

Figure 1: Original plots with SNP sliding window

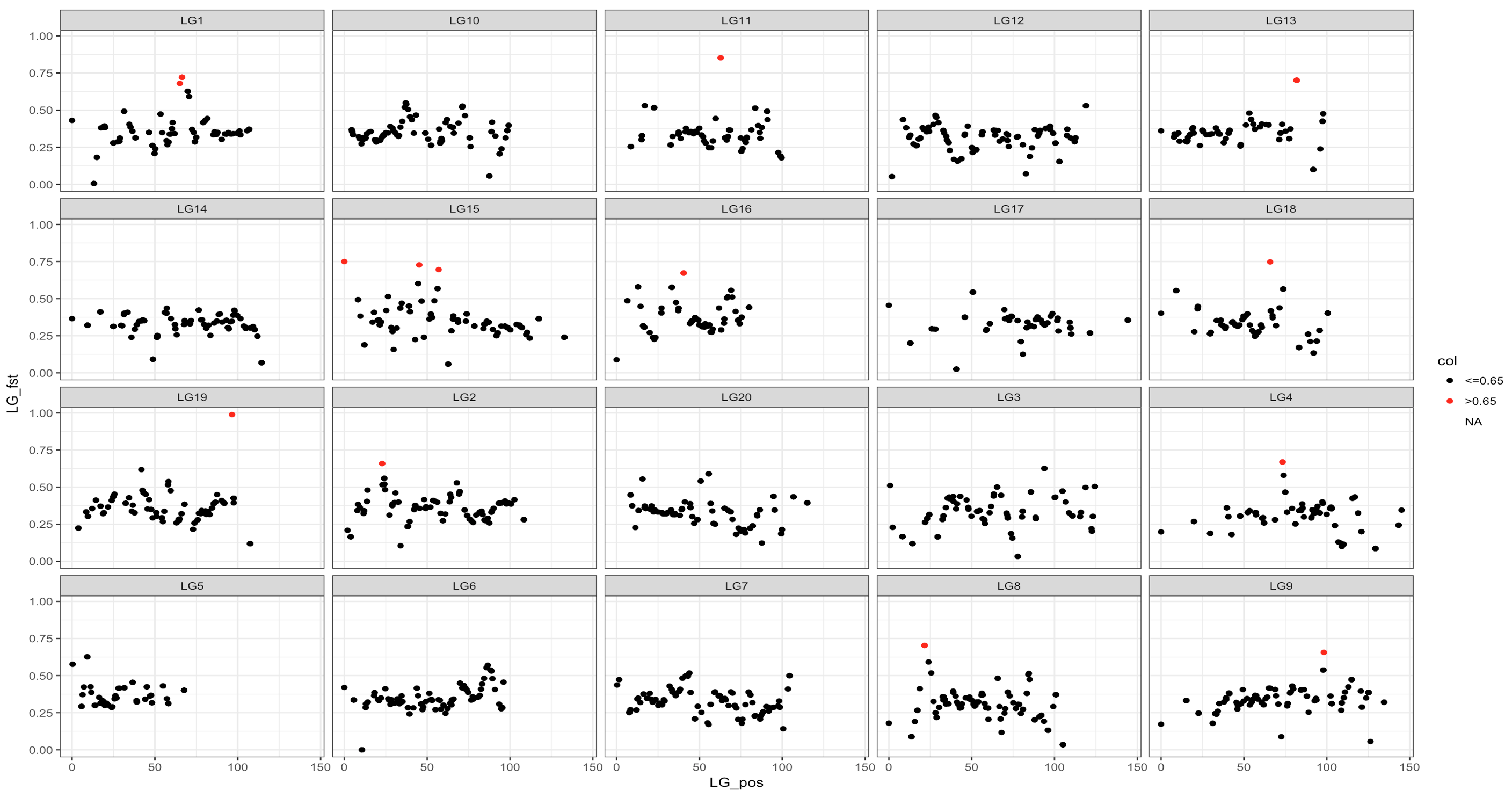


Figure 2: Plots with Linkage Group Position sliding window

Figure 2 above shows the plots with the original code rewritten to slide the window by genomic position instead of SNP count. The plots changed significantly with this different sliding window. The number of total points decreased and there are different and more numerous red points visible. The overall shape of the plots is quite different as well– in general, there seems to be less variation and there are far less points stacked on top of each other. The drastic change makes sense since the sliding window now only averages data points that are within a certain genomic distance, rather than whatever 20 data points happened to be closest, regardless of how far away they physically were. This new sliding window based on genomic positions should be a lot more meaningful in what it shows.

# Source Code:

---

title: "Roda et al\_tips"

output:

pdf\_document: default

html\_document: default

---

install.packages("zoo")

#Tips for the Roda et al data:

```{r}

#Getting started:

library(ggplot2)

roda<-read.csv("data/Fst\_BSA\_wLinkageGrp.csv", header=T)

roda<-roda[order(roda$LinkageGroup, roda$LinkageGroup\_Position),]

#We want to look at each linkage group at a time:

rd<-split(roda, roda$LinkageGroup)

SWlwg = function(data\_x, data\_y, window, step\_size){

data\_x = as.numeric(data\_x)

data\_y = as.numeric(data\_y)

x\_mean=rep(0,144)

y\_mean=rep(0,length(x\_mean))

floor = 0

for(i in 1:length(x\_mean)){

vals=which(data\_x>=floor & data\_x<=floor+window)

x\_mean[i]=sum(data\_x[vals])/length(vals)

y\_mean[i]=sum(data\_y[vals])/length(vals)

floor = floor + step\_size

}

out <- data.frame(x\_mean, y\_mean)

}

```

```{r}

SWlwg\_LG1 <- SWlwg(rd$GR1$LinkageGroup\_Position, rd$GR1$Fst, 5, 1)

LG1\_pos<-SWlwg\_LG1[[1]]

LG1\_Fst<-SWlwg\_LG1[[2]]

LG1\_LG<-rep("LG1",length(LG1\_pos))

SWlwg\_LG2 <- SWlwg(rd$GR2$LinkageGroup\_Position, rd$GR2$Fst, 5, 1)

LG2\_pos<-SWlwg\_LG2[[1]]

LG2\_Fst<-SWlwg\_LG2[[2]]

LG2\_LG<-rep("LG2",length(LG2\_pos))

SWlwg\_LG3 <- SWlwg(rd$GR3$LinkageGroup\_Position, rd$GR3$Fst, 5, 1)

LG3\_pos<-SWlwg\_LG3[[1]]

LG3\_Fst<-SWlwg\_LG3[[2]]

LG3\_LG<-rep("LG3",length(LG3\_pos))

SWlwg\_LG4 <- SWlwg(rd$GR4$LinkageGroup\_Position, rd$GR4$Fst, 5, 1)

LG4\_pos<-SWlwg\_LG4[[1]]

LG4\_Fst<-SWlwg\_LG4[[2]]

LG4\_LG<-rep("LG4",length(LG4\_pos))

SWlwg\_LG5 <- SWlwg(rd$GR5$LinkageGroup\_Position, rd$GR5$Fst, 5, 1)

LG5\_pos<-SWlwg\_LG5[[1]]

LG5\_Fst<-SWlwg\_LG5[[2]]

LG5\_LG<-rep("LG5",length(LG5\_pos))

SWlwg\_LG6 <- SWlwg(rd$GR6$LinkageGroup\_Position, rd$GR6$Fst, 5, 1)

LG6\_pos<-SWlwg\_LG6[[1]]

LG6\_Fst<-SWlwg\_LG6[[2]]

LG6\_LG<-rep("LG6",length(LG6\_pos))

SWlwg\_LG7 <- SWlwg(rd$GR7$LinkageGroup\_Position, rd$GR7$Fst, 5, 1)

LG7\_pos<-SWlwg\_LG7[[1]]

LG7\_Fst<-SWlwg\_LG7[[2]]

LG7\_LG<-rep("LG7",length(LG7\_pos))

SWlwg\_LG8 <- SWlwg(rd$GR8$LinkageGroup\_Position, rd$GR8$Fst, 5, 1)

LG8\_pos<-SWlwg\_LG8[[1]]

LG8\_Fst<-SWlwg\_LG8[[2]]

LG8\_LG<-rep("LG8",length(LG8\_pos))

SWlwg\_LG9 <- SWlwg(rd$GR9$LinkageGroup\_Position, rd$GR9$Fst, 5, 1)

LG9\_pos<-SWlwg\_LG9[[1]]

LG9\_Fst<-SWlwg\_LG9[[2]]

LG9\_LG<-rep("LG9",length(LG9\_pos))

SWlwg\_LG10 <- SWlwg(rd$GR10$LinkageGroup\_Position, rd$GR10$Fst, 5, 1)

LG10\_pos<-SWlwg\_LG10[[1]]

LG10\_Fst<-SWlwg\_LG10[[2]]

LG10\_LG<-rep("LG10",length(LG10\_pos))

SWlwg\_LG11 <- SWlwg(rd$GR11$LinkageGroup\_Position, rd$GR11$Fst, 5, 1)

LG11\_pos<-SWlwg\_LG11[[1]]

LG11\_Fst<-SWlwg\_LG11[[2]]

LG11\_LG<-rep("LG11",length(LG11\_pos))

SWlwg\_LG12 <- SWlwg(rd$GR12$LinkageGroup\_Position, rd$GR12$Fst, 5, 1)

LG12\_pos<-SWlwg\_LG12[[1]]

LG12\_Fst<-SWlwg\_LG12[[2]]

LG12\_LG<-rep("LG12",length(LG12\_pos))

SWlwg\_LG13 <- SWlwg(rd$GR13$LinkageGroup\_Position, rd$GR13$Fst, 5, 1)

LG13\_pos<-SWlwg\_LG13[[1]]

LG13\_Fst<-SWlwg\_LG13[[2]]

LG13\_LG<-rep("LG13",length(LG13\_pos))

SWlwg\_LG14 <- SWlwg(rd$GR14$LinkageGroup\_Position, rd$GR14$Fst, 5, 1)

LG14\_pos<-SWlwg\_LG14[[1]]

LG14\_Fst<-SWlwg\_LG14[[2]]

LG14\_LG<-rep("LG14",length(LG14\_pos))

SWlwg\_LG15 <- SWlwg(rd$GR15$LinkageGroup\_Position, rd$GR15$Fst, 5, 1)

LG15\_pos<-SWlwg\_LG15[[1]]

LG15\_Fst<-SWlwg\_LG15[[2]]

LG15\_LG<-rep("LG15",length(LG15\_pos))

SWlwg\_LG16 <- SWlwg(rd$GR16$LinkageGroup\_Position, rd$GR16$Fst, 5, 1)

LG16\_pos<-SWlwg\_LG16[[1]]

LG16\_Fst<-SWlwg\_LG16[[2]]

LG16\_LG<-rep("LG16",length(LG16\_pos))

SWlwg\_LG17 <- SWlwg(rd$GR17$LinkageGroup\_Position, rd$GR17$Fst, 5, 1)

LG17\_pos<-SWlwg\_LG17[[1]]

LG17\_Fst<-SWlwg\_LG17[[2]]

LG17\_LG<-rep("LG17",length(LG17\_pos))

SWlwg\_LG18 <- SWlwg(rd$GR18$LinkageGroup\_Position, rd$GR18$Fst, 5, 1)

LG18\_pos<-SWlwg\_LG18[[1]]

LG18\_Fst<-SWlwg\_LG18[[2]]

LG18\_LG<-rep("LG18",length(LG18\_pos))

SWlwg\_LG19 <- SWlwg(rd$GR19$LinkageGroup\_Position, rd$GR19$Fst, 5, 1)

LG19\_pos<-SWlwg\_LG19[[1]]

LG19\_Fst<-SWlwg\_LG19[[2]]

LG19\_LG<-rep("LG19",length(LG19\_pos))

SWlwg\_LG20 <- SWlwg(rd$GR20$LinkageGroup\_Position, rd$GR20$Fst, 5, 1)

LG20\_pos<-SWlwg\_LG20[[1]]

LG20\_Fst<-SWlwg\_LG20[[2]]

LG20\_LG<-rep("LG20",length(LG20\_pos))

```

```{r}

#putting all the windowed values back in 1 dataframe:

LG\_names<-rbind(LG1\_LG, LG2\_LG, LG3\_LG, LG4\_LG, LG5\_LG, LG6\_LG, LG7\_LG, LG8\_LG, LG9\_LG, LG10\_LG, LG11\_LG, LG12\_LG, LG13\_LG, LG14\_LG, LG15\_LG, LG16\_LG, LG17\_LG, LG18\_LG, LG19\_LG, LG20\_LG)

LG\_pos<-rbind(LG1\_pos, LG2\_pos, LG3\_pos, LG4\_pos, LG5\_pos, LG6\_pos, LG7\_pos, LG8\_pos, LG9\_pos, LG10\_pos, LG11\_pos, LG12\_pos, LG13\_pos, LG14\_pos, LG15\_pos, LG16\_pos, LG17\_pos, LG18\_pos, LG19\_pos, LG20\_pos)

LG\_fst<-rbind(LG1\_Fst, LG2\_Fst, LG3\_Fst, LG4\_Fst, LG5\_Fst, LG6\_Fst, LG7\_Fst, LG8\_Fst, LG9\_Fst, LG10\_Fst, LG11\_Fst, LG12\_Fst, LG13\_Fst, LG14\_Fst, LG15\_Fst, LG16\_Fst, LG17\_Fst, LG18\_Fst, LG19\_Fst, LG20\_Fst)

LG\_names<-as.vector(LG\_names)

LG\_pos<-as.vector(LG\_pos)

LG\_fst<-as.vector(LG\_fst)

data<-cbind(LG\_names, LG\_pos, LG\_fst)

data<-as.data.frame(data)

```

```{r}

# we've coerced the data a few times, so we need to reformat our numeric data in order for it to plot in ggplot

data[[2]] <- as.numeric(as.character(data[[2]]))

data[[3]] <- as.numeric(as.character(data[[3]]))

#Let's plot it now:

ggplot(data, aes(x=LG\_pos, y=LG\_fst))+geom\_point()+facet\_wrap(~LG\_names)

#lets plot some high Fst values in another color. We'll use 0.65 as a cutoff

data$col <- cut(data$LG\_fst,

breaks = c(-Inf, 0.65, Inf),

labels = c("<=0.65", ">0.65"))

ggplot(data, aes(x=LG\_pos, y=LG\_fst, color=col))+ geom\_point()+facet\_wrap(~LG\_names)+scale\_color\_manual(values=c("black", "red"))+theme\_bw()

#

```