

# FSTs in hybrid poplars

Vikram Chhatre

07/20/2022

## 1. Create a Fst Density Plot

- Import libraries and data into R

```
library(ggplot2)
library(reshape2)

sig.fst <- read.table("sig.fst", header=F)
neu.fst <- read.table("neu.fst", header=F)
```

- Format and merge data

```
sig <- data.frame(x=sig.fst$V2, label=rep("Significant (1764 SNPs)", 1764))

neu <- data.frame(x=neu.fst$V2, label=rep("Neutral (1912 SNPs)", 1912))

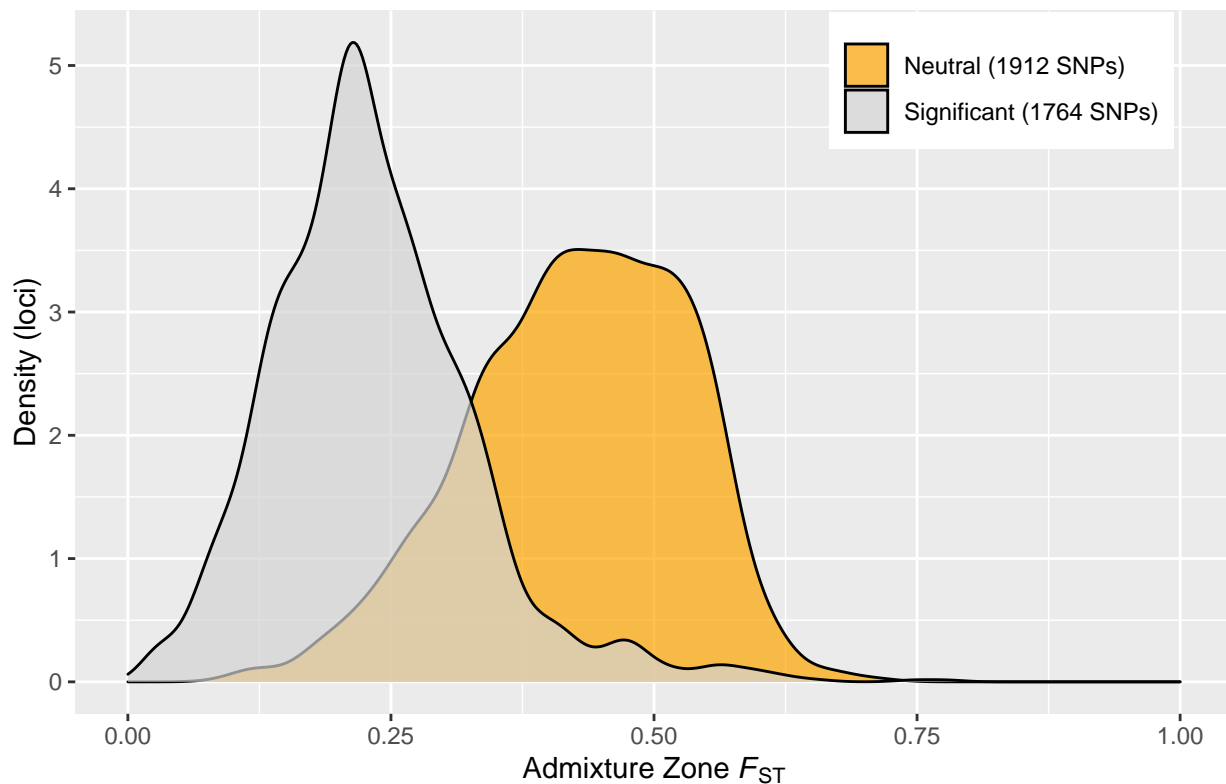
fst <- rbind(sig, neu)
```

- Make density plot

```
ggplot(fst, aes(x, y=..density.., fill=label)) + xlim(0,1) +
  geom_density(color="black", alpha=0.7) +
  scale_fill_manual(values=c("orange", "lightgray")) +
  labs(x=expression(paste("Admixture Zone"~italic(F)[ST])), y="Density (loci)",
        title=expression(paste("(b) "~italic(F)[ST]~" in "~italic("P. angustifolia")~" derived alleles")),
        theme(legend.position=c(0.8,0.9), legend.title=element_blank())
```

```
## Warning: Removed 49 rows containing non-finite values (stat_density).
```

(b)  $F_{ST}$  in *P. angustifolia* derived alleles



## 2. Reproducibility of results

- Read the data file

```
pct <- read.table("pval_counts.txt", header=T)
```

- Assign colors to numbers

```
pct$fill <- ifelse(pct$counts_pvalzero == 0, "lightgray",
  ifelse(pct$counts_pvalzero == 100, "orange", "black"))
```

- Draw a histogram to show reproducibility

```
ggplot(pct, aes(counts_pvalzero, fill=fill)) +
  geom_histogram(binwidth=1) +
  scale_fill_manual(values=c("black", "gray", "orange"),
    labels=c("Variably Sig. (3847 SNPs)", "Never Sig. (1912 SNPs)", "Always Sig. (1764 SNPs)")) +
  labs(x="Num. of Introgress Replicates", y="Significantly introgressed loci",
    title="(a) Reproducibility of Introgress results") +
  theme(legend.position=c(0.8,0.9), legend.title=element_blank(), legend.text=element_text(size=8))
```

