

Starling-May18

Projects/Katarina Stuart/KStuart.Starling-Aug18/Sv6\_Morphology/Analysis/2021.05.25.Baypass\_7pheno

PDF Version generated by

Katarina Stuart (z5188231@ad.unsw.edu.au)

on

Feb 12, 2022 @03:42 PM AEDT

## Table of Contents

2021.05.25.Baypass_7pheno .....	2
---------------------------------	---



## BayPass Analysis 212 ind

Consider for if I change in the future:

Use just AUX model

Calibrate XtX and only keep outliers above this threshold

Do morph associated analysis on that.

```
cd /srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno
```

```
PLINK=/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/GenoPhenoVariation/sv6_PhenoFilter_14pop_Tarsus_maf005_miss50_r
```

```
module add vcftools/0.1.16
```

```
module load plink/1.90b6.7
```

```
mkdir plink_files
```

```
for PHENO in mass tarsus head beak wing spleen heart
```

```
do
```

```
plink --file $PLINK --keep ../GCTA212_7pheno/GCTA_${PHENO}_subset.txt --allow-extra-chr --freq counts --family --out plink_files/plink_${PHENO}
```

```
done
```

change file structure to extract what I need

```
for PHENO in mass tarsus head beak wing spleen heart
```

```
do
```

```
tail -n +2 plink_files/plink_${PHENO}.frq.strat | awk '{ $9 = $8 - $7 } 1' | awk '{print $7,$9}' | tr "\n" " " | sed 's/ \n/28; P; D/'> genofile_baypass_${PHENO}.txt
```

```
done
```

## Baypass with PCA axis

Baypass runs:

#pop averages efile is alphabetical order (I think...) corresponding to the geno file stats order.

```
#!/bin/bash
```

```
#PBS -N 2021-05-25.baypass_7pheno.pbs
```

```
#PBS -l nodes=1:ppn=16
```

```
#PBS -l mem=120gb
```

```
#PBS -l walltime=48:00:00
```

```
#PBS -j oe
```

```
#PBS -M katarina.stuart@student.unsw.edu.au
```

```
#PBS -m ae
```

```
#PBS -J 01-07
```

```
module load baypass
```

```
cd /srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno
```

```
echo "working with pheno:"
```

```
sed "${PBS_ARRAY_INDEX}q;d" 7pheno_list.txt
```

```
PHENO=$(sed "${PBS_ARRAY_INDEX}q;d" 7pheno_list.txt)
```

```
mkdir Pheno_${PHENO}
```

```
cd Pheno_${PHENO}
```

```
g_bypass -npop 14 -gfile ../genofile_bypass_${PHENO}.txt -outprefix Pheno_${PHENO}_anacore -nthreads 16
g_bypass -npop 14 -gfile ../genofile_bypass_${PHENO}.txt -efile ../bypass_pheno_${PHENO}.txt -scalecov -auxmodel -nthreads 16 -omegafile Pheno_${PHENO}_anacore_mat_omeg
g_bypass -npop 14 -gfile ../genofile_bypass_${PHENO}.txt -efile ../bypass_pheno_${PHENO}.txt -scalecov -nthreads 16 -omegafile Pheno_${PHENO}_anacore_mat_omega.out -out
```

**Running in R to make the anapod data: (haven't redone this bit as it shouldn't be different).**

```
module load R/3.6.3
R
setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212/Pheno212_axis1")
source("/apps/bypass/2.1/utis/bypass_utils.R")
library("ape")
library("corrplot")

omega=as.matrix(read.table("Pheno212_axis1_anacore_mat_omega.out"))
anacore.snp.res=read.table("Pheno212_axis1_anacore_summary_pi_xtx.out",h=T)
pdf("Sv6_mass_XtXdiff.pdf")
plot(anacore.snp.res$M_XtX)
dev.off()

pdf("Sv6_mass_anacore_pval.pdf")
hist(10**(-1*anacore.snp.res$ACC_P),freq=F,breaks=50)
abline(h=1)
layout(matrix(1:2,2,1))
plot(anacore.snp.res$M_XtX)
plot(anacore.snp.res$ACC_P,ylab="XtX P-value (-log10 scale)")
abline(h=3,lty=2) #0.001 p--value threshold
dev.off()

pi.beta.coef=read.table("Pheno212_axis1_anacore_summary_beta_params.out",h=T)$Mean
bta14.data<-geno2YN("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212/genofile_bypass.txt")
simu.bta<-simulate.bypass(omega.mat=omega, nsnp=20000, sample.size=bta14.data$NN, beta.pi=pi.beta.coef,pi.maf=0,suffix="btapods")
coredir="/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212/Pheno212_axis1/"
file.rename("G.btapods", paste(coredir, "G.btapods", sep=""))
```

## Covis

```
setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212_7pheno/Pheno_mass")
covis.snp.res.mass=read.table("Pheno_mass_anacovis_scaled_summary_betai_reg.out",h=T)
graphics.off()
pdf("Pheno_mass_covis_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covis.snp.res.mass$BF.dB.,xlab="SNP",ylab="BFis (in dB)")
abline(h=20, col="red")
plot(covis.snp.res.mass$eBPis,xlab="SNP",ylab="eBPis")
plot(covis.snp.res.mass$Beta_is,xlab="SNP",ylab=expression(beta~"coefficient"))
dev.off()

setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212_7pheno/Pheno_tarsus")
covis.snp.res.tarsus=read.table("Pheno_tarsus_anacovis_scaled_summary_betai_reg.out",h=T)
graphics.off()
pdf("Pheno_tarsus_covis_scaled.pdf")
```

```

layout(matrix(1:3,3,1))
plot(covis.snp.res.tarsus$BF.dB.,xlab="SNP",ylab="BFis (in dB)")
abline(h=20, col="red")
plot(covis.snp.res.tarsus$eBPis,xlab="SNP",ylab="eBPis")
plot(covis.snp.res.tarsus$Beta_is,xlab="SNP",ylab=expression(beta~"coefficient"))
dev.off()

```

```

setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_head")
covis.snp.res.head=read.table("Pheno_head_anacovis_scaled_summary_betai_reg.out",h=T)
graphics.off()
pdf("Pheno_head_covis_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covis.snp.res.head$BF.dB.,xlab="SNP",ylab="BFis (in dB)")
abline(h=20, col="red")
plot(covis.snp.res.head$eBPis,xlab="SNP",ylab="eBPis")
plot(covis.snp.res.head$Beta_is,xlab="SNP",ylab=expression(beta~"coefficient"))
dev.off()

```

```

setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_beak")
covis.snp.res.beak=read.table("Pheno_beak_anacovis_scaled_summary_betai_reg.out",h=T)
graphics.off()
pdf("Pheno_beak_covis_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covis.snp.res.beak$BF.dB.,xlab="SNP",ylab="BFis (in dB)")
abline(h=20, col="red")
plot(covis.snp.res.beak$eBPis,xlab="SNP",ylab="eBPis")
plot(covis.snp.res.beak$Beta_is,xlab="SNP",ylab=expression(beta~"coefficient"))
dev.off()

```

```

setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_wing")
covis.snp.res.wing=read.table("Pheno_wing_anacovis_scaled_summary_betai_reg.out",h=T)
graphics.off()
pdf("Pheno_wing_covis_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covis.snp.res.wing$BF.dB.,xlab="SNP",ylab="BFis (in dB)")
abline(h=20, col="red")
plot(covis.snp.res.wing$eBPis,xlab="SNP",ylab="eBPis")
plot(covis.snp.res.wing$Beta_is,xlab="SNP",ylab=expression(beta~"coefficient"))
dev.off()

```

```

setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_spleen")
covis.snp.res.spleen=read.table("Pheno_spleen_anacovis_scaled_summary_betai_reg.out",h=T)
graphics.off()
pdf("Pheno_spleen_covis_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covis.snp.res.spleen$BF.dB.,xlab="SNP",ylab="BFis (in dB)")
abline(h=20, col="red")
plot(covis.snp.res.spleen$eBPis,xlab="SNP",ylab="eBPis")
plot(covis.snp.res.spleen$Beta_is,xlab="SNP",ylab=expression(beta~"coefficient"))
dev.off()

```

```

setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_heart")
covis.snp.res.heart=read.table("Pheno_heart_anacovis_scaled_summary_betai_reg.out",h=T)
graphics.off()
pdf("Pheno_heart_covis_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covis.snp.res.heart$BF.dB.,xlab="SNP",ylab="BFis (in dB)")
abline(h=20, col="red")
plot(covis.snp.res.heart$eBPis,xlab="SNP",ylab="eBPis")
plot(covis.snp.res.heart$Beta_is,xlab="SNP",ylab=expression(beta~"coefficient"))
dev.off()

```

# Covaux

```
setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_mass")
covaux.snp.res.mass=read.table("Pheno_mass_anacovaux_scaled_summary_betai.out",h=T)
covaux.snp.xtx.mass=read.table("Pheno_mass_anacovaux_scaled_summary_pi_xtx.out",h=T)$M_XtX
graphics.off()
pdf("Pheno_mass_covaux_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covaux.snp.res.mass$BF.dB.,xlab="Mass",ylab="BFmc (in dB)")
abline(h=20, col="red")
plot(covaux.snp.res.mass$M_Beta,xlab="SNP",ylab=expression(beta~"coefficient"))
plot(covaux.snp.xtx.mass, xlab="SNP",ylab="XtX corrected for SMS")
dev.off()
```

```
setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_tarsus")
covaux.snp.res.tarsus=read.table("Pheno_tarsus_anacovaux_scaled_summary_betai.out",h=T)
covaux.snp.xtx.tarsus=read.table("Pheno_tarsus_anacovaux_scaled_summary_pi_xtx.out",h=T)$M_XtX
graphics.off()
pdf("Pheno_tarsus_covaux_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covaux.snp.res.tarsus$BF.dB.,xlab="Tarsus",ylab="BFmc (in dB)")
abline(h=20, col="red")
plot(covaux.snp.res.tarsus$M_Beta,xlab="SNP",ylab=expression(beta~"coefficient"))
plot(covaux.snp.xtx.tarsus, xlab="SNP",ylab="XtX corrected for SMS")
dev.off()
```

```
setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_head")
covaux.snp.res.head=read.table("Pheno_head_anacovaux_scaled_summary_betai.out",h=T)
covaux.snp.xtx.head=read.table("Pheno_head_anacovaux_scaled_summary_pi_xtx.out",h=T)$M_XtX
graphics.off()
pdf("Pheno_head_covaux_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covaux.snp.res.head$BF.dB.,xlab="head",ylab="BFmc (in dB)")
abline(h=20, col="red")
plot(covaux.snp.res.head$M_Beta,xlab="SNP",ylab=expression(beta~"coefficient"))
plot(covaux.snp.xtx.head, xlab="SNP",ylab="XtX corrected for SMS")
dev.off()
```

```
setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_beak")
covaux.snp.res.beak=read.table("Pheno_beak_anacovaux_scaled_summary_betai.out",h=T)
covaux.snp.xtx.beak=read.table("Pheno_beak_anacovaux_scaled_summary_pi_xtx.out",h=T)$M_XtX
graphics.off()
pdf("Pheno_beak_covaux_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covaux.snp.res.beak$BF.dB.,xlab="beak",ylab="BFmc (in dB)")
abline(h=20, col="red")
plot(covaux.snp.res.beak$M_Beta,xlab="SNP",ylab=expression(beta~"coefficient"))
plot(covaux.snp.xtx.beak, xlab="SNP",ylab="XtX corrected for SMS")
dev.off()
```

```
setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/Pheno_wing")
covaux.snp.res.wing=read.table("Pheno_wing_anacovaux_scaled_summary_betai.out",h=T)
covaux.snp.xtx.wing=read.table("Pheno_wing_anacovaux_scaled_summary_pi_xtx.out",h=T)$M_XtX
graphics.off()
pdf("Pheno_wing_covaux_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covaux.snp.res.wing$BF.dB.,xlab="wing",ylab="BFmc (in dB)")
abline(h=20, col="red")
plot(covaux.snp.res.wing$M_Beta,xlab="SNP",ylab=expression(beta~"coefficient"))
plot(covaux.snp.xtx.wing, xlab="SNP",ylab="XtX corrected for SMS")
dev.off()
```

```
setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212_7pheno/Pheno_spleen")
covaux.snp.res.spleen=read.table("Pheno_spleen_anacovaux_scaled_summary_betai.out",h=T)
covaux.snp.xtx.spleen=read.table("Pheno_spleen_anacovaux_scaled_summary_pi_xtx.out",h=T)$M_XtX
graphics.off()
pdf("Pheno_spleen_covaux_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covaux.snp.res.spleen$BF.dB.,xlab="spleen",ylab="BFmc (in dB)")
abline(h=20, col="red")
plot(covaux.snp.res.spleen$M_Beta,xlab="SNP",ylab=expression(beta~"coefficient"))
plot(covaux.snp.xtx.spleen, xlab="SNP",ylab="XtX corrected for SMS")
dev.off()

setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212_7pheno/Pheno_heart")
covaux.snp.res.heart=read.table("Pheno_heart_anacovaux_scaled_summary_betai.out",h=T)
covaux.snp.xtx.heart=read.table("Pheno_heart_anacovaux_scaled_summary_pi_xtx.out",h=T)$M_XtX
graphics.off()
pdf("Pheno_heart_covaux_scaled.pdf")
layout(matrix(1:3,3,1))
plot(covaux.snp.res.heart$BF.dB.,xlab="heart",ylab="BFmc (in dB)")
abline(h=20, col="red")
plot(covaux.snp.res.heart$M_Beta,xlab="SNP",ylab=expression(beta~"coefficient"))
plot(covaux.snp.xtx.heart, xlab="SNP",ylab="XtX corrected for SMS")
dev.off()
```

#### Adding corresponding line numbers to the SNP list file that match Bypass snp line numbers

```
cd /srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212_7pheno
SNPLIST=/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/GenoPhenoVariation/sv6_PhenoFilter_14pop_Tarsus_maf005_miss5
tail -n +2 ${SNPLIST} | awk '{print $0,NR}' > Stacks_filtered_bypass_128217snps_snplist_numbered.txt
```

#### Annotation file:

```
GFF=/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv3_Genome/Sv3.4_GenomeAnnotation/annotation/2020-10-22.vAUMAKER/results_run3_nopred/m
GFF_GO=/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv3_Genome/Sv3.4_GenomeAnnotation/annotation/2020-10-22.vAUMAKER/results_run3_nopr
```

#### Filtering the data sets for SNPS above BFmc threshold and grab gene list

```
module load bedtools/2.27.1

for PHENO in mass tarsus head beak wing spleen heart;
do
echo "${PHENO} about to do doooowwwwnn"
cd Pheno_${PHENO}/
cat Pheno_${PHENO}_anacovaux_scaled_summary_betai.out | awk '$6>20' > outliers_${PHENO}_anacovaux_BF20.txt
cat Pheno_${PHENO}_anacovis_scaled_summary_betai_reg.out | awk '$5>20' > outliers_${PHENO}_anacovis_BF20.txt

awk '{print $2}' outliers_${PHENO}_anacovis_BF20.txt | grep -f - outliers_${PHENO}_anacovaux_BF20.txt > outliers_${PHENO}_both_BF20.txt

awk 'FNR==NR{a[$2];next} (($4) in a)' outliers_${PHENO}_both_BF20.txt ../Stacks_filtered_bypass_128217snps_snplist_numbered.txt > outliers_${PHE
awk 'BEGIN{FS=OFS="\t"}{if(NR>1){ print $1,$2-1,$2 }}' outliers_${PHENO}_both_BF20_SNPlist.txt | bedtools intersect -wb -a $GFF -b stdin | awk '$3=="
cd ../
done
```

#### Grabbing gene list in combined data set for GO enrichment analysis:

```
for PHENO in mass tarsus head beak wing spleen heart;
do
echo "${PHENO} about to do doooowwwwnn"
cd Pheno_${PHENO}/

awk 'FNR==NR{a[$2];next} (($4) in a)' outliers_${PHENO}_anacovaux_BF20.txt ../Stacks_filtered_bypass_128217snps_snplist_numbered.txt > outliers_${PHENO}_anacovaux_BF20_SNPlist.txt
awk 'FNR==NR{a[$2];next} (($4) in a)' outliers_${PHENO}_anacovis_BF20.txt ../Stacks_filtered_bypass_128217snps_snplist_numbered.txt > outliers_${PHENO}_anacovis_BF20_SNPlist.txt
sort outliers_${PHENO}_anacovis_BF20_SNPlist.txt outliers_${PHENO}_anacovaux_BF20_SNPlist.txt | uniq > outliers_${PHENO}_combine_BF20_SNPlist.txt
awk 'BEGIN{FS=OFS="\t"}{if(NR>1){ print $1,$2-1,$2 }}' outliers_${PHENO}_combine_BF20_SNPlist.txt | bedtools intersect -wb -a $GFF_GO -b stdin | awk '{print $1}' > outliers_${PHENO}_combine_BF20_SNPlist.txt
cd ../
done
```

## Gene ontology

<http://geneontology.org/docs/go-enrichment-analysis/>

Add the gene list from the main VCF file, compare to the SNPs from the outlier lists to see if there is any statistically significant gene enrichment.

```
cd /srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212_7pheno/GOenrich

module load bedtools/2.27.1

VCF=/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/GenoPhenoVariation/sv6_Phenofilter_14pop_Tarsus_maf005_miss50_r2_20210525.vcf.gz
grep -v "^##" $VCF | cut -f1-2 > snplist_chrompos_ALLSNPS.txt

GFF=/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv3_Genome/Sv3.4_GenomeAnnotation/annotation/2020-10-22.vAUMAKER/results_run3_nopred/m
awk 'BEGIN{FS=OFS="\t"}{if(NR>1){ print $1,$2-1,$2 }}' snplist_chrompos_ALLSNPS.txt | bedtools intersect -wb -a $GFF -b stdin | awk '{print $3}' > snplist_chrompos_ALLSNPS.gff
sed -nr 's/.*Similar to +([^\s]+) .*/p' snplist_chrompos_ALLSNPS.gff | sed 's|[:,:]||g' > genelist_snplist_chrompos_ALLSNPS.txt

for PHENO in mass tarsus head beak wing spleen heart;
do
sed -nr 's/.*Similar to +([^\s]+) .*/p' ../Pheno_${PHENO}/outliers_${PHENO}_combine_genelist.txt | sed 's|[:,:]||g' > genelist_${PHENO}.txt
done
```

no significant results :(

[BiNGO \(ugent.be\)](http://ugent.be)

David and Gorilla

[topGO.pdf \(bioconductor.org\)](http://topGO.pdf(bioconductor.org))

## Environment regressions:

[Quick-R: Multiple Regression \(statmethods.net\)](http://statmethods.net)

```
cd /srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/bypass_phenofilter212_7pheno/env_regressions
```

```
module load vcftools/0.1.16
```

```
module load plink/1.90b6.7
```

```
VCF=/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/GenoPhenoVariation/sv6_Phenofilter_14pop_Tarsus_maf005_miss50_r2_20210525.vcf.gz
```

```
mkdir plink_files
for PHENO in mass tarsus head beak wing spleen heart;
do
cut -f3 ../Pheno_${PHENO}/outliers_${PHENO}_combine_BF20_SNPlist.txt > ${PHENO}_snps.txt
vcftools --vcf ${VCF} --snps ${PHENO}_snps.txt --recode --out ${PHENO}_snps
vcftools --vcf ${PHENO}_snps.recode.vcf --plink --out plink_files/${PHENO}_snps.plink
plink --file plink_files/${PHENO}_snps.plink --pca --out plink_files/${PHENO} --make-rel
awk '{print $NR}' plink_files/${PHENO}.rel > plink_files/${PHENO}.rel.diag
done
```

```
head -n 5 *val
```

```
==> beak.eigenval <==
```

```
35.3494
```

```
17.5209
```

```
16.4536
```

```
15.9777
```

```
15.0378
```

```
==> head.eigenval <==
```

```
34.681
```

```
19.7652
```

```
10.6485
```

```
5.75051
```

```
5.2798
```

```
==> heart.eigenval <==
```

```
22.3332
```

```
9.90272
```

```
7.97538
```

```
7.14724
```

```
6.79292
```

```
==> mass.eigenval <==
```

```
36.865
```

```
19.7109
```

```
17.8826
```

```
17.4028
```

```
15.6454
```

```
==> spleen.eigenval <==
```



26.4095

15.1798

6.73227

6.37949

6.21996

==&gt; tarsus.eigenval &lt;==

30.5707

14.4554

14.0551

13.318

12.6836

==&gt; wing.eigenval &lt;==

30.1239

11.6065

11.2747

10.3612

10.0981

```
module load bioconductor/3.10
```

```
R
```

```
library(dplyr)
```

```
library(psych)
```

```
#library(tibble)
```

```
#library(tidyverse)
```

```
library(data.table)
```

```
library(relaimpo)
```

```
setwd("/srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/env_regressions")
```

```
Starling <- read.csv("Sample_Metadata_Kat_212.csv",stringsAsFactors=TRUE,sep=",")
```

```
str(Starling)
```

```
Starling_pheno_check <- filter(Starling, PhenoDataTarsus_212 == "YES") %>% select(c(location, INDV, bio01, bio02, bio03, bio04, bio05, bio06, bio07, bi
```

```
Starling_pheno_check2 <- filter(Starling, PhenoDataTarsus_212 == "YES") %>% select(c(bio01, bio03, bio04, bio07, bio08, bio09))
```

```
Starling_pheno_check3 <- filter(Starling, PhenoDataTarsus_212 == "YES") %>% select(c(bio12, bio13, bio14, bio15, bio18, bio19, elev))
```

```
Starling_pheno_check4 <- filter(Starling, PhenoDataTarsus_212 == "YES") %>% select(c(bio01, bio03, bio04, bio08, bio09, bio12, bio14, bio15, bio18, bi
```

```
pdf("Sv6_env_pairspanel.pdf")
```

```
pairs.panels(Starling_pheno_check, scale=T, cex.labels=1, pch=20, cex=2.5)
```

```
pairs.panels(Starling_pheno_check2, scale=T, cex.labels=1, pch=20, cex=2.5)
```

```
pairs.panels(Starling_pheno_check3, scale=T, cex.labels=1, pch=20, cex=2.5)
```

```
pairs.panels(Starling_pheno_check4, scale=T, cex.labels=1, pch=20, cex=2.5)
```

```
dev.off()
```

```
Starling_pheno_212_env <- filter(Starling, PhenoDataTarsus_212 == "YES") %>% select(c(location, INDV, bio01, bio03, bio04, bio08, bio09, bio12, bio14,
str(Starling_pheno_212_env)
```

### #MASS

```
pca.mass.import <- read.table("plink_files/mass.eigenvec", sep=" ", header=F)
lmer.all.mass <- NA
lmer.all.mass.variance <- NA

for (i in 1:20) {
  j <- i + 2
  pca.mass <- pca.mass.import[,c(2,j)]
  names(pca.mass)[1] <- "INDV"
  names(pca.mass)[2] <- "PCA"
  modeldata.mass <- merge(Starling_pheno_212_env, pca.mass, by.x = "INDV")
  fit.mass <- lm( PCA ~ bio01 + bio03 + bio04 + bio08 + bio09 + bio12 + bio14 + bio15 + bio18 + bio19 + elev + mnNDVI,data=modeldata.mass)
  imp.mass <- calc.relimp(fit.mass,type=c("lmg"), rela=TRUE)
  lmer.all.mass <- data.frame(lmer.all.mass,imp.mass@lmg)
  z <- i + 1
  names(lmer.all.mass)[z] <- paste0("PCA", i)
  lmer.all.mass.variance <- data.frame(lmer.all.mass.variance,imp.mass@R2)
  z <- i + 1
  names(lmer.all.mass.variance)[z] <- paste0("PCA", i)
}

pca.mass.vals <- read.table("plink_files/mass.eigenval", sep=" ", header=F)
pca.mass.diag <- read.table("plink_files/mass.rel.diag", sep=" ", header=F)

for (i in 1:20) {
  lmer.all.mass <- lmer.all.mass %>% mutate(!as.name(paste0("PCA",i,".cor"))) := !as.name(paste0("PCA",i)) * (pca.mass.vals[i,]/sum(as.numeric(pca.mass.vals[i,])))
  lmer.all.mass.variance <- lmer.all.mass.variance %>% mutate(!as.name(paste0("PCA",i,".var"))) := !as.name(paste0("PCA",i)) * (pca.mass.vals[i,]/sum(as.numeric(pca.mass.vals[i,])))
}

#1 axis for mass (2% cutoff)
lmer.all.mass <- lmer.all.mass %>% mutate( PCAallmass.cor = rowSums(.[,22:22]))
lmer.all.mass.variance <- lmer.all.mass.variance %>% mutate( PCAallmass.var = rowSums(.[,22:22]))
```

### #TARSUS

```
pca.tarsus.import <- read.table("plink_files/tarsus.eigenvec", sep=" ", header=F)
lmer.all.tarsus <- NA
lmer.all.tarsus.variance <- NA

for (i in 1:20) {
  j <- i + 2
  pca.tarsus <- pca.tarsus.import[,c(2,j)]
  names(pca.tarsus)[1] <- "INDV"
  names(pca.tarsus)[2] <- "PCA"
  modeldata.tarsus <- merge(Starling_pheno_212_env, pca.tarsus, by.x = "INDV")
  fit.tarsus <- lm( PCA ~ bio01 + bio03 + bio04 + bio08 + bio09 + bio12 + bio14 + bio15 + bio18 + bio19 + elev + mnNDVI,data=modeldata.tarsus)
  imp.tarsus <- calc.relimp(fit.tarsus,type=c("lmg"), rela=TRUE)
  lmer.all.tarsus <- data.frame(lmer.all.tarsus,imp.tarsus@lmg)
  z <- i + 1
  names(lmer.all.tarsus)[z] <- paste0("PCA", i)
  lmer.all.tarsus.variance <- data.frame(lmer.all.tarsus.variance,imp.tarsus@R2)
  z <- i + 1
  names(lmer.all.tarsus.variance)[z] <- paste0("PCA", i)
}

pca.tarsus.vals <- read.table("plink_files/tarsus.eigenval", sep=" ", header=F)
pca.tarsus.diag <- read.table("plink_files/tarsus.rel.diag", sep=" ", header=F)
```

```
for (i in 1:20) {
  lmer.all.tarsus <- lmer.all.tarsus %>% mutate(!as.name(paste0("PCA",i,".cor"))) := !as.name(paste0("PCA",i)) * (pca.tarsus.vals[i,]/sum(as.numeric(pca.tarsus.vals[i,])))
  lmer.all.tarsus.variance <- lmer.all.tarsus.variance %>% mutate(!as.name(paste0("PCA",i,".var"))) := !as.name(paste0("PCA",i)) * (pca.tarsus.vals[i,]/sum(as.numeric(pca.tarsus.vals[i,])))
}
```

```
#1 axis for tarsus (2% cutoff)
```

```
lmer.all.tarsus <- lmer.all.tarsus %>% mutate( PCAalltarsus.cor = rowSums(.[,22:22]))
lmer.all.tarsus.variance <- lmer.all.tarsus.variance %>% mutate( PCAalltarsus.var = rowSums(.[,22:22]))
```

## #HEAD

```
pca.head.import <- read.table("plink_files/head.eigenvec", sep=" ", header=F)
lmer.all.head <- NA
lmer.all.head.variance <- NA

for (i in 1:20) {
  j <- i + 2
  pca.head <- pca.head.import[,c(2,j)]
  names(pca.head)[1] <- "INDV"
  names(pca.head)[2] <- "PCA"
  modeldata.head <- merge(Starling_pheno_212_env, pca.head, by.x = "INDV")
  fit.head <- lm( PCA ~ bio01 + bio03 + bio04 + bio08 + bio09 + bio12 + bio14 + bio15 + bio18 + bio19 + elev + mnNDVI, data=modeldata.head)
  imp.head <- calc.relimp(fit.head, type=c("lmg"), rela=TRUE)
  lmer.all.head <- data.frame(lmer.all.head, imp.head@lmg)
  z <- i + 1
  names(lmer.all.head)[z] <- paste0("PCA", i)
  lmer.all.head.variance <- data.frame(lmer.all.head.variance, imp.head@R2)
  z <- i + 1
  names(lmer.all.head.variance)[z] <- paste0("PCA", i)
}
```

```
pca.head.vals <- read.table("plink_files/head.eigenval", sep=" ", header=F)
pca.head.diag <- read.table("plink_files/head.rel.diag", sep=" ", header=F)
```

```
for (i in 1:20) {
  lmer.all.head <- lmer.all.head %>% mutate(!as.name(paste0("PCA",i,".cor"))) := !as.name(paste0("PCA",i)) * (pca.head.vals[i,]/sum(as.numeric(pca.head.vals[i,])))
  lmer.all.head.variance <- lmer.all.head.variance %>% mutate(!as.name(paste0("PCA",i,".var"))) := !as.name(paste0("PCA",i)) * (pca.head.vals[i,]/sum(as.numeric(pca.head.vals[i,])))
}
```

```
#2 axis for head (2% cutoff)
```

```
lmer.all.head <- lmer.all.head %>% mutate( PCAallhead.cor = rowSums(.[,22:25]))
lmer.all.head.variance <- lmer.all.head.variance %>% mutate( PCAallhead.var = rowSums(.[,22:25]))
```

## #BEAK

```
pca.beak.import <- read.table("plink_files/beak.eigenvec", sep=" ", header=F)
lmer.all.beak <- NA
lmer.all.beak.variance <- NA

for (i in 1:20) {
  j <- i + 2
  pca.beak <- pca.beak.import[,c(2,j)]
  names(pca.beak)[1] <- "INDV"
  names(pca.beak)[2] <- "PCA"
  modeldata.beak <- merge(Starling_pheno_212_env, pca.beak, by.x = "INDV")
  fit.beak <- lm( PCA ~ bio01 + bio03 + bio04 + bio08 + bio09 + bio12 + bio14 + bio15 + bio18 + bio19 + elev + mnNDVI, data=modeldata.beak)
  imp.beak <- calc.relimp(fit.beak, type=c("lmg"), rela=TRUE)
  lmer.all.beak <- data.frame(lmer.all.beak, imp.beak@lmg)
  z <- i + 1
  names(lmer.all.beak)[z] <- paste0("PCA", i)
}
```

```

lmer.all.beak.variance <- data.frame(lmer.all.beak.variance, imp.beak@R2)
  z <- i + 1
names(lmer.all.beak.variance)[z] <- paste0("PCA", i)
}

pca.beak.vals <- read.table("plink_files/beak.eigenval", sep=" ", header=F)
pca.beak.diag <- read.table("plink_files/beak.rel.diag", sep=" ", header=F)

for (i in 1:20) {
lmer.all.beak <- lmer.all.beak %>% mutate(!as.name(paste0("PCA", i, ".cor"))) := !as.name(paste0("PCA", i)) * (pca.beak.vals[i,]/sum(as.numeric(pca.beak.vals[i,])))
lmer.all.beak.variance <- lmer.all.beak.variance %>% mutate(!as.name(paste0("PCA", i, ".var"))) := !as.name(paste0("PCA", i)) * (pca.beak.vals[i,]/sum(as.numeric(pca.beak.vals[i,])))
}

```

#1 axis for beak (2% cutoff)

```

lmer.all.beak <- lmer.all.beak %>% mutate( PCAallbeak.cor = rowSums(.[22:22]))
lmer.all.beak.variance <- lmer.all.beak.variance %>% mutate( PCAallbeak.var = rowSums(.[22:22]))

```

### #WING

```

pca.wing.import <- read.table("plink_files/wing.eigenvec", sep=" ", header=F)
lmer.all.wing <- NA
lmer.all.wing.variance <- NA

for (i in 1:20) {
  j <- i + 2
pca.wing <- pca.wing.import[,c(2,j)]
  names(pca.wing)[1] <- "INDV"
names(pca.wing)[2] <- "PCA"
modeldata.wing <- merge(Starling_pheno_212_env, pca.wing, by.x = "INDV")
fit.wing <- lm( PCA ~ bio01 + bio03 + bio04 + bio08 + bio09 + bio12 + bio14 + bio15 + bio18 + bio19 + elev + mnNDVI, data=modeldata.wing)
  imp.wing <- calc.relimp(fit.wing, type=c("lm"), rela=TRUE)
lmer.all.wing <- data.frame(lmer.all.wing, imp.wing@lm)
  z <- i + 1
names(lmer.all.wing)[z] <- paste0("PCA", i)
lmer.all.wing.variance <- data.frame(lmer.all.wing.variance, imp.wing@R2)
  z <- i + 1
names(lmer.all.wing.variance)[z] <- paste0("PCA", i)
}

```

```

pca.wing.vals <- read.table("plink_files/wing.eigenval", sep=" ", header=F)
pca.wing.diag <- read.table("plink_files/wing.rel.diag", sep=" ", header=F)

```

```

for (i in 1:20) {
lmer.all.wing <- lmer.all.wing %>% mutate(!as.name(paste0("PCA", i, ".cor"))) := !as.name(paste0("PCA", i)) * (pca.wing.vals[i,]/sum(as.numeric(pca.wing.vals[i,])))
lmer.all.wing.variance <- lmer.all.wing.variance %>% mutate(!as.name(paste0("PCA", i, ".var"))) := !as.name(paste0("PCA", i)) * (pca.wing.vals[i,]/sum(as.numeric(pca.wing.vals[i,])))
}

```

#1 axis for wing (2% cutoff)

```

lmer.all.wing <- lmer.all.wing %>% mutate( PCAallwing.cor = rowSums(.[22:22]))
lmer.all.wing.variance <- lmer.all.wing.variance %>% mutate( PCAallwing.var = rowSums(.[22:22]))

```

### #SPLEEN

```

pca.spleen.import <- read.table("plink_files/spleen.eigenvec", sep=" ", header=F)
lmer.all.spleen <- NA
lmer.all.spleen.variance <- NA

for (i in 1:20) {
  j <- i + 2
pca.spleen <- pca.spleen.import[,c(2,j)]
  names(pca.spleen)[1] <- "INDV"
names(pca.spleen)[2] <- "PCA"
modeldata.spleen <- merge(Starling_pheno_212_env, pca.spleen, by.x = "INDV")
fit.spleen <- lm( PCA ~ bio01 + bio03 + bio04 + bio08 + bio09 + bio12 + bio14 + bio15 + bio18 + bio19 + elev + mnNDVI, data=modeldata.spleen)

```

```

    imp.spleen <- calc.relimp(fit.spleen,type=c("lmg"), rela=TRUE)
lmer.all.spleen <- data.frame(lmer.all.spleen,imp.spleen@lmg)
    z <- i + 1
names(lmer.all.spleen)[z] <- paste0("PCA", i)
lmer.all.spleen.variance <- data.frame(lmer.all.spleen.variance,imp.spleen@R2)
    z <- i + 1
names(lmer.all.spleen.variance)[z] <- paste0("PCA", i)
}

pca.spleen.vals <- read.table("plink_files/spleen.eigenval", sep=" ", header=F)
pca.spleen.diag <- read.table("plink_files/spleen.rel.diag", sep=" ", header=F)

for (i in 1:20) {
lmer.all.spleen <- lmer.all.spleen %>% mutate(!as.name(paste0("PCA",i,".cor"))) := !!as.name(paste0("PCA",i)) * (pca.spleen.vals[i,]/sum(as.numeric(pca.spleen.vals[i,])))
lmer.all.spleen.variance <- lmer.all.spleen.variance %>% mutate(!as.name(paste0("PCA",i,".var"))) := !!as.name(paste0("PCA",i)) * (pca.spleen.vals[i,]/sum(as.numeric(pca.spleen.vals[i,])))
}

#2 axis for spleen (2% cutoff)
lmer.all.spleen <- lmer.all.spleen %>% mutate( PCAallspleen.cor = rowSums(.[22:23]))
lmer.all.spleen.variance <- lmer.all.spleen.variance %>% mutate( PCAallspleen.var = rowSums(.[22:23]))

#HEART
pca.heart.import <- read.table("plink_files/heart.eigenvec", sep=" ", header=F)
lmer.all.heart <- NA
lmer.all.heart.variance <- NA

for (i in 1:20) {
    j <- i + 2
pca.heart <- pca.heart.import[,c(2,j)]
    names(pca.heart)[1] <- "INDV"
names(pca.heart)[2] <- "PCA"
modeldata.heart <- merge(Starling_pheno_212_env, pca.heart, by.x = "INDV")
fit.heart <- lm( PCA ~bio01 + bio03 + bio04 + bio08 + bio09 + bio12 + bio14 + bio15 + bio18 + bio19 + elev + mnNDVI,data=modeldata.heart)
    imp.heart <- calc.relimp(fit.heart,type=c("lmg"), rela=TRUE)
lmer.all.heart <- data.frame(lmer.all.heart,imp.heart@lmg)
    z <- i + 1
names(lmer.all.heart)[z] <- paste0("PCA", i)
lmer.all.heart.variance <- data.frame(lmer.all.heart.variance,imp.heart@R2)
    z <- i + 1
names(lmer.all.heart.variance)[z] <- paste0("PCA", i)
}

pca.heart.vals <- read.table("plink_files/heart.eigenval", sep=" ", header=F)
pca.heart.diag <- read.table("plink_files/heart.rel.diag", sep=" ", header=F)

for (i in 1:20) {
lmer.all.heart <- lmer.all.heart %>% mutate(!as.name(paste0("PCA",i,".cor"))) := !!as.name(paste0("PCA",i)) * (pca.heart.vals[i,]/sum(as.numeric(pca.heart.vals[i,])))
lmer.all.heart.variance <- lmer.all.heart.variance %>% mutate(!as.name(paste0("PCA",i,".var"))) := !!as.name(paste0("PCA",i)) * (pca.heart.vals[i,]/sum(as.numeric(pca.heart.vals[i,])))
}

#3 axis for heart (2% cutoff)
lmer.all.heart <- lmer.all.heart %>% mutate( PCAallheart.cor = rowSums(.[22:25]))
lmer.all.heart.variance <- lmer.all.heart.variance %>% mutate( PCAallheart.var = rowSums(.[22:25]))

#summary of how much genetic variance is used over all analysed axis

sum(as.numeric(pca.mass.vals[1:1,1]))/sum(as.numeric(pca.mass.diag[,1]))
sum(as.numeric(pca.tarsus.vals[1:1,1]))/sum(as.numeric(pca.tarsus.diag[,1]))
sum(as.numeric(pca.head.vals[1:4,1]))/sum(as.numeric(pca.head.diag[,1]))
sum(as.numeric(pca.beak.vals[1:1,1]))/sum(as.numeric(pca.beak.diag[,1]))

```

```
sum(as.numeric(pca.wing.vals[1:1,1]))/sum(as.numeric(pca.wing.diag[,1]))
sum(as.numeric(pca.spleen.vals[1:2,1]))/sum(as.numeric(pca.spleen.diag[,1]))
sum(as.numeric(pca.heart.vals[1:4,1]))/sum(as.numeric(pca.heart.diag[,1]))
```

```
[1] 0.1054824
```

```
> sum(as.numeric(pca.tarsus.vals[1:1,1]))/sum(as.numeric(pca.tarsus.diag[,1]))
```

```
[1] 0.08683995
```

```
> sum(as.numeric(pca.head.vals[1:4,1]))/sum(as.numeric(pca.head.diag[,1]))
```

```
[1] 0.206143
```

```
> sum(as.numeric(pca.beak.vals[1:1,1]))/sum(as.numeric(pca.beak.diag[,1]))
```

```
[1] 0.1039676
```

```
> sum(as.numeric(pca.wing.vals[1:1,1]))/sum(as.numeric(pca.wing.diag[,1]))
```

```
[1] 0.08859117
```

```
> sum(as.numeric(pca.spleen.vals[1:2,1]))/sum(as.numeric(pca.spleen.diag[,1]))
```

```
[1] 0.1229814
```

```
> sum(as.numeric(pca.heart.vals[1:4,1]))/sum(as.numeric(pca.heart.diag[,1]))
```

```
[1] 0.1387845
```

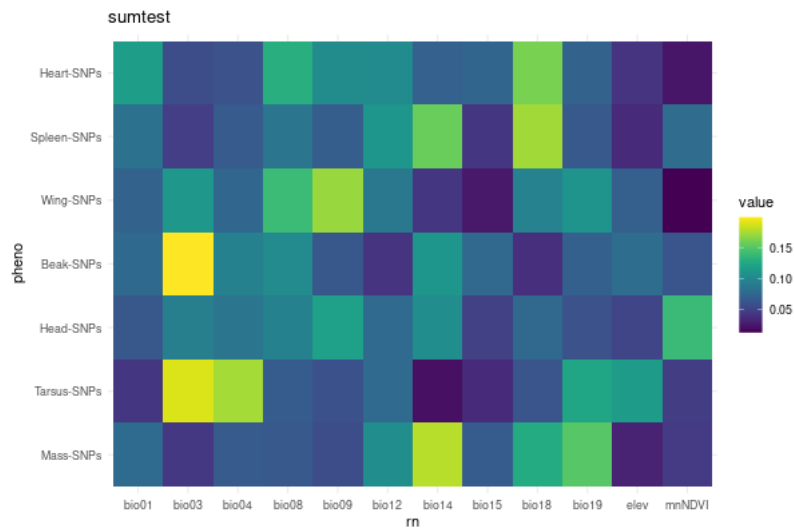
## #SUMMARISE AND PLOT

```
lmer.all <- data.frame(lmer.all.mass$PCAallmass.cor,lmer.all.tarsus$PCAalltarsus.cor,lmer.all.head$PCAallhead.cor,lmer.all.beak$PCAallbeak.cor,lmer.all.wing$PCAallwing.cor,lmer.all.spleen$PCAallspleen.cor,lmer.all.heart$PCAallheart.cor)
library(data.table)
lmer.all2 <- setDT(lmer.all, keep.rownames = TRUE)[,]
lmer.all2[,1] <- c("bio01","bio03","bio04","bio08","bio09","bio12","bio14","bio15","bio18","bio19","elev","mnNDVI")
lmer.all2.long <- melt(setDT(lmer.all2), id.vars = c("rn"), variable.name = "pheno")

library(ggplot2)
library(viridis)

xlab.phenos <- c("Mass-SNPs", "Tarsus-SNPs", "Head-SNPs", "Beak-SNPs", "Wing-SNPs", "Spleen-SNPs", "Heart-SNPs")

png("Sv6_geno_env_heatpanels_variableaxis.png", width = 800, height = 500)
ggplot(lmer.all2.long, aes(rn, pheno, fill= value)) +
  geom_tile() +
  scale_fill_viridis(discrete=FALSE) +
  scale_y_discrete(labels= xlab.phenos) +
  theme_minimal(base_size = 18) + theme(axis.text.x = element_text(angle =45, hjust=1),axis.title.x = element_blank(),axis.title.y = element_blank())
dev.off()
```



```
#lmer.all.var <-
data.frame(lmer.all.mass.variance$PCAallmass.var,lmer.all.tarsus.variance$PCAalltarsus.var,lmer.all.head.variance$PCAallhead.var,lmer.all.beak.variance

#lmer.all.var2 <- as.data.frame(t(lmer.all.var))

#values obtained by multiple total gen var used in model (important axis) by overall env exp across the important axis
lmer.all.var2 <- c(0.0088153, 0.008126831, 0.01973585, 0.008118384, 0.009433739, 0.014707997, 0.013428924)

rownames2 <- c("1-Mass", "2-Tarsus", "3-Head", "4-Beak", "5-Wing", "6-Spleen", "7-Heart")

lmer.all.var3 <- as.data.frame(cbind(rownames2, lmer.all.var2))

lmer.all.var3$lmer.all.var2 <- as.numeric(as.character(lmer.all.var3$lmer.all.var2))

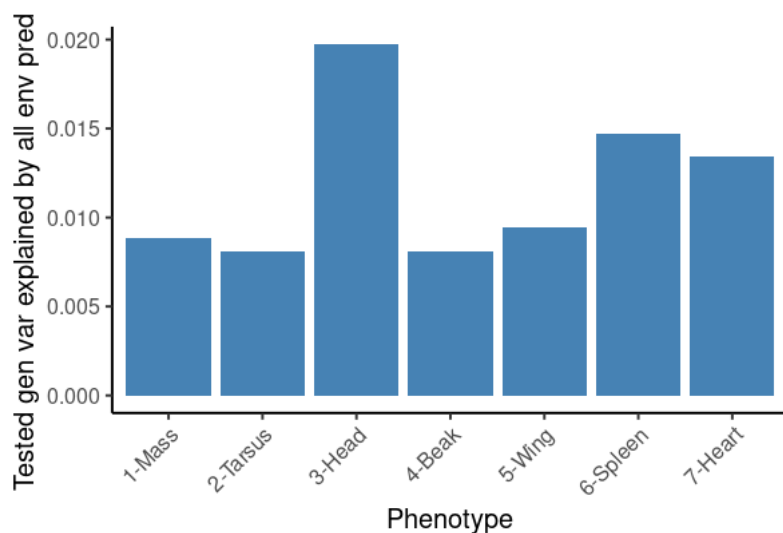
png("Sv6_geno_env_heatpanels_var2.png", width = 600, height = 400)

ggplot(data=lmer.all.var3, aes(x=rownames2, y=lmer.all.var2)) +

  geom_bar(stat="identity", fill="steelblue")+

  theme_classic(base_size = 20) + xlab("Phenotype") + ylab("Tested gen var explained by all env predictors") +
  theme(axis.text.x = element_text(angle =45, hjust=1))

dev.off()
```



#Total genetic variance this captures:

sum the vec file over the diag to get overall var.

```
rownames <- c("Axis1", "Axis2", "Axis3", "Axis4", "Axis5", "Axis6", "Axis7")
```

```
lmer.all.var.trend <- as.data.frame(cbind(as.numeric(lmer.all.mass.variance[,22:30]),as.numeric(lmer.all.tarsus.variance[,22:30]),as.numeric(lmer.all.head.variance[,22:30]),as.numeric(lmer.all.spleen.variance[,22:30]),as.numeric(lmer.all.heart.variance[,22:30])))
```

```
colnames(lmer.all.var.trend) <- c("Mass", "Tarsus", "Head", "Beak", "Wing", "Spleen", "Heart")
```

```
lmer.all.var.trend$Axis <- c("Axis1", "Axis2", "Axis3", "Axis4", "Axis5", "Axis6", "Axis7", "Axis8", "Axis9")
```

```
WGT1 <- ggplot(data=lmer.all.var.trend, aes(x=Axis, y=Mass, group=1)) + geom_line() + theme_classic(base_size = 20) + xlab("Genetic Axis") + ylab("Weight")
theme(axis.text.x = element_text(angle = 45, hjust=1))
```

```
WGT2 <- ggplot(data=lmer.all.var.trend, aes(x=Axis, y=Tarsus, group=1)) + geom_line() + theme_classic(base_size = 20) + xlab("Genetic Axis") + ylab("Weight")
theme(axis.text.x = element_text(angle = 45, hjust=1))
```

```
WGT3 <- ggplot(data=lmer.all.var.trend, aes(x=Axis, y=Head, group=1)) + geom_line() + theme_classic(base_size = 20) + xlab("Genetic Axis") + ylab("Weight")
theme(axis.text.x = element_text(angle = 45, hjust=1))
```

```
WGT4 <- ggplot(data=lmer.all.var.trend, aes(x=Axis, y=Beak, group=1)) + geom_line() + theme_classic(base_size = 20) + xlab("Genetic Axis") + ylab("Weight")
theme(axis.text.x = element_text(angle = 45, hjust=1))
```

```
WGT5 <- ggplot(data=lmer.all.var.trend, aes(x=Axis, y=Wing, group=1)) + geom_line() + theme_classic(base_size = 20) + xlab("Genetic Axis") + ylab("Weight")
theme(axis.text.x = element_text(angle = 45, hjust=1))
```

```
WGT6 <- ggplot(data=lmer.all.var.trend, aes(x=Axis, y=Spleen, group=1)) + geom_line() + theme_classic(base_size = 20) + xlab("Genetic Axis") + ylab("Weight")
theme(axis.text.x = element_text(angle = 45, hjust=1))
```

```
WGT7 <- ggplot(data=lmer.all.var.trend, aes(x=Axis, y=Heart, group=1)) + geom_line() + theme_classic(base_size = 20) + xlab("Genetic Axis") + ylab("Weight")
theme(axis.text.x = element_text(angle = 45, hjust=1))
```

```
png("Sv6_geno_env_heatpanels_axisvars.png", width=800, height=1000)
```

```
grid.arrange(WGT1, WGT2, WGT3, WGT4, WGT5, WGT6, WGT7, ncol=3)
```

```
dev.off()
```

```
library(gridExtra)
```

```
library(grid)
```

```
library(lattice)
```

```
png("Sv6_EnvReg_3panel.png", width=800, height=500)
```

```
pushViewport(viewport(layout = grid.layout(1, 3)))
```

```
vplayout <- function(x, y) viewport(layout.pos.row = x, layout.pos.col = y)
```

```
print(plotA, vp = vplayout(1, 1:2))
```

```
print(plotB, vp = vplayout(1, 3))
```

```
dev.off()
```



## FASTSTRUCTURE

```
module load plink/1.90b6.7
cd /srv/scratch/z5188231/KStuart.Starling-Aug18/Sv6_Morphology/analysis/baypass_phenofilter212_7pheno/env_regressions/fastSTRUCTURE
plink --file ../plink_files/mass_snps.plink --make-bed --noweb --out mass_snps
```

```
mkdir mypythonenv
module load python/2.7.15
module add gsl/2.5
virtualenv mypythonenv
source mypythonenv/bin/activate
pip install Cython==0.27.3
pip install NumPy==1.16.2
pip install SciPy==1.2.1
pip install matplotlib
```

```
mkdir proj
cd proj
git clone https://github.com/rajanil/fastStructure
cd fastStructure/vars
python setup.py build_ext --inplace
cd ../
python setup.py build_ext --inplace
```

```
for SEED in {1};
do
for K in 2;
do
echo working with k=$K seed=${SEED};
python proj/fastStructure/structure.py -K $K --input=mass_snps --format=bed --prior=simple --output=mass_snps_${SEED}_out --full - - seed=${SEED}
done;
done
```

```

for SEED in {1...2};
do
for K in 1 2;
do
echo working with k=$K seed=${SEED};
admixture -s ${SEED} -B -j16 --cv mass_snps.bed $K | tee log${K}.${SEED}.all.out; done
done;
done

```

```

for SEED in {1...2};
do
for K in 1 2;
do
echo working with k=$K seed=${SEED};
admixture -s ${SEED} --cv mass_snps.bed $K | tee log${K}.${SEED}.all.out; done
done;
done

```

#####Uisng fastSTRUCTURE stuff

#MASS

```

pca.mass.import <- read.table("plink_files/mass.eigenvec", sep=" ", header=F)
pca1.mass.import <- read.table("fastSTRUCTURE/mass_snps.2.Q", sep=" ", header=F)
pca1.mass <- data.frame(pca.mass.import[,1], pca1.mass.import[,1])
names(pca1.mass)[1] <- "INDV"
names(pca1.mass)[2] <- "FS1"
modeldata.mass <- merge(Starling_pheno_212_env, pca1.mass, by.x = "INDV")
fit.mass <- lm(FS1~bio01 + bio03 + bio04 + bio08 + bio09 + bio14 + bio15 + bio18 + bio19 + elev,data=modeldata.mass)
imp.mass <- calc.relimp(fit.mass,type=c("lmg"), rela=TRUE)
##Proportion of variance explained by model: 52.85%
#####

```

FS:

lmg

bio01 0.05919943

bio03 0.05952670

bio04 0.07222592

bio08 0.05576394

bio09 0.05571183

bio14 0.22490678

bio15 0.08249799

bio18 0.18000418

bio19 0.18392955

elev 0.02623366

PCA1:

0.06396111  
0.04916456  
0.08208704  
0.05823958  
0.05515386  
0.21035363  
0.07387125  
0.17975560  
0.19786961  
0.02954376