allFill_Kvary_AlleleFreq

December 20, 2017

```
In [2]: library(plyr)
        library(ggplot2)
        library(data.table)
        library(Rmisc)
        library(dplyr)
```

```
Loading required package: lattice
-------------------------------------------------------------------------------
data.table + dplyr code now lives in dtplyr.
Please library(dtplyr)!
-------------------------------------------------------------------------------

Attaching package: dplyr

The following objects are masked from package:data.table:

    between, first, last

The following objects are masked from package:plyr:

    arrange, count, desc, failwith, id, mutate, rename, summarise,
    summarize

The following objects are masked from package:stats:

    filter, lag

The following objects are masked from package:base:

    intersect, setdiff, setequal, union
```

```
In [8]: #Working directory for datasets varying migration rate and bottleneck proportion
        setwd('/Users/jamessantangelo/Documents/Academia/Doctorate_PhD/Projects/SEC_Simulating.e
        volutionary.clines/SEC_Data/summary-datasets/allFill_Kvary_AlleleFreq')

        # Dataset for mean slope and proportions by parameter combination
        MeansProps <- fread('20171218_MeansProps_allFill_Kvary_AlleleFreq.csv', header = T)
        MeansProps$AlleleFreqs <- paste(MeansProps$pA_start, MeansProps$pB_start, sep = "-")
        MeansProps$AlleleFreqs <- as.factor(as.character(MeansProps$AlleleFreqs))
        MeansProps$AlleleFreqs <- factor(MeansProps$AlleleFreqs, levels=c("0.1-0.1", "0.1-0.5",
        "0.5-0.1", "0.1-0.9",
                                                                          "0.9-0.1", "0.5-0.5",
        "0.5-0.9", "0.9-0.5",
                                                                          "0.9-0.9"))
        MeansProps$Mig_rate = as.factor(as.character(MeansProps$Mig_rate))
        MeansProps$prop_sigPos[MeansProps$prop_sigPos == 1.000] <- 0.0
```

```r
# datRegSummary <- fread("20171218_RegSummary_allFill_Kvary_AlleleFreq.csv")
# datRegSummary$AlleleFreqs <- paste(datRegSummary$pA_start, datRegSummary$pB_start, sep
= "-")
# datRegSummary$AlleleFreqs <- as.factor(as.character(datRegSummary$AlleleFreqs))
# datRegSummary$AlleleFreqs <- factor(datRegSummary$AlleleFreqs, levels=c("0.1-0.1",
"0.1-0.5", "0.5-0.1", "0.1-0.9",
#                                                        "0.9-0.1", "0.5-0.5",
"0.5-0.9", "0.9-0.5",
#                                                        "0.9-0.9"))
# datRegSummary$Mig_rate = as.factor(as.character(datRegSummary$Mig_rate))
```

In [37]: #Data subsets
```r
MeansProps_Gen250 <- subset(MeansProps, seq == "250")
MeansProps_NoMig <- subset(MeansProps, Mig_rate == "0")
MeansProps_NoMig_Gen250 <- subset(MeansProps, seq == "250" & Mig_rate == "0")
```

In [26]: str(MeansProps)

```
Classes data.table and 'data.frame':  13500 obs. of  22 variables:
 $ Mig_rate   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ pA_start   : num  0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 ...
 $ pB_start   : num  0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 ...
 $ seq        : int  1 2 3 4 5 6 7 8 9 10 ...
 $ mean       : num  7.61e-05 8.87e-05 9.68e-05 1.06e-04 1.22e-04 ...
 $ sd         : num  0 0.000226 0.000314 0.000386 0.000435 ...
 $ n          : int  1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 ...
 $ se         : num  0.00 7.13e-06 9.92e-06 1.22e-05 1.38e-05 ...
 $ ci_mean    : num  0.00 1.40e-05 1.94e-05 2.39e-05 2.70e-05 ...
 $ prop_sigPos: num  1 0.242 0.221 0.221 0.24 0.222 0.233 0.234 0.234 0.23 ...
 $ prop_pos   : num  1 0.71 0.715 0.721 0.731 0.735 0.739 0.728 0.73 0.727 ...
 $ se_pos     : num  0 0.0143 0.0143 0.0142 0.014 ...
 $ ci_pos     : num  0 0.0281 0.028 0.0278 0.0275 ...
 $ se_sigPos  : num  0 0.0135 0.0131 0.0131 0.0135 ...
 $ ci_sigPos  : num  0 0.0265 0.0257 0.0257 0.0265 ...
 $ prop_sigNeg: num  0 0.018 0.014 0.009 0.008 0.012 0.008 0.01 0.01 0.008 ...
 $ prop_neg   : num  0 0.29 0.285 0.279 0.269 0.265 0.261 0.272 0.27 0.273 ...
 $ se_neg     : num  0 0.0143 0.0143 0.0142 0.014 ...
 $ ci_neg     : num  0 0.0281 0.028 0.0278 0.0275 ...
 $ se_sigNeg  : num  0 0.0042 0.00372 0.00299 0.00282 ...
 $ ci_sigNeg  : num  0 0.00824 0.00728 0.00585 0.00552 ...
 $ AlleleFreqs: Factor w/ 9 levels "0.1-0.1","0.1-0.5",..: 1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, ".internal.selfref")=<externalptr>
```

In [6]: #Themes used for plotting
```r
ng1=theme(aspect.ratio=0.7,panel.background = element_blank(),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border=element_blank(),
        axis.line.x = element_line(color="black",size=1),
        axis.line.y = element_line(color="black",size=1),
        axis.ticks=element_line(color="black"),
        axis.text=element_text(color="black",size=15),
        axis.title=element_text(color="black",size=1),
        axis.title.y=element_text(vjust=2,face="bold",size=15),
        axis.title.x=element_text(vjust=0.1,face="bold",size=15),
        axis.text.x=element_text(size=13),
        axis.text.y=element_text(size=13),
        legend.position = "right", legend.direction="vertical",
        legend.text=element_text(size=11), legend.key = element_rect(fill = "white"),
        legend.title = element_text(size=13,face="bold"),legend.key.size = unit(0.5,
"cm"))
```
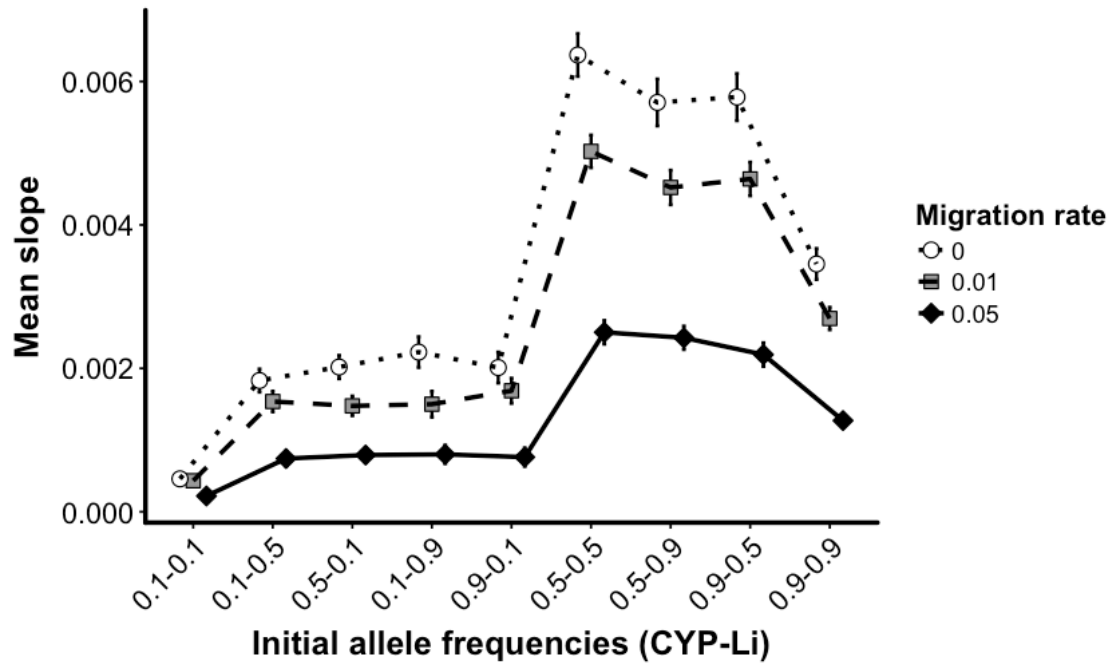
```
ng1.45=theme(aspect.ratio=0.7,panel.background = element_blank(),
            panel.grid.major = element_blank(),
            panel.grid.minor = element_blank(),
            panel.border=element_blank(),
            axis.line.x = element_line(color="black",size=1),
            axis.line.y = element_line(color="black",size=1),
            axis.ticks=element_line(color="black"),
            axis.text=element_text(color="black",size=15),
            axis.title=element_text(color="black",size=1),
            axis.title.y=element_text(vjust=2,face="bold",size=15),
            axis.title.x=element_text(vjust=0.1,face="bold",size=15),
            axis.text.x=element_text(size=13,angle=45,hjust=1),
            axis.text.y=element_text(size=13),
            legend.position = "right", legend.direction="vertical",
            legend.text=element_text(size=11), legend.key = element_rect(fill = "white"),
            legend.title = element_text(size=13,face="bold"),legend.key.size = unit(0.5,
"cm"))
```

1 Mean slope by allele frequencies and migration rate
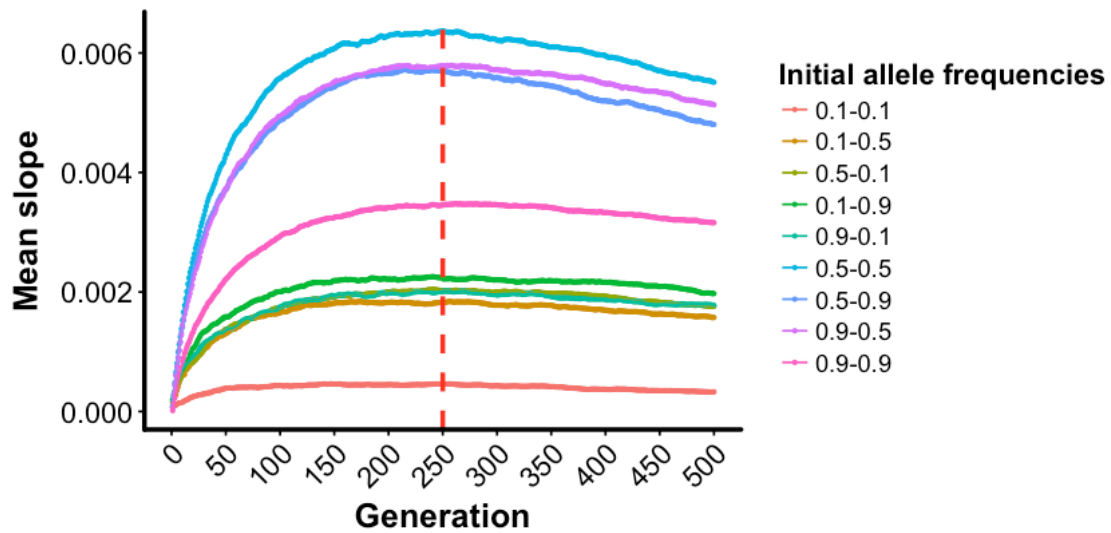
```
In [22]: MeanSlope_AlleleFreq_Mig <- ggplot(MeansProps_Gen250, aes(x = factor(AlleleFreqs), y =
    mean, group = factor(Mig_rate))) +
        geom_errorbar(aes(ymin = mean - ci_mean, ymax = mean + ci_mean), width=0.15,
size=0.7,
        position = position_dodge(width = 0.5)) +
        geom_point(size = 3, aes(fill = factor(Mig_rate), shape = factor(Mig_rate)),
position = position_dodge(width = 0.5)) +
        geom_line(size = 1, aes(linetype = factor(Mig_rate)), position =
position_dodge(width = 0.5)) +
        scale_shape_manual(labels = c("0", "0.01", "0.05"), values = c(21, 22, 23)) +
        scale_fill_manual(labels = c("0", "0.01", "0.05"), values = c("white", "grey60",
"black")) +
        scale_linetype_manual(labels = c("0", "0.01", "0.05"), values = c("dotted",
"dashed", "solid")) +
        ylab("Mean slope") + xlab("Initial allele frequencies (CYP-Li)") +
        labs(fill = "Migration rate", shape = "Migration rate", linetype = "Migration rate")
    + ng1.45
    MeanSlope_AlleleFreq_Mig
```

1.1 Effects of allele frequencies on proportion of negative and positive clines

```
In [34]: PropSigPos_AlleleFreq_Mig <- ggplot(MeansProps_Gen250, aes(x = factor(AlleleFreqs), y =
         prop_sigPos, group = factor(Mig_rate))) +
             geom_errorbar(aes(ymin = prop_sigPos - ci_sigPos, ymax = prop_sigPos + ci_sigPos),
         width=0.15, size=0.7,
             position = position_dodge(width = 0.5)) +
             geom_point(size = 3, aes(fill = factor(Mig_rate), shape = factor(Mig_rate)),
         position = position_dodge(width = 0.5)) +
             geom_line(size = 1, aes(linetype = factor(Mig_rate)), position =
         position_dodge(width = 0.5)) +
             scale_shape_manual(labels = c("0", "0.01", "0.05"), values = c(21, 22, 23)) +
             scale_fill_manual(labels = c("0", "0.01", "0.05"), values = c("white", "grey60",
         "black")) +
             scale_linetype_manual(labels = c("0", "0.01", "0.05"), values = c("dotted",
         "dashed", "solid")) +
             ylab("Proportion sig. positive slopes") + xlab("Initial allele frequencies (CYP-
         Li)") +
```

```
        labs(fill = "Migration rate", shape = "Migration rate", linetype = "Migration rate")
    + ng1.45
    PropSigPos_AlleleFreq_Mig
```
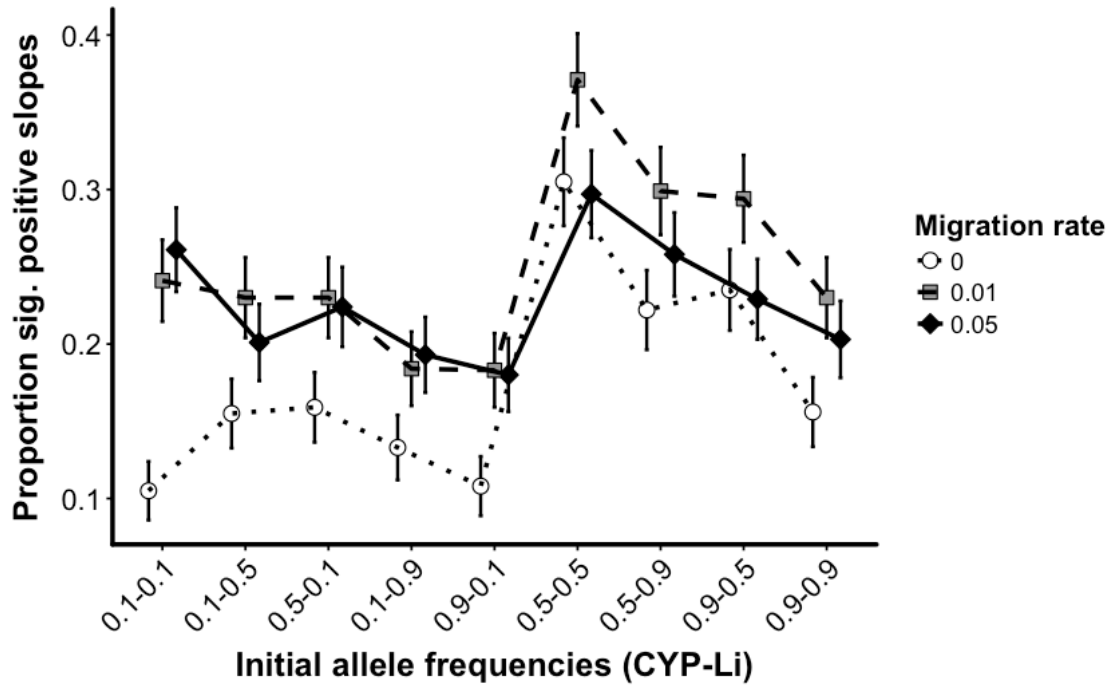
In [43]:
```
PropSigPos_byGen <- ggplot(MeansProps_NoMig, aes(x = seq, y = prop_sigPos, group =
    AlleleFreqs, color = AlleleFreqs)) +
        ylab("Proportion sig. positive slopes") + xlab("Generation") + geom_point(size =
    0.5, aes(color = AlleleFreqs)) +
        geom_line(size = 0.5, aes(color = AlleleFreqs)) +
        coord_cartesian(xlim = c(0, 500)) + scale_x_continuous(breaks = seq(from = 0, to =
    500, by = 50)) +
        labs(color = "Initial allele frequencies") +
        geom_vline(data = Kvary_Mig_Summary_Cyan, xintercept = 250, size = 1, colour =
    "#FF3721", linetype = "dashed") + ng1.45
    PropSigPos_byGen
```
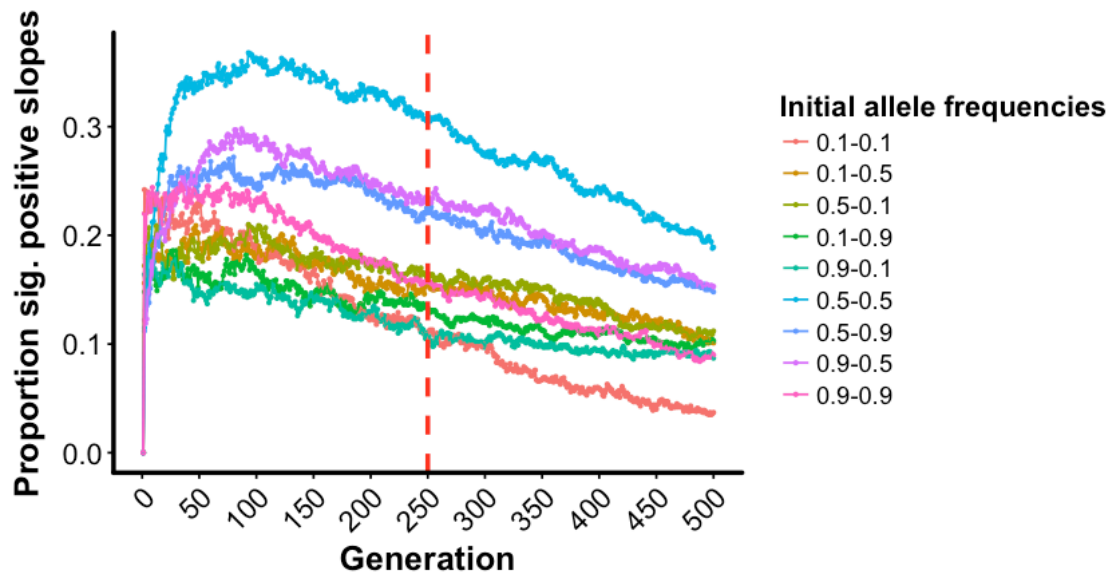
In [38]: *#Extra columns that will be used to create melting dataset for plotting proportion of*
positive and negative slopes
MeansProps_melt <- dplyr::select(MeansProps_NoMig_Gen250, AlleleFreqs, pA_start,
pB_start, prop_sigPos, ci_sigPos, prop_sigNeg,
  ci_sigNeg)

#Melt dataframe
dm1 <- melt(MeansProps_melt[,c("AlleleFreqs", "prop_sigPos", "ci_sigPos")],
          id=c("AlleleFreqs", "ci_sigPos"))
dm2 <- melt(MeansProps_melt[,c("AlleleFreqs", "prop_sigNeg", "ci_sigNeg")],
   id=c("AlleleFreqs", "ci_sigNeg"))

#Rename columns
setnames(dm1, old = c("ci_sigPos"), new = "ci")
setnames(dm2, old = c("ci_sigNeg"), new = "ci")

#Merge melted dataframes
MeansProps_melt <- rbind(dm1, dm2)
MeansProps_melt

AlleleFreqs	ci	variable	value
0.1-0.1	0.019000378	prop_sigPos	0.105
0.1-0.5	0.022431085	prop_sigPos	0.155
0.1-0.9	0.021047060	prop_sigPos	0.133
0.5-0.1	0.022664839	prop_sigPos	0.159
0.5-0.5	0.028536348	prop_sigPos	0.305
0.5-0.9	0.025758606	prop_sigPos	0.222
0.9-0.1	0.019237577	prop_sigPos	0.108
0.9-0.5	0.026279719	prop_sigPos	0.235
0.9-0.9	0.022490007	prop_sigPos	0.156
0.1-0.1	0.000000000	prop_sigNeg	0.000
0.1-0.5	0.002769085	prop_sigNeg	0.002
0.1-0.9	0.003389724	prop_sigNeg	0.003
0.5-0.1	0.001959020	prop_sigNeg	0.001
0.5-0.5	0.000000000	prop_sigNeg	0.000
0.5-0.9	0.001959020	prop_sigNeg	0.001
0.9-0.1	0.004371723	prop_sigNeg	0.005
0.9-0.5	0.003912152	prop_sigNeg	0.004
0.9-0.9	0.006166996	prop_sigNeg	0.010

```
In [40]: #Plot proportion of significant slopes by migration rate
        PropSig_AlleleFreq <- ggplot(MeansProps_melt, aes(x = AlleleFreqs, y = value,
            shape = variable, fill = variable)) +
            geom_errorbar(aes(ymin = value - ci, ymax = value + ci), width=0.15, size = 0.7,
        position = position_dodge(width = 0.55)) +
            ylab("Proportion sig. slopes") + xlab("Initial allele frequencies (A-B)") +
            geom_point(size = 3.5, color = "black", position = position_dodge(width = 0.55)) +
            scale_shape_manual(labels = c("Positive", "Negative"),values=c(22, 23))+
            scale_fill_manual(labels = c("Positive", "Negative"),values=c("white", "black")) +
            labs(shape = 'Cline direction', fill = 'Cline direction') + ng1.45
        PropSig_AlleleFreq
```
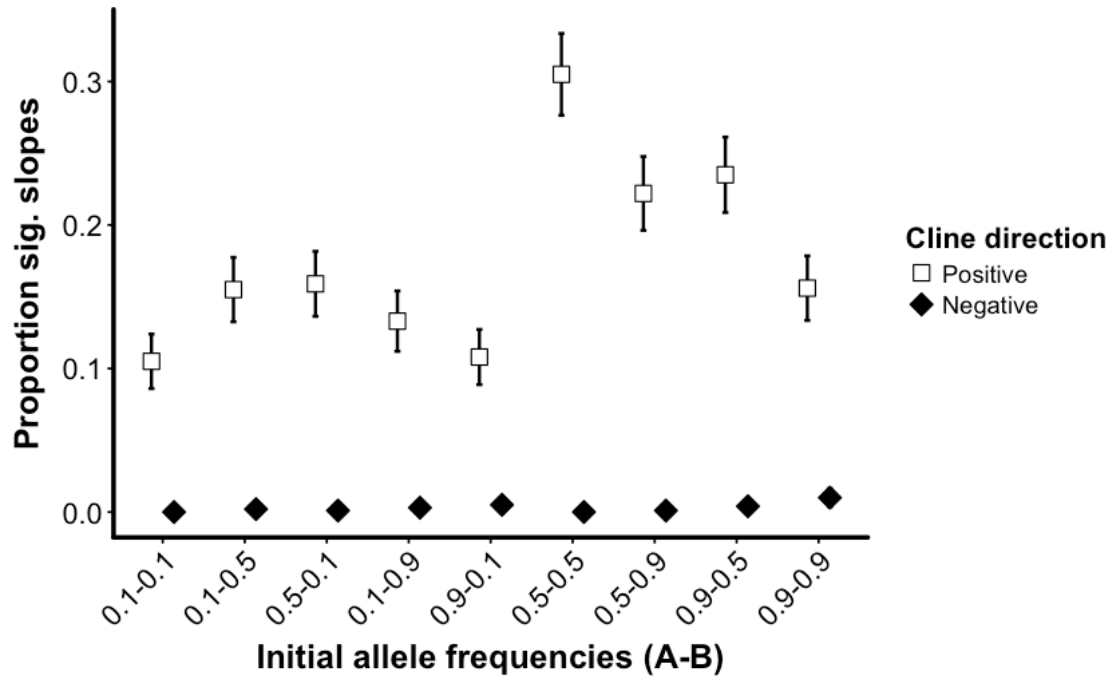
```
In [ ]: path = "/scratch/research/projects/trifolium/SEC_Simulation.Evolutionary.Clines/SEC_Sync
        /SEC_Figures/Drift.Migration/Mig_Bot_Vary"

        # ggsave("PropSig_Mig.pdf", plot = PropSig_Mig, device = "pdf", width = 6.0, height =
        6.0, path = path, dpi = 600)
        # ggsave("HistSlopes_pA_NoMig.pdf", plot = HistSlopes_pA_NoMig, device = "pdf", width =
        6.0, height = 6.0, path = path, dpi = 600)
        # ggsave("HistSlopes_HCN_NoMig.pdf", plot = HistSlopes_Cyan_NoMig, device = "pdf", width
        = 6.0, height = 6.0, path = path, dpi = 600)
        # ggsave("MeanSlope_Mig_Type.pdf", plot = MeanSlope_Mig_Type, device = "pdf", width =
        6.0, height = 6.0, path = path, dpi = 600)

        ggsave("MeanSlope_Mig_Type.pdf", plot = MeanSlope_Mig_Type, device = "pdf", width = 6.0,
        height = 6.0, path = path, dpi = 600)
        ggsave("PropSig_Mig.pdf", plot = PropSig_Mig, device = "pdf", width = 6.0, height = 6.0,
        path = path, dpi = 600)
```