Remove the two extra values since they are 0 and 1  
  
p\_df = data.frame(observed = -log10(observed\_pvals),  
 expected = -log10(expected\_pvals))  
  
qq <- ggplot(p\_df, aes(x = expected, y = observed)) +  
 geom\_point() +  
 geom\_abline(intercept = 0, slope = 1, color = 'red') +  
 labs(x = '-log10 Expected p-val',  
 y = '-log10 Observed p-val',  
 title = 'GWAS QQ plot - GFM1',  
 subtitle = 'Covariates Included')  
grid.arrange(man4,qq, ncol=2)



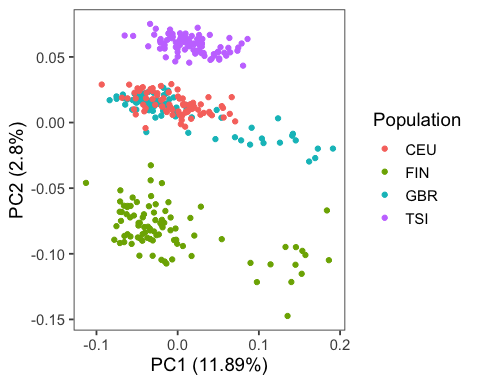
## phenotype 5 - MARCHF7  
observed\_pvals = sort(results.5$p)  
expected\_pvals = qunif(seq(0, 1, length.out = length(observed\_pvals) + 2), min = 0, max = 1) ## Generate expected values. Note that we are using length+2  
expected\_pvals = expected\_pvals[expected\_pvals != 0 & expected\_pvals != 1] ## Remove the two extra values since they are 0 and 1  
  
p\_df = data.frame(observed = -log10(observed\_pvals),  
 expected = -log10(expected\_pvals))  
  
qq <- ggplot(p\_df, aes(x = expected, y = observed)) +  
 geom\_point() +  
 geom\_abline(intercept = 0, slope = 1, color = 'red') +  
 labs(x = '-log10 Expected p-val',  
 y = '-log10 Observed p-val',  
 title = 'GWAS QQ plot - MARCHF7',  
 subtitle = 'Covariates Included')  
grid.arrange(man5,qq, ncol=2)



library(ggfortify)  
xc.pca<- prcomp(xa\_mat %\*% t(xa\_mat))  
  
autoplot(xc.pca)



autoplot(xc.pca, data = xc\_mat, colour = "Population")



autoplot(xc.pca, data = xc\_mat, colour = "Sex")

