CeNDR lifespans

2019_10_31

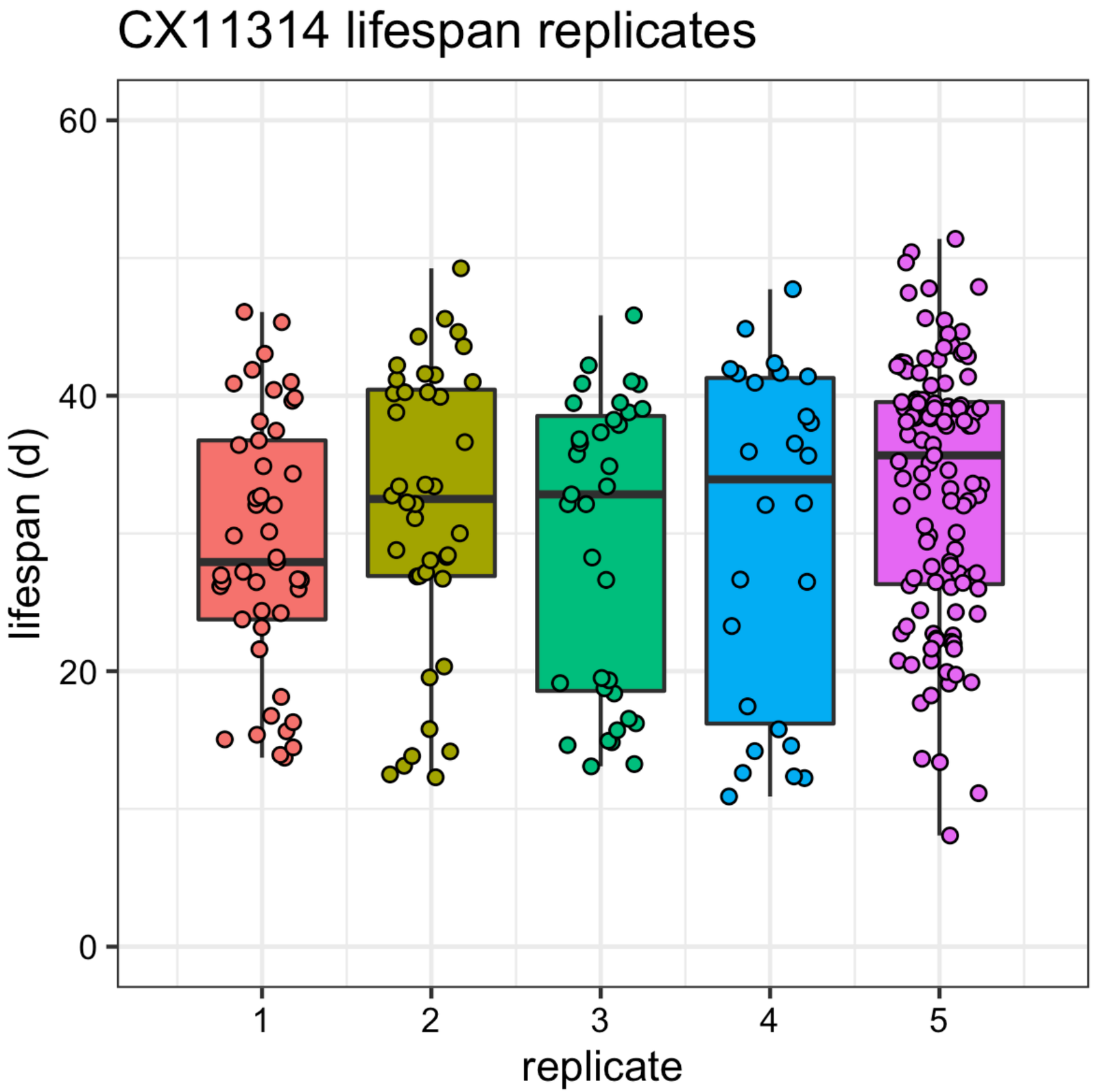
Background

- Original lifespan data are from Patrick Phillips' life machines. These include lifespan measures for 20 strains replicated 4 times each with ~60 individuals per replicate.

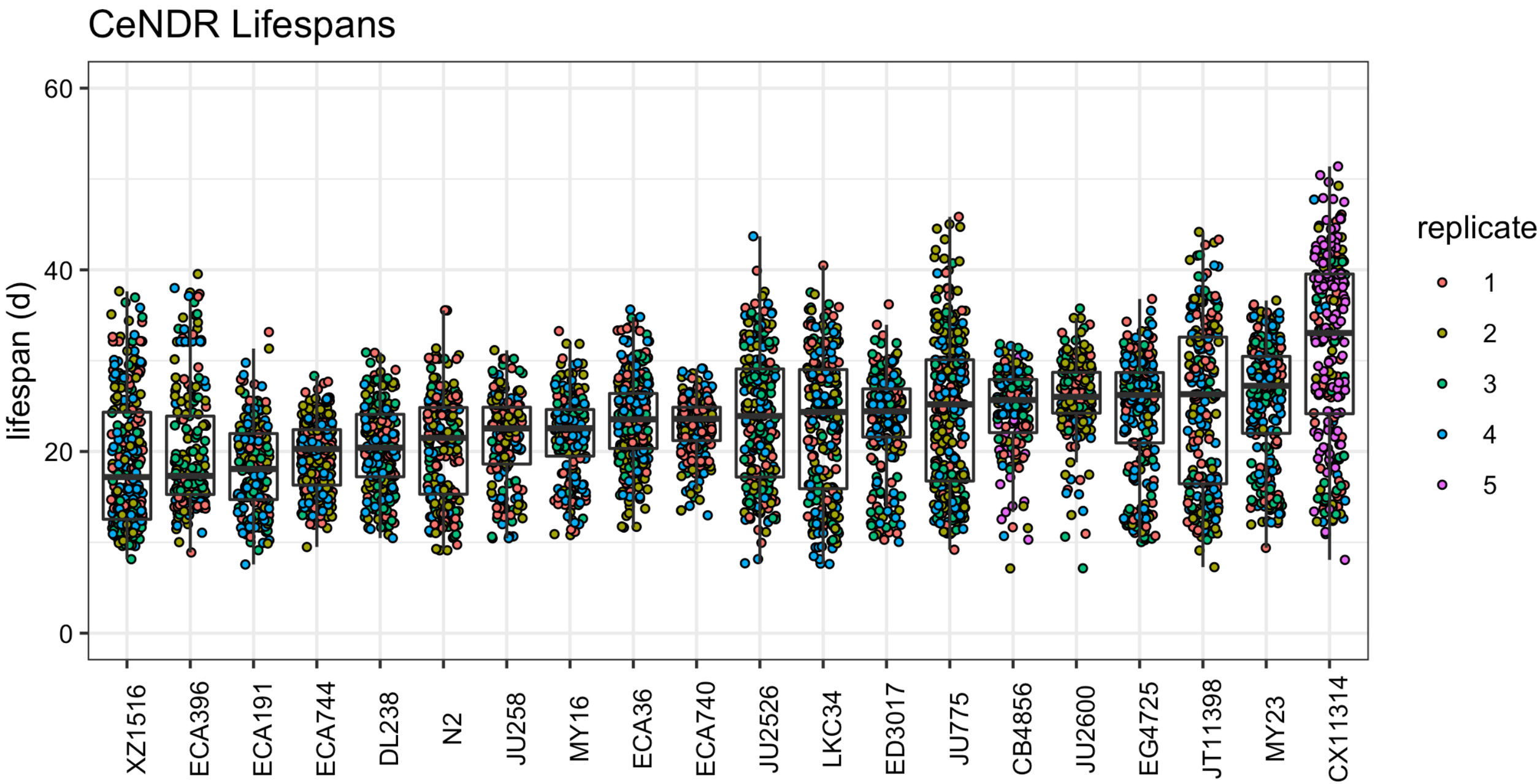
- CX11314 was the longest lived strain among the 20 but this is inconsistent with observations in the Feng-Yen lab.

- The Phillips Lab performed a fifth replicate assay to measure CX11314 lifespan and confirm its long lived phenotype.

All replicate measures of CX11314 have long-lived individuals. It looks like there is bimodality in lifespans for reps 1 - 4 but much less in rep 5.

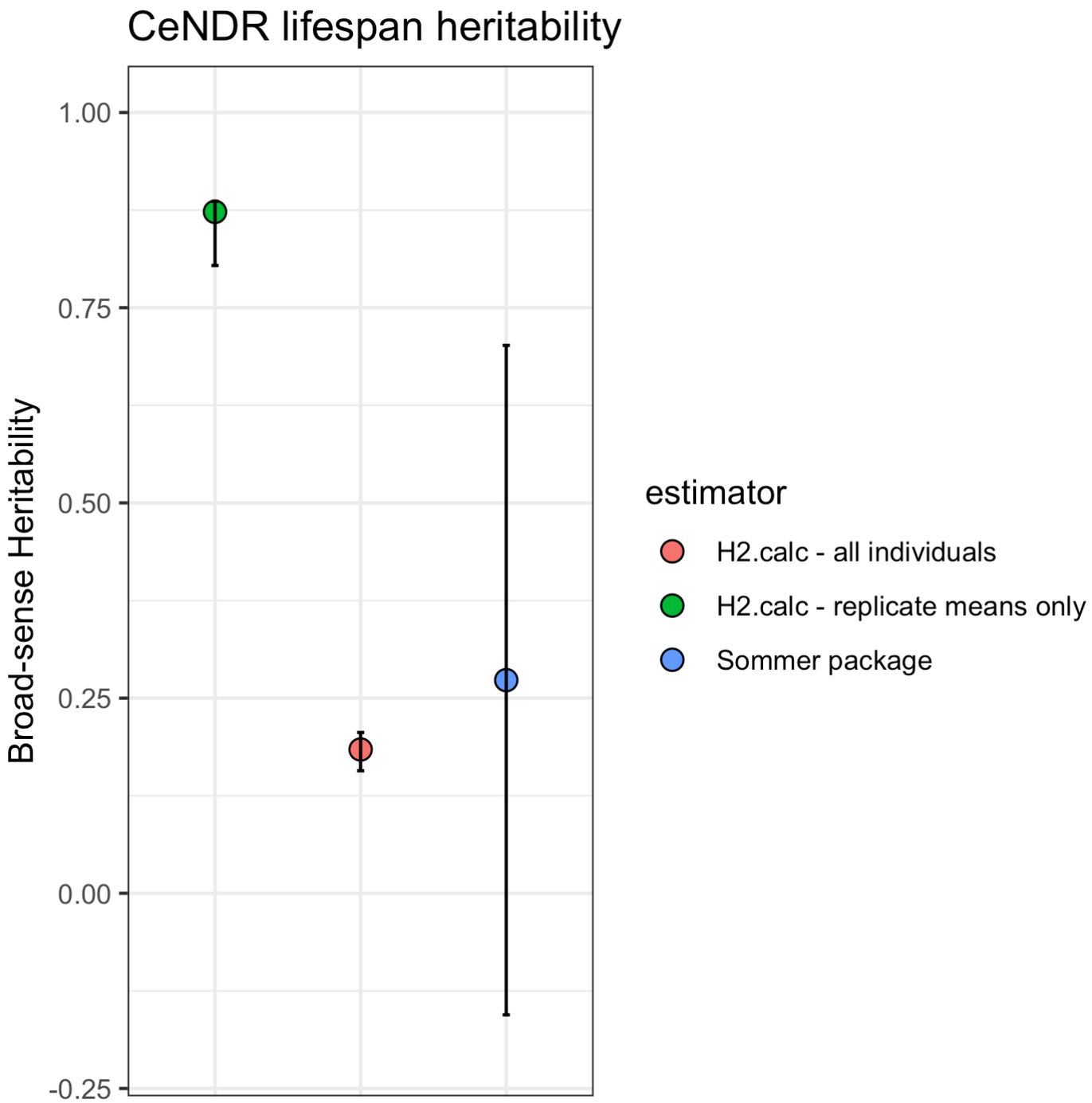


CX11314 is still the longest lived CeNDR strain as measured with the life machines. This is true if we include the 5th replicate or not.



Broad-sense heritability estimates for the lifespan phenotype. I've measured broad sense heritability three ways.

1. **H2.calc - all individuals**: uses lifespans of all individuals in the heritability calculation.
2. **H2.calc - replicate means only:** uses the mean lifespans from the replicates in the H2 calculation.
3. **Sommer package:** uses genotype matrix and mean phenotypes for each strain (no replicate measures into model) to estimate genetic variance (additive and epistatic) and environmental variance (error variance).



Heritability with H2.calc function

```r
# data is data frame that contains strain and phenotype column
# indicies are used by the boot function to sample from the 'data' data.frame
H2.test.boot <- function(data, indicies){

  d <- data[indicies,]

  pheno <- as.data.frame(dplyr::select(d, phenotype))[,1]
  strain <- as.factor(d$strain)

  reffMod <- lme4::lmer(pheno ~ 1 + (1|strain))

  Variances <- as.data.frame(lme4::VarCorr(reffMod, comp = "Variance"))

  Vg <- Variances$vcov[1]
  Ve <- Variances$vcov[2]
  H2 <- Vg/(Vg+Ve)

  # errors <- sqrt(diag(lme4::VarCorr(reffMod, comp = "Variance")$strain))

  return(H2)
}

# df is data frame that contains strain and phenotype column. H2.test function not shown
H2.calc <- function(df, boot = T){


  if(boot == T){
    # bootstrapping with R replications
    results <- boot(data=df, statistic=H2.test.boot, R=10000)

    # get 95% confidence interval
    ci <- boot.ci(results, type="bca")

    H2_errors <- data.frame(H2 = ci$t0, ci_l = ci$bca[4], ci_r = ci$bca[5])

    return(H2_errors)

  } else {

    H2 <- data.frame(H2 = H2.test(data = df), ci_l = NA, ci_r = NA)
    return(H2)
  }

}
```

Sommer package

```r
# load in genotype matrix from cegwas2-nf (WI.20180527.impute.vcf.gz used to generate matrix)
geno_matrix <- data.table::fread("data/life_machine_Genotype_Matrix.tsv")

# df with strain, strain, trait value
df_y <- traitfile_small %>%
  dplyr::rename(value = mean_ls) %>%
  dplyr::mutate(strain1=strain, strain2=strain) %>%
  dplyr::select(-strain)

A <- sommer::A.mat(t(dplyr::select(geno_matrix, -CHROM, -POS, -REF, -ALT)))
E <- sommer::E.mat(t(dplyr::select(geno_matrix, -CHROM, -POS, -REF, -ALT)))

df_H2 <- sommer::mmer(value~1, random=~vs(strain1,Gu=A)+vs(strain2,Gu=E), data=df_y)

(summary(df_H2)$varcomp)

# narrow-sense H2 (additive only). Note, we never worry about dominance because we assume homozygous.
lm_sommer_narrow_h2 <- pin(df_H2, H2 ~ (V1) / (V1+V2+V3))

# Broad-sense H2 (additive + epistatic variance) / (additive, epistatic, error)
lm_sommer_broad_h2 <- pin(df_H2, H2 ~ (V1+V2) / (V1+V2+V3))
```