calculate fitness and genetic variance in inbred and hybrid population1

FANG WANG

April 1st, 2020

#IDEA: the interested traits are ‘Days to anthesis’, ‘Plant Yield’ and ‘Total kernel weight’ in inbred population and population 1 (hybrid).Because the %BPH of these traits in hybrid are all pretty high (100%,98% and 99%). All these traits show significant heterosis depending on the literature.

#Hypothesis: Heterosis exits in the traits ‘Days to anthesis’, ‘Plant Yield’ and ‘Total kernel weight’ which has high better-parent heterosis percentage (%BPH) in the hybrid population.

## Normalize the path:

## Flint-Garcia et al., 2009 data

f <- read.delim("data/journal.pone.0007433.s001.txt", header=TRUE)  
# Convert missing data to NA  
f[f=="."] <- NA

###variance of each trait in inbred population

###model:Y=u+Fi+Ej+e (u:overall mean, Fi:individuals genetic background, Ej: effect by jth environment, e: error)

f$INBRED <- as.factor(f$INBRED)  
f$Env <- as.factor(f$Env)  
f$PltYield\_Inbred<-as.factor(f$PltYield\_Inbred)  
f$DTT\_Inbred<-as.factor(f$DTT\_Inbred)  
f$TotKnlWt\_Inbred<-as.factor(f$TotKnlWt\_Inbred)  
  
#calculate ANOVA for each trait in inbred population  
##Total kernel weight  
f$TotKnlWt\_Inbred<-as.numeric(as.character((f$TotKnlWt\_Inbred)))  
fit\_TotKnlwt <- lm(TotKnlWt\_Inbred ~ INBRED + Env, data=f)  
a\_TotKnlwt<-anova(fit\_TotKnlwt)  
a\_TotKnlwt

## Analysis of Variance Table  
##   
## Response: TotKnlWt\_Inbred  
## Df Sum Sq Mean Sq F value Pr(>F)   
## INBRED 303 182960 603.8 1.9637 2.214e-12 \*\*\*  
## Env 3 18749 6249.7 20.3247 1.548e-12 \*\*\*  
## Residuals 581 178653 307.5   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##Days to Anthesis  
f$DTT\_Inbred<-as.numeric(as.character((f$DTT\_Inbred)))  
fit\_DTT <- lm(DTT\_Inbred ~ INBRED + Env, data=f)  
a\_DTT<-anova(fit\_DTT)  
a\_DTT

## Analysis of Variance Table  
##   
## Response: DTT\_Inbred  
## Df Sum Sq Mean Sq F value Pr(>F)   
## INBRED 310 71943 232.1 18.083 < 2.2e-16 \*\*\*  
## Env 3 51597 17198.9 1340.137 < 2.2e-16 \*\*\*  
## Residuals 733 9407 12.8   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##Plant Yield  
f$PltYield\_Inbred<-as.numeric(as.character((f$PltYield\_Inbred)))  
fit\_PltYield<-lm(PltYield\_Inbred ~ INBRED + Env, data=f)  
a\_PltYield<-anova(fit\_PltYield)  
a\_PltYield

## Analysis of Variance Table  
##   
## Response: PltYield\_Inbred  
## Df Sum Sq Mean Sq F value Pr(>F)   
## INBRED 227 124640 549.08 1.5295 0.003149 \*\*  
## Env 1 2203 2202.82 6.1362 0.014426 \*   
## Residuals 141 50617 358.99   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

###H2 of each trait in inbred population

Vgt = (603.8 - 307.5)/4  
Vet = 307.5  
H2\_TotKnlWt\_Inbred = Vgt/(Vgt + Vet)  
H2\_TotKnlWt\_Inbred

## [1] 0.1941296

Vgd = (232.1 - 12.8)/4  
Ved = 12.8  
H2\_DTT\_Inbred = Vgd/(Vgd + Ved)  
H2\_DTT\_Inbred

## [1] 0.8107209

Vgp = (549.08 - 358.99)/2  
Vep = 358.99  
H2\_PltYield\_Inbred = Vgp/(Vgp + Vep)  
H2\_PltYield\_Inbred

## [1] 0.2093341

###variance of each trait in hybrid population1

###For population1, it is a half-sib mating design that all the inbred lines were considered as male parent and making cross with B73.

###model:Y=u+Fi+Ej\_e(u:overall mean, Fi:effect of ith parent, Ej: effect by jth environment, e: error)

f$INBRED <- as.factor(f$INBRED)  
f$Env <- as.factor(f$Env)  
f$TotKnlWt\_Hyb <- as.factor(f$KnlHgt\_Hyb)  
f$DTT\_Hyb<-as.factor(f$DTT\_Hyb)  
f$PltYield\_Hyb<-as.factor(f$PltYield\_Hyb)  
  
#calculate ANOVA for each trait in hybrid population  
##Total kernel weight  
f$TotKnlWt\_Hyb <- as.numeric(as.character((f$TotKnlWt\_Hyb)))  
fit\_TotKnlwt\_Hyb <- lm(TotKnlWt\_Hyb ~ INBRED + Env, data=f)  
a\_TotKnlwt\_Hyb <- anova(fit\_TotKnlwt\_Hyb)  
a\_TotKnlwt\_Hyb

## Analysis of Variance Table  
##   
## Response: TotKnlWt\_Hyb  
## Df Sum Sq Mean Sq F value Pr(>F)   
## INBRED 309 2970.7 9.61 3.994 < 2.2e-16 \*\*\*  
## Env 3 1613.6 537.87 223.452 < 2.2e-16 \*\*\*  
## Residuals 748 1800.5 2.41   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##Day to anthesis  
f$DTT\_Hyb <- as.numeric(as.character((f$DTT\_Hyb)))  
fit\_DTT\_Hyb <- lm(DTT\_Hyb ~ INBRED + Env, data=f)  
a\_DTT\_Hyb<-anova(fit\_DTT\_Hyb)  
a\_DTT\_Hyb

## Analysis of Variance Table  
##   
## Response: DTT\_Hyb  
## Df Sum Sq Mean Sq F value Pr(>F)   
## INBRED 309 22415 72.5 22.846 < 2.2e-16 \*\*\*  
## Env 3 38829 12943.1 4076.388 < 2.2e-16 \*\*\*  
## Residuals 774 2458 3.2   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##Plant yield  
f$PltYield\_Hyb <- as.numeric(as.character((f$PltYield\_Hyb)))  
fit\_PltYield\_Hyb <- lm(PltYield\_Hyb ~ INBRED + Env, data=f)  
a\_PltYield\_Hyb<-anova(fit\_PltYield\_Hyb)  
a\_PltYield\_Hyb

## Analysis of Variance Table  
##   
## Response: PltYield\_Hyb  
## Df Sum Sq Mean Sq F value Pr(>F)   
## INBRED 261 608627 2332 1.9277 2.441e-07 \*\*\*  
## Env 1 161885 161885 133.8275 < 2.2e-16 \*\*\*  
## Residuals 228 275801 1210   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

###the covariance between relatives for each trait: cov=1/4*Vf=1/4*V(inbred)

Vgth = (9.61 - 2.41)/4 #this is the covariance between male parents(selected from inbred population) and B73 (female parent)  
Veth = 2.41  
H2\_TotKnlWt\_Hyb = Vgth/(Vgth + Veth)  
H2\_TotKnlWt\_Hyb

## [1] 0.4275534

Vgdh = (72.5 - 3.2)/4 #this is the covariance between male parents(selected from inbred population) and B73 (female parent)  
Vedh = 3.2  
H2\_DTT\_Hyb = Vgdh/(Vgdh + Vedh)  
H2\_DTT\_Hyb

## [1] 0.8440926

Vgph = (2332 - 1210)/2 #this is the covariance between male parents(selected from inbred population) and B73 (female parent)  
Veph = 1210  
H2\_PltYield\_Hyb = Vgph/(Vgph + Veph)  
H2\_PltYield\_Hyb

## [1] 0.3167702

#H2\_NUMBER<-c(H2\_TotKnlWt\_Inbred,H2\_TotKnlWt\_Hyb,H2\_DTT\_Inbred,H2\_DTT\_Hyb,H2\_PltYield\_Inbred,H2\_PltYield\_Hyb)  
H2\_NUMBER<-c(0.1941296,0.4275534, 0.8107209, 0.8440926, 0.2093341, 0.3167702)  
H2\_NUMBER

## [1] 0.1941296 0.4275534 0.8107209 0.8440926 0.2093341 0.3167702

H2\_name<-c("H2\_TotKnlWt\_Inbred","H2\_TotKnlWt\_Hyb","H2\_DTT\_Inbred","H2\_DTT\_Hyb","H2\_PltYield\_Inbred","H2\_PltYield\_Hyb")  
H2\_name

## [1] "H2\_TotKnlWt\_Inbred" "H2\_TotKnlWt\_Hyb" "H2\_DTT\_Inbred"   
## [4] "H2\_DTT\_Hyb" "H2\_PltYield\_Inbred" "H2\_PltYield\_Hyb"

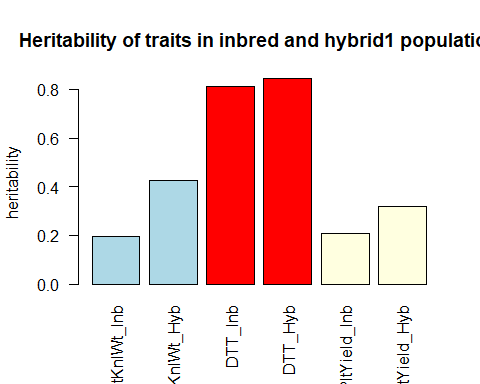
data.matrix(H2\_name,H2\_NUMBER)

## [,1]   
## [1,] "H2\_TotKnlWt\_Inbred"  
## [2,] "H2\_TotKnlWt\_Hyb"   
## [3,] "H2\_DTT\_Inbred"   
## [4,] "H2\_DTT\_Hyb"   
## [5,] "H2\_PltYield\_Inbred"  
## [6,] "H2\_PltYield\_Hyb"

#H2=data.frame(H2\_name,H2\_NUMBER)  
H2<-cbind(H2\_name,H2\_NUMBER)  
H2

## H2\_name H2\_NUMBER   
## [1,] "H2\_TotKnlWt\_Inbred" "0.1941296"  
## [2,] "H2\_TotKnlWt\_Hyb" "0.4275534"  
## [3,] "H2\_DTT\_Inbred" "0.8107209"  
## [4,] "H2\_DTT\_Hyb" "0.8440926"  
## [5,] "H2\_PltYield\_Inbred" "0.2093341"  
## [6,] "H2\_PltYield\_Hyb" "0.3167702"

library(ggplot2)  
barplot(H2\_NUMBER,  
 main = "Heritability of traits in inbred and hybrid1 population",  
 col = c("light blue","light blue","red","red","light yellow","light yellow"),  
 ylab = "heritability",  
 names.arg = c("TotKnlWt\_Inb","TotKnlWt\_Hyb","DTT\_Inb","DTT\_Hyb","PltYield\_Inb","PltYield\_Hyb"),   
 las=2)



##interpretation of the result: ##Overall, the H2 of DTT is significantly higher than other two trait which means the “Days to Anthesis” is a more stable and genetic-controlled trait. The Heritability of hybrid is higher than it in inbred populations for three traits. This indicates the heterosis do exit in the hybrid offsprings which lays the theoretical foundation for breeding.