What shapes genome-wide diversity?

A case study in Drosophila melanogaster

Gustavo V. Barroso and Julien Y. Dutheil

September 1st 2021

Contents

1 Drosophila-like neutral simulations of 2L (True Landscapes) 4
1.1 50 kb scale . . . 4
1.1.1 Replicate 1 . . . 4
1.1.2 Replicate 2 . . . 9
1.1.3 Replicate 3 . . . 12
1.1.4 Replicate 4 . . . 15
1.1.5 Replicate 5 . . . 18
1.1.6 Replicate 6 . . . 21
1.1.7 Replicate 7 . . . 24
1.1.8 Replicate 8 . . . 27
1.1.9 Replicate 9 . . . 30
1.1.10 Replicate 10 . . . 33
1.1.11 all replicates: . . . 36
1.2 200 kb scale . . . 39
1.2.1 Replicate 1 . . . 39
1.2.2 Replicate 2 . . . 43
1.2.3 Replicate 3 . . . 46
1.2.4 Replicate 4 . . . 49
1.2.5 Replicate 5 . . . 52
1.2.6 Replicate 6 . . . 55
1.2.7 Replicate 7 . . . 58
1.2.8 Replicate 8 . . . 61
1.2.9 Replicate 9 . . . 64
1.2.10 Replicate 10 . . . 67
1.2.11 all replicates: . . . 70
1.3 1 Mb scale . . . 72
1.3.1 Replicate 1 . . . 73
1.3.2 Replicate 2 . . . 76
1.3.3 Replicate 3 . . . 80
1.3.4 Replicate 4 . . . 84
1.3.5 Replicate 5 . . . 88
1.3.6 Replicate 6 . . . 91
1.3.7 Replicate 7 . . . 94
1.3.8 Replicate 8 . . . 97
1.3.9 Replicate 9 . . . 100
1.3.10 Replicate 10 . . . 103
1.3.11 all replicates: . . . 106

2 Drosophila-like neutral simulations of chr 2L (Inferred Landscapes) 108
2.1 50 kb scale . . . 109
2.1.1 Replicate 1 . . . 109
2.1.2 Replicate 2 . . . 114
2.1.3 Replicate 3 . . . 119
2.1.4 Replicate 4 . . . 124
2.1.5 Replicate 5 . . . 129
2.1.6 Replicate 6 . . . 134
2.1.7 Replicate 7 . . . 139
2.1.8 Replicate 8 . . . 144
2.1.9 Replicate 9 . . . 149
2.1.10 Replicate 10 . . . 154
2.1.11 all replicates: . . . 159
2.1.12 Plots . . . 162
2.2 200 kb scale . . . 163
2.2.1 Replicate 1 . . . 163
2.2.2 Replicate 2 . . . 168
2.2.3 Replicate 3 . . . 173
2.2.4 Replicate 4 . . . 178
2.2.5 Replicate 5 . . . 183
2.2.6 Replicate 6 . . . 188
2.2.7 Replicate 7 . . . 193
2.2.8 Replicate 8 . . . 198
2.2.9 Replicate 9 . . . 203
2.2.10 Replicate 10 . . . 208
2.2.11 all replicates: . . . 213
2.2.12 Plots . . . 216
2.3 1 Mb scale . . . 217
2.3.1 Replicate 1 . . . 217
2.3.2 Replicate 2 . . . 222
2.3.3 Replicate 3 . . . 227
2.3.4 Replicate 4 . . . 232
2.3.5 Replicate 5 . . . 237
2.3.6 Replicate 6 . . . 242
2.3.7 Replicate 7 . . . 247
2.3.8 Replicate 8 . . . 252
2.3.9 Replicate 9 . . . 257
2.3.10 Replicate 10 . . . 262
2.3.11 all replicates: . . . 267
2.3.12 Plots . . . 270

3 Real Drosophila data. 272
3.1 50 kb windows . . . 275
3.2 200 kb windows . . . 290
3.3 1 Mb windows . . . 305
3.4 Evolutionary (Protein) Rates . . . 320

4 R^2 plot with real, simulated (true) and simulated (inferred) data. 326

5 Simulations with BGS and flat Mutation Landscape in SLiM 328
5.1 50 kb scale . . . 328
5.1.1 Replicate 1 . . . 328
5.1.2 Replicate 2 . . . 332
5.1.3 Replicate 3 . . . 336

5.1.4 Replicate 4 . . . 340
5.1.5 Replicate 5 . . . 344
5.1.6 Replicate 6 . . . 348
5.1.7 Replicate 7 . . . 352
5.1.8 Replicate 8 . . . 356
5.1.9 Replicate 9 . . . 360
5.1.10 Replicate 10 . . . 364
5.2 200 kb scale . . . 368
5.2.1 Replicate 1 . . . 368
5.2.2 Replicate 2 . . . 372
5.2.3 Replicate 3 . . . 376
5.2.4 Replicate 4 . . . 380
5.2.5 Replicate 5 . . . 384
5.2.6 Replicate 6 . . . 388
5.2.7 Replicate 7 . . . 392
5.2.8 Replicate 8 . . . 396
5.2.9 Replicate 9 . . . 400
5.2.10 Replicate 10 . . . 404
5.3 1 Mb scale . . . 408
5.3.1 Replicate 1 . . . 408
5.3.2 Replicate 2 . . . 412
5.3.3 Replicate 3 . . . 416
5.3.4 Replicate 4 . . . 420
5.3.5 Replicate 5 . . . 425
5.3.6 Replicate 6 . . . 429
5.3.7 Replicate 7 . . . 434
5.3.8 Replicate 8 . . . 438
5.3.9 Replicate 9 . . . 443
5.3.10 Replicate 10 . . . 447
5.4 all replicates: . . . 452

6 Simulations with BGS and Gamma Mutation Landscape in SLiM 453
6.1 50 kb scale . . . 453
6.1.1 Replicate 1 . . . 454
6.1.2 Replicate 2 . . . 458
6.1.3 Replicate 3 . . . 462
6.1.4 Replicate 4 . . . 466
6.1.5 Replicate 5 . . . 470
6.1.6 Replicate 6 . . . 474
6.1.7 Replicate 7 . . . 478
6.1.8 Replicate 8 . . . 482
6.1.9 Replicate 9 . . . 486
6.1.10 Replicate 10 . . . 490
6.2 all replicates: . . . 495
6.3 200 kb scale . . . 496
6.3.1 Replicate 1 . . . 496
6.3.2 Replicate 2 . . . 501
6.3.3 Replicate 3 . . . 505
6.3.4 Replicate 4 . . . 509
6.3.5 Replicate 5 . . . 513
6.3.6 Replicate 6 . . . 517
6.3.7 Replicate 7 . . . 521
6.3.8 Replicate 8 . . . 525
6.3.9 Replicate 9 . . . 530

6.3.10 Replicate 10 . . . 534
6.4 all replicates: . . . 538
6.5 1 Mb scale . . . 539
6.5.1 Replicate 1 . . . 539
6.5.2 Replicate 2 . . . 544
6.5.3 Replicate 3 . . . 548
6.5.4 Replicate 4 . . . 552
6.5.5 Replicate 5 . . . 556
6.5.6 Replicate 6 . . . 560
6.5.7 Replicate 7 . . . 564
6.5.8 Replicate 8 . . . 568
6.5.9 Replicate 9 . . . 572
6.5.10 Replicate 10 . . . 577
6.6 all replicates: . . . 581

7 Plots **582**

1 Drosophila-like neutral simulations of 2L (True Landscapes)

Second batch of 10 replicate datasets simulated using parameters as inferred from real Drosophila data.

Here we present results using true (simulated) genomic landscapes.

1.1 50 kb scale

```
r2.sim.50kb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.sim.50kb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.sim.50kb) <- reps

sim.rho.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/dm.sim.rho.50000.txt", header = T)
sim.theta.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/dm.sim.theta.50000.txt", header = T
```

1.1.1 Replicate 1

```
rep_1.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_1/rep_1.pi.50kb.bedgraph", header
rep_1.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_1/sim.tmrca.50k.map", header =

sim.lands.50kb <- as.data.frame(cbind(rep_1.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_1.tmr
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 36665911, p-value = 0.6511
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.01850036
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 35649572, p-value = 0.8119
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.00973136
```

```
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)

sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)

# for merging:
sim.lands.50kb.rep5 <- sim.lands.50kb
sim.lands.50kb.rep5$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = sim.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = sim.lands.50kb)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7492.648
## m.div.50kb.2  7 -7491.365
## m.div.50kb.3  8 -7491.767
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 2.1111, p-value = 0.9056
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.39214, p-value < 2.2e-16
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.839e-03 -2.676e-04 -1.006e-05  2.647e-04  2.976e-03
##
## Coefficients:
##                Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   2.069e-02  1.908e-05 1084.381   <2e-16 ***
## thetaC        1.309e+00  2.283e-03  573.279   <2e-16 ***
## rhoC          1.435e-02  6.455e-03    2.223   0.0266 *
## tmrcaC        2.342e-02  2.686e-04   87.195   <2e-16 ***
## thetaC:tmrcaC 1.483e+00  3.029e-02   48.983   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004673 on 595 degrees of freedom
## Multiple R-squared:  0.9982, Adjusted R-squared:  0.9982
## F-statistic: 8.37e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.50kb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```r
r2.sim.50kb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 1] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = sim.lands.50kb, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), method =

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = sim.lands.50kb, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), method

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = sim.lands.50kb, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = sim.lands.50kb, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7490.623
## g.div.50kb.2  8 -7570.956
## g.div.50kb.3  7 -7572.341
## g.div.50kb.4  7 -7490.654
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: sim.lands.50kb
##         AIC       BIC  logLik
##   -7572.341 -7541.562 3793.17
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.4774622
##
## Coefficients:
##                   Value  Std.Error   t-value p-value
## (Intercept)   0.0206901 0.00001905 1086.3241  0.0000
## thetaC        1.3155987 0.00236880  555.3850  0.0000
## rhoC          0.0087173 0.00554727    1.5715  0.1166
## tmrcaC        0.0234411 0.00026760   87.5965  0.0000
## thetaC:tmrcaC 1.4652196 0.03311231   44.2500  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.497
## rhoC           0.001  0.052
## tmrcaC         0.013  0.026 -0.004
## thetaC:tmrcaC  0.026  0.025  0.036  0.494
```

```
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -2.955335880 -0.608690299  0.006577071  0.610719662  4.719725050
##
## Residual standard error: 0.003393872
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.003575      1.004633      1.323896      1.325415
```

1.1.2 Replicate 2

```
rep_2.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_2/rep_2.pi.50kb.bedgraph", header
rep_2.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_2/sim.tmrca.50k.map", header =

sim.lands.50kb <- as.data.frame(cbind(rep_2.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_2.tmr
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 36499595, p-value = 0.7344
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.01388046
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 33177308, p-value = 0.05494
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.07840555
```

```r
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)

sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)

# for merging:
sim.lands.50kb.rep2 <- sim.lands.50kb
sim.lands.50kb.rep2$Replicate <- 2

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = sim.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = sim.lands.50kb)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7408.739
## m.div.50kb.2  7 -7409.225
## m.div.50kb.3  8 -7407.388
```

```r
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.9715, p-value = 0.3445
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.51438, p-value = 0.677
```

```r
hist(resid(m.div.50kb))
```

```r
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.857e-03 -2.640e-04  1.830e-06  2.990e-04  2.087e-03
##
## Coefficients:
##                 Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    2.057e-02  2.047e-05 1005.301   <2e-16 ***
## thetaC         1.299e+00  2.448e-03  530.815   <2e-16 ***
## rhoC          -1.009e-02  6.919e-03   -1.459    0.145
## tmrcaC         2.385e-02  2.831e-04   84.273   <2e-16 ***
## thetaC:tmrcaC  1.515e+00  3.222e-02   47.015   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.0005011 on 595 degrees of freedom
## Multiple R-squared:  0.9979, Adjusted R-squared:  0.9979
## F-statistic: 7.16e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.50kb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.sim.50kb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 2] <- anova.diversity$VarExp[4] * 100
```

1.1.3 Replicate 3

```
rep_3.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_3/rep_3.pi.50kb.bedgraph", header
rep_3.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_3/sim.tmrca.50k.map", header =

sim.lands.50kb <- as.data.frame(cbind(rep_3.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_3.tmr
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 37873663, p-value = 0.203
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.05204912
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 35383984, p-value = 0.6757
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
```

```
##        rho
## 0.01710883
```

```r
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)

sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)

# for merging:
sim.lands.50kb.rep3 <- sim.lands.50kb
sim.lands.50kb.rep3$Replicate <- 3

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.9214, p-value = 0.1563
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.50434, p-value = 0.595
```

```r
hist(resid(m.div.50kb))
```

```r
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.081e-03 -2.636e-04 -7.680e-06  2.600e-04  1.833e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0206720  0.0000187 1105.47   <2e-16 ***
## thetaC        1.3103632  0.0022312  587.29   <2e-16 ***
## rhoC          0.0064418  0.0063164    1.02    0.308
## tmrcaC        0.0236099  0.0002297  102.80   <2e-16 ***
## thetaC:tmrcaC 1.5012872  0.0254605   58.97   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004578 on 595 degrees of freedom
## Multiple R-squared:  0.9983, Adjusted R-squared:  0.9983
## F-statistic: 8.857e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```r
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.sim.50kb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.sim.50kb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 3] <- anova.diversity$VarExp[4] * 100
```

1.1.4 Replicate 4

```
rep_4.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_4/rep_4.pi.50kb.bedgraph", header
rep_4.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_4/sim.tmrca.50k.map", header =

sim.lands.50kb <- as.data.frame(cbind(rep_4.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_4.tmr
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 36135074, p-value = 0.9269
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##          rho
## -0.003754843
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 37039432, p-value = 0.4801
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.02887597
```

```
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)
```

```r
sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)

# for merging:
sim.lands.50kb.rep4 <- sim.lands.50kb
sim.lands.50kb.rep4$Replicate <- 4

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 2.0985, p-value = 0.8779
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.53836, p-value = 0.896
```

```r
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.787e-03 -2.851e-04 -3.163e-05  2.685e-04  1.746e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.050e-02  2.077e-05 987.094   <2e-16 ***
## thetaC        1.300e+00  2.476e-03 525.118   <2e-16 ***
## rhoC          1.386e-02  7.016e-03   1.976   0.0486 *
## tmrcaC        2.349e-02  2.643e-04  88.857   <2e-16 ***
## thetaC:tmrcaC 1.493e+00  3.134e-02  47.634   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005088 on 595 degrees of freedom
## Multiple R-squared:  0.9979, Adjusted R-squared:  0.9979
## F-statistic: 7.205e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.50kb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```
r2.sim.50kb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 4] <- anova.diversity$VarExp[4] * 100
```

1.1.5 Replicate 5

```
rep_5.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_5/rep_5.pi.50kb.bedgraph", header
rep_5.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_5/sim.tmrca.50k.map", header =

sim.lands.50kb <- as.data.frame(cbind(rep_5.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_5.tmr
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 35425520, p-value = 0.6965
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.01595505
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 37018948, p-value = 0.4888
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.02830697
```

```
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)

sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)
```

```r
# for merging:
sim.lands.50kb.rep5 <- sim.lands.50kb
sim.lands.50kb.rep5$Replicate <- 5

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 2.0931, p-value = 0.864
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.45962, p-value = 0.091
```

```r
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.966e-03 -2.902e-04 -1.674e-05  2.879e-04  2.110e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.062e-02  1.967e-05 1048.31   <2e-16 ***
## thetaC        1.308e+00  2.348e-03  557.06   <2e-16 ***
## rhoC          1.528e-03  6.645e-03    0.23    0.818
## tmrcaC        2.374e-02  2.674e-04   88.80   <2e-16 ***
## thetaC:tmrcaC 1.471e+00  3.126e-02   47.05   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004816 on 595 degrees of freedom
## Multiple R-squared:  0.9981, Adjusted R-squared:  0.9981
## F-statistic: 7.932e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.50kb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```
r2.sim.50kb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 5] <- anova.diversity$VarExp[4] * 100
```

1.1.6 Replicate 6

```
rep_6.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_6/rep_6.pi.50kb.bedgraph", header
rep_6.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_6/sim.tmrca.50k.map", header =

sim.lands.50kb <- as.data.frame(cbind(rep_6.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_6.tmr
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 36901608, p-value = 0.5403
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0250475
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 37734896, p-value = 0.2384
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04819447
```

```
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)

sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)
```

```r
# for merging:
sim.lands.50kb.rep6 <- sim.lands.50kb
sim.lands.50kb.rep6$Replicate <- 6

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 2.1478, p-value = 0.9618
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.46094, p-value = 0.083
```

```r
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0023630 -0.0002937 -0.0000107  0.0002498  0.0036466
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.058e-02  2.073e-05 993.114   <2e-16 ***
## thetaC         1.304e+00  2.479e-03 526.030   <2e-16 ***
## rhoC          -1.179e-02  7.005e-03  -1.683   0.0929 .
## tmrcaC         2.363e-02  2.752e-04  85.861   <2e-16 ***
## thetaC:tmrcaC  1.439e+00  3.174e-02  45.345   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005077 on 595 degrees of freedom
## Multiple R-squared:  0.998,  Adjusted R-squared:  0.9979
## F-statistic: 7.29e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.50kb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```
r2.sim.50kb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 6] <- anova.diversity$VarExp[4] * 100
```

1.1.7 Replicate 7

```
rep_7.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_7/rep_7.pi.50kb.bedgraph", header
rep_7.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_7/sim.tmrca.50k.map", header =

sim.lands.50kb <- as.data.frame(cbind(rep_7.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_7.tmr
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 38638560, p-value = 0.0728
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0732963
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 36196758, p-value = 0.8936
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##          rho
## -0.005468293
```

```
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)

sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)
```

```r
# for merging:
sim.lands.50kb.rep7 <- sim.lands.50kb
sim.lands.50kb.rep7$Replicate <- 7

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.959, p-value = 0.2918
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.47957, p-value = 0.238
```

```r
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.928e-03 -2.610e-04 -1.731e-05  2.569e-04  3.020e-03
##
## Coefficients:
##                 Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    2.061e-02  1.945e-05 1059.614   <2e-16 ***
## thetaC         1.313e+00  2.320e-03  565.892   <2e-16 ***
## rhoC          -5.753e-03  6.560e-03   -0.877    0.381
## tmrcaC         2.389e-02  2.547e-04   93.801   <2e-16 ***
## thetaC:tmrcaC  1.491e+00  3.019e-02   49.398   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004759 on 595 degrees of freedom
## Multiple R-squared:  0.9982, Adjusted R-squared:  0.9982
## F-statistic: 8.263e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.50kb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```
r2.sim.50kb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 7] <- anova.diversity$VarExp[4] * 100
```

1.1.8 Replicate 8

```
rep_8.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_8/rep_8.pi.50kb.bedgraph", header
rep_8.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_8/sim.tmrca.50k.map", header =

sim.lands.50kb <- as.data.frame(cbind(rep_8.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_8.tmr
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
## Warning in cor.test.default(x = mf[[1L]], y = mf[[2L]], ...): Cannot compute
## exact p-value with ties
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 37183069, p-value = 0.4216
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.03286589
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
## Warning in cor.test.default(x = mf[[1L]], y = mf[[2L]], ...): Cannot compute
## exact p-value with ties
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 37073930, p-value = 0.4656
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.02983425
```

```r
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)

sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)

# for merging:
sim.lands.50kb.rep8 <- sim.lands.50kb
sim.lands.50kb.rep8$Replicate <- 8

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 2.0323, p-value = 0.6384
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.48525, p-value = 0.289
```

```r
hist(resid(m.div.50kb))
```

```r
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.605e-03 -2.691e-04 -1.686e-05  2.620e-04  2.428e-03
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.057e-02  2.065e-05 996.284   <2e-16 ***
## thetaC         1.303e+00  2.468e-03 527.803   <2e-16 ***
## rhoC          -8.730e-04  6.970e-03  -0.125      0.9
## tmrcaC         2.380e-02  2.848e-04  83.583   <2e-16 ***
## thetaC:tmrcaC  1.383e+00  3.162e-02  43.722   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005051 on 595 degrees of freedom
## Multiple R-squared:  0.9979, Adjusted R-squared:  0.9979
## F-statistic: 7.078e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```r
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.sim.50kb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.sim.50kb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 8] <- anova.diversity$VarExp[4] * 100
```

1.1.9 Replicate 9

```
rep_9.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_9/rep_9.pi.50kb.bedgraph", header
rep_9.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_9/sim.tmrca.50k.map", header =

sim.lands.50kb <- as.data.frame(cbind(rep_9.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_9.tmr
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 35536739, p-value = 0.7531
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.01286561
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 37091756, p-value = 0.4583
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.03032942
```

```
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)
```

```r
sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)

# for merging:
sim.lands.50kb.rep9 <- sim.lands.50kb
sim.lands.50kb.rep9 $Replicate <- 9

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.9547, p-value = 0.2758
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.55455, p-value = 0.973
```

```r
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.591e-03 -2.409e-04 -2.982e-05  2.668e-04  1.822e-03
##
## Coefficients:
##                 Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    2.065e-02  1.899e-05 1087.271   <2e-16 ***
## thetaC         1.308e+00  2.265e-03  577.567   <2e-16 ***
## rhoC          -7.769e-03  6.412e-03   -1.212    0.226
## tmrcaC         2.404e-02  2.565e-04   93.731   <2e-16 ***
## thetaC:tmrcaC  1.508e+00  2.888e-02   52.235   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000465 on 595 degrees of freedom
## Multiple R-squared:  0.9983, Adjusted R-squared:  0.9983
## F-statistic: 8.777e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.50kb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```
r2.sim.50kb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 9] <- anova.diversity$VarExp[4] * 100
```

1.1.10 Replicate 10

```
rep_10.pi.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_10/rep_10.pi.50kb.bedgraph", hea
rep_10.tmrca.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_10/sim.tmrca.50k.map", header

sim.lands.50kb <- as.data.frame(cbind(rep_10.pi.50kb$pi, sim.theta.50kb$sim, sim.rho.50kb$sim, rep_10.tm
names(sim.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 37655834, p-value = 0.2606
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0459983
```

```
cor.test(~theta+rho, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 37660747, p-value = 0.2592
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.04613476
```

```
cor.test(~rho+tmrca, data = sim.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 36687846, p-value = 0.6403
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.01910966
```

```
# centering
sim.lands.50kb$thetaC <- sim.lands.50kb$theta - mean(sim.lands.50kb$theta)
sim.lands.50kb$tmrcaC <- sim.lands.50kb$tmrca - mean(sim.lands.50kb$tmrca)
sim.lands.50kb$rhoC <- sim.lands.50kb$rho - mean(sim.lands.50kb$rho)

sim.lands.50kb$bin <- 1:nrow(sim.lands.50kb)
```

```r
# for merging:
sim.lands.50kb.rep50 <- sim.lands.50kb
sim.lands.50kb.rep50$Replicate <- 10

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.50kb)
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.9775, p-value = 0.3738
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.43153, p-value = 0.006
```

```r
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.473e-03 -2.632e-04  1.203e-05  2.861e-04  1.648e-03
##
## Coefficients:
##                Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   2.070e-02  1.924e-05 1075.884   <2e-16 ***
## thetaC        1.311e+00  2.304e-03  568.969   <2e-16 ***
## rhoC          7.558e-03  6.493e-03    1.164    0.245
## tmrcaC        2.376e-02  2.376e-04  100.018   <2e-16 ***
## thetaC:tmrcaC 1.554e+00  2.642e-02   58.816   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004703 on 595 degrees of freedom
## Multiple R-squared:  0.9982, Adjusted R-squared:  0.9982
## F-statistic: 8.211e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.50kb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.50kb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.sim.50kb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.sim.50kb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.sim.50kb[5, 10] <- anova.diversity$VarExp[4] * 100

r2.sim.50kb$average <- rowMeans(r2.sim.50kb)
r2.sim.50kb <- transform(r2.sim.50kb, sd=apply(r2.sim.50kb, 1, sd, na.rm = TRUE))
```

1.1.11 all replicates:

```
sim.lands.50kb.all <- rbind(sim.lands.50kb.rep5, sim.lands.50kb.rep2, sim.lands.50kb.rep3, sim.lands.50k

m.div.50kb <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)*as.factor(Replicate), data = sim.l
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 2.0313, p-value = 0.8378
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.50138, p-value = 0.561
```

```r
hist(resid(m.div.50kb))
```

```r
m.div.50kb.2 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC)*as.factor(Replica
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC + rhoC:tmrcaC)*as.

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb   46 -74487.09
## m.div.50kb.2 55 -74476.67
## m.div.50kb.3 64 -74469.15
```

```r
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC) *
##     as.factor(Replicate), data = sim.lands.50kb.all)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0029659 -0.0002718 -0.0000139  0.0002767  0.0036466
##
## Coefficients:
##                                        Estimate Std. Error  t value Pr(>|t|)
## (Intercept)                           2.057e-02  1.983e-05 1037.362  < 2e-16
## thetaC                                1.299e+00  2.372e-03  547.744  < 2e-16
## rhoC                                 -1.009e-02  6.706e-03   -1.505 0.132298
## tmrcaC                                2.385e-02  2.743e-04   86.961  < 2e-16
## as.factor(Replicate)3                 9.708e-05  2.805e-05    3.461 0.000543
```

```
## as.factor(Replicate)4                -6.998e-05  2.804e-05   -2.495 0.012614
## as.factor(Replicate)5                 4.112e-05  2.429e-05    1.693 0.090555
## as.factor(Replicate)6                 8.307e-06  2.804e-05    0.296 0.767075
## as.factor(Replicate)7                 3.621e-05  2.806e-05    1.290 0.197011
## as.factor(Replicate)8                -2.113e-06  2.806e-05   -0.075 0.939967
## as.factor(Replicate)9                 7.290e-05  2.805e-05    2.599 0.009365
## as.factor(Replicate)10                1.253e-04  2.807e-05    4.465 8.14e-06
## thetaC:tmrcaC                         1.515e+00  3.122e-02   48.515  < 2e-16
## thetaC:as.factor(Replicate)3          1.091e-02  3.351e-03    3.254 0.001143
## thetaC:as.factor(Replicate)4          9.175e-04  3.349e-03    0.274 0.784112
## thetaC:as.factor(Replicate)5          8.367e-03  2.904e-03    2.882 0.003971
## thetaC:as.factor(Replicate)6          4.459e-03  3.354e-03    1.329 0.183796
## thetaC:as.factor(Replicate)7          1.356e-02  3.352e-03    4.045 5.29e-05
## thetaC:as.factor(Replicate)8          3.160e-03  3.355e-03    0.942 0.346264
## thetaC:as.factor(Replicate)9          8.544e-03  3.350e-03    2.551 0.010780
## thetaC:as.factor(Replicate)10         1.150e-02  3.360e-03    3.422 0.000626
## rhoC:as.factor(Replicate)3            1.654e-02  9.480e-03    1.744 0.081155
## rhoC:as.factor(Replicate)4            2.396e-02  9.476e-03    2.528 0.011491
## rhoC:as.factor(Replicate)5            1.162e-02  8.210e-03    1.415 0.156986
## rhoC:as.factor(Replicate)6           -1.695e-03  9.479e-03   -0.179 0.858054
## rhoC:as.factor(Replicate)7            4.341e-03  9.475e-03    0.458 0.646844
## rhoC:as.factor(Replicate)8            9.221e-03  9.479e-03    0.973 0.330721
## rhoC:as.factor(Replicate)9            2.325e-03  9.476e-03    0.245 0.806184
## rhoC:as.factor(Replicate)10           1.765e-02  9.482e-03    1.862 0.062717
## tmrcaC:as.factor(Replicate)3         -2.443e-04  3.669e-04   -0.666 0.505586
## tmrcaC:as.factor(Replicate)4         -3.682e-04  3.727e-04   -0.988 0.323169
## tmrcaC:as.factor(Replicate)5         -1.113e-04  3.340e-04   -0.333 0.738997
## tmrcaC:as.factor(Replicate)6         -2.224e-04  3.802e-04   -0.585 0.558571
## tmrcaC:as.factor(Replicate)7          3.626e-05  3.779e-04    0.096 0.923556
## tmrcaC:as.factor(Replicate)8         -5.299e-05  3.875e-04   -0.137 0.891250
## tmrcaC:as.factor(Replicate)9          1.871e-04  3.834e-04    0.488 0.625634
## tmrcaC:as.factor(Replicate)10        -9.171e-05  3.680e-04   -0.249 0.803197
## thetaC:tmrcaC:as.factor(Replicate)3  -1.346e-02  4.128e-02   -0.326 0.744413
## thetaC:tmrcaC:as.factor(Replicate)4  -2.192e-02  4.324e-02   -0.507 0.612153
## thetaC:tmrcaC:as.factor(Replicate)5  -4.384e-02  3.836e-02   -1.143 0.253171
## thetaC:tmrcaC:as.factor(Replicate)6  -7.532e-02  4.355e-02   -1.729 0.083798
## thetaC:tmrcaC:as.factor(Replicate)7  -2.354e-02  4.386e-02   -0.537 0.591443
## thetaC:tmrcaC:as.factor(Replicate)8  -1.321e-01  4.358e-02   -3.032 0.002438
## thetaC:tmrcaC:as.factor(Replicate)9  -6.256e-03  4.341e-02   -0.144 0.885413
## thetaC:tmrcaC:as.factor(Replicate)10  3.941e-02  4.146e-02    0.950 0.341925
##
## (Intercept)                          ***
## thetaC                               ***
## rhoC
## tmrcaC                               ***
## as.factor(Replicate)3                ***
## as.factor(Replicate)4                *
## as.factor(Replicate)5                .
## as.factor(Replicate)6
## as.factor(Replicate)7
## as.factor(Replicate)8
## as.factor(Replicate)9                **
## as.factor(Replicate)10               ***
## thetaC:tmrcaC                        ***
```

```
## thetaC:as.factor(Replicate)3         **
## thetaC:as.factor(Replicate)4
## thetaC:as.factor(Replicate)5         **
## thetaC:as.factor(Replicate)6
## thetaC:as.factor(Replicate)7         ***
## thetaC:as.factor(Replicate)8
## thetaC:as.factor(Replicate)9         *
## thetaC:as.factor(Replicate)10        ***
## rhoC:as.factor(Replicate)3           .
## rhoC:as.factor(Replicate)4           *
## rhoC:as.factor(Replicate)5
## rhoC:as.factor(Replicate)6
## rhoC:as.factor(Replicate)7
## rhoC:as.factor(Replicate)8
## rhoC:as.factor(Replicate)9
## rhoC:as.factor(Replicate)10          .
## tmrcaC:as.factor(Replicate)3
## tmrcaC:as.factor(Replicate)4
## tmrcaC:as.factor(Replicate)5
## tmrcaC:as.factor(Replicate)6
## tmrcaC:as.factor(Replicate)7
## tmrcaC:as.factor(Replicate)8
## tmrcaC:as.factor(Replicate)9
## tmrcaC:as.factor(Replicate)10
## thetaC:tmrcaC:as.factor(Replicate)3
## thetaC:tmrcaC:as.factor(Replicate)4
## thetaC:tmrcaC:as.factor(Replicate)5
## thetaC:tmrcaC:as.factor(Replicate)6  .
## thetaC:tmrcaC:as.factor(Replicate)7
## thetaC:tmrcaC:as.factor(Replicate)8  **
## thetaC:tmrcaC:as.factor(Replicate)9
## thetaC:tmrcaC:as.factor(Replicate)10
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004856 on 5955 degrees of freedom
## Multiple R-squared:  0.9981, Adjusted R-squared:  0.9981
## F-statistic: 7.12e+04 on 44 and 5955 DF,  p-value: < 2.2e-16
```

1.2 200 kb scale

```
r2.sim.200kb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.sim.200kb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.sim.200kb) <- reps

sim.rho.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/dm.sim.rho.2e+05.txt", header = T)
sim.theta.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/dm.sim.theta.2e+05.txt", header = T
```

1.2.1 Replicate 1

```
rep_1.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_1/rep_1.pi.200kb.bedgraph", head
rep_1.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_1/sim.tmrca.200k.map", header
```

```r
sim.lands.200kb <- as.data.frame(cbind(rep_1.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_1
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 569023, p-value = 0.8876
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.01164184
```

```r
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```r
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 570624, p-value = 0.8602
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.01448776
```

```r
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)

# for merging:
sim.lands.200kb.rep5 <- sim.lands.200kb
sim.lands.200kb.rep5$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.9751, p-value = 0.4234
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.49982, p-value = 0.503
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.306e-03 -1.718e-04  1.801e-05  1.985e-04  9.793e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.067e-02  2.788e-05 741.370   <2e-16 ***
## thetaC        1.306e+00  3.769e-03 346.532   <2e-16 ***
## rhoC          3.201e-02  1.870e-02   1.712    0.089 .
## tmrcaC        2.526e-02  7.654e-04  32.997   <2e-16 ***
## thetaC:tmrcaC 1.536e+00  1.072e-01  14.325   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003415 on 145 degrees of freedom
## Multiple R-squared:  0.9988, Adjusted R-squared:  0.9988
## F-statistic: 3.049e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 1] <- anova.diversity$VarExp[4] * 100
```

1.2.2 Replicate 2

```
rep_2.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_2/rep_2.pi.200kb.bedgraph", hea
rep_2.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_2/sim.tmrca.200k.map", heade

sim.lands.200kb <- as.data.frame(cbind(rep_2.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_2
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 551432, p-value = 0.8115
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.01963361
```

```
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 533372, p-value = 0.529
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.05174097
```

```
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)
```

```r
# for merging:
sim.lands.200kb.rep2 <- sim.lands.200kb
sim.lands.200kb.rep2$Replicate <- 2

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.5928, p-value = 0.005174
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.59119, p-value = 0.933
```

```r
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.096e-03 -1.706e-04 -2.456e-05  1.813e-04  1.306e-03
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.056e-02  2.494e-05  824.14   <2e-16 ***
## thetaC         1.297e+00  3.394e-03  382.15   <2e-16 ***
## rhoC          -3.118e-02  1.667e-02   -1.87   0.0635 .
## tmrcaC         2.455e-02  6.157e-04   39.88   <2e-16 ***
## thetaC:tmrcaC  1.643e+00  8.199e-02   20.04   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003054 on 145 degrees of freedom
## Multiple R-squared:  0.999,  Adjusted R-squared:  0.999
## F-statistic: 3.71e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 2] <- anova.diversity$VarExp[4] * 100
```

1.2.3 Replicate 3

```
rep_3.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_3/rep_3.pi.200kb.bedgraph", hea
rep_3.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_3/sim.tmrca.200k.map", heade

sim.lands.200kb <- as.data.frame(cbind(rep_3.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_3
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 592174, p-value = 0.5211
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.05280076
```

```
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 594376, p-value = 0.4902
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.05671541
```

```
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)
```

```r
# for merging:
sim.lands.200kb.rep3 <- sim.lands.200kb
sim.lands.200kb.rep3$Replicate <- 3

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.9205, p-value = 0.2965
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.45845, p-value = 0.231
```

```r
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.928e-03 -1.460e-04  3.044e-05  2.110e-04  2.099e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.064e-02  3.245e-05 636.263   <2e-16 ***
## thetaC        1.309e+00  4.394e-03 297.825   <2e-16 ***
## rhoC          3.167e-02  2.173e-02   1.458    0.147
## tmrcaC        2.488e-02  7.859e-04  31.662   <2e-16 ***
## thetaC:tmrcaC 1.472e+00  1.097e-01  13.422   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003972 on 145 degrees of freedom
## Multiple R-squared:  0.9984, Adjusted R-squared:  0.9984
## F-statistic: 2.269e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 3] <- anova.diversity$VarExp[4] * 100
```

1.2.4 Replicate 4

```
rep_4.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_4/rep_4.pi.200kb.bedgraph", hea
rep_4.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_4/sim.tmrca.200k.map", heade

sim.lands.200kb <- as.data.frame(cbind(rep_4.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_4
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 554491, p-value = 0.8631
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.01419494
```

```
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 559798, p-value = 0.9539
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## 0.004759323
```

```
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)
```

```r
# for merging:
sim.lands.200kb.rep4 <- sim.lands.200kb
sim.lands.200kb.rep4$Replicate <- 4

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.9887, p-value = 0.4477
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.4058, p-value = 0.046
```

```r
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.257e-03 -1.693e-04  3.530e-06  1.847e-04  1.055e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0205018  0.0000271 756.559   <2e-16 ***
## thetaC        1.3010756  0.0036626 355.236   <2e-16 ***
## rhoC          0.0087841  0.0181026   0.485    0.628
## tmrcaC        0.0244436  0.0006814  35.873   <2e-16 ***
## thetaC:tmrcaC 1.4688735  0.0873643  16.813   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003317 on 145 degrees of freedom
## Multiple R-squared:  0.9989, Adjusted R-squared:  0.9989
## F-statistic: 3.248e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 4] <- anova.diversity$VarExp[4] * 100
```

1.2.5 Replicate 5

```
rep_5.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_5/rep_5.pi.200kb.bedgraph", hea
rep_5.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_5/sim.tmrca.200k.map", heade

sim.lands.200kb <- as.data.frame(cbind(rep_5.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_5
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 570758, p-value = 0.858
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.01472654
```

```
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 582370, p-value = 0.6671
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.03537046
```

```
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)
```

```r
# for merging:
sim.lands.200kb.rep5 <- sim.lands.200kb
sim.lands.200kb.rep5$Replicate <- 5

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.9065, p-value = 0.2691
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.42468, p-value = 0.103
```

```r
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -7.823e-04 -1.890e-04  3.370e-06  1.582e-04  7.148e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.061e-02  2.367e-05 870.811   <2e-16 ***
## thetaC        1.306e+00  3.205e-03 407.532   <2e-16 ***
## rhoC          8.924e-03  1.586e-02   0.563    0.575
## tmrcaC        2.402e-02  6.327e-04  37.957   <2e-16 ***
## thetaC:tmrcaC 1.380e+00  7.367e-02  18.728   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0002897 on 145 degrees of freedom
## Multiple R-squared:  0.9991, Adjusted R-squared:  0.9991
## F-statistic: 4.197e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 5] <- anova.diversity$VarExp[4] * 100
```

1.2.6 Replicate 6

```
rep_6.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_6/rep_6.pi.200kb.bedgraph", hea
rep_6.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_6/sim.tmrca.200k.map", heade

sim.lands.200kb <- as.data.frame(cbind(rep_6.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_6
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 577585, p-value = 0.7442
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.02686264
```

```
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 606554, p-value = 0.3401
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.07836615
```

```
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)
```

```r
# for merging:
sim.lands.200kb.rep6 <- sim.lands.200kb
sim.lands.200kb.rep6$Replicate <- 6

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 2.0877, p-value = 0.6996
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.59616, p-value = 0.964
```

```r
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.463e-03 -1.507e-04  2.405e-05  1.519e-04  7.279e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.055e-02  2.337e-05 879.460   <2e-16 ***
## thetaC        1.299e+00  3.288e-03 395.145   <2e-16 ***
## rhoC          3.569e-03  1.559e-02   0.229    0.819
## tmrcaC        2.392e-02  6.333e-04  37.771   <2e-16 ***
## thetaC:tmrcaC 1.533e+00  6.794e-02  22.563   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000286 on 145 degrees of freedom
## Multiple R-squared:  0.9992, Adjusted R-squared:  0.9992
## F-statistic: 4.499e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 6] <- anova.diversity$VarExp[4] * 100
```

1.2.7 Replicate 7

```
rep_7.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_7/rep_7.pi.200kb.bedgraph", hea
rep_7.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_7/sim.tmrca.200k.map", heade

sim.lands.200kb <- as.data.frame(cbind(rep_7.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_7
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 647408, p-value = 0.06511
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.150999
```

```
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 523746, p-value = 0.402
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.06885462
```

```
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)
```

```r
# for merging:
sim.lands.200kb.rep7 <- sim.lands.200kb
sim.lands.200kb.rep7$Replicate <- 7

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.9801, p-value = 0.4361
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.40495, p-value = 0.044
```

```r
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.220e-03 -1.868e-04 -1.610e-06  1.848e-04  1.913e-03
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.062e-02  3.141e-05 656.642   <2e-16 ***
## thetaC         1.315e+00  4.235e-03 310.525   <2e-16 ***
## rhoC          -1.620e-02  2.076e-02  -0.781    0.436
## tmrcaC         2.369e-02  8.316e-04  28.486   <2e-16 ***
## thetaC:tmrcaC  1.422e+00  1.203e-01  11.823   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003796 on 145 degrees of freedom
## Multiple R-squared:  0.9985, Adjusted R-squared:  0.9985
## F-statistic: 2.432e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 7] <- anova.diversity$VarExp[4] * 100
```

1.2.8 Replicate 8

```
rep_8.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_8/rep_8.pi.200kb.bedgraph", hea
rep_8.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_8/sim.tmrca.200k.map", heade

sim.lands.200kb <- as.data.frame(cbind(rep_8.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_8
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 588671, p-value = 0.5714
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0465727
```

```
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 624456, p-value = 0.1793
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.1101933
```

```
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)
```

```r
# for merging:
sim.lands.200kb.rep8 <- sim.lands.200kb
sim.lands.200kb.rep8$Replicate <- 8

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 2.036, p-value = 0.577
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.54355, p-value = 0.784
```

```r
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0010021 -0.0001570 -0.0000008  0.0001486  0.0009608
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.057e-02  2.534e-05 811.819   <2e-16 ***
## thetaC         1.309e+00  3.478e-03 376.496   <2e-16 ***
## rhoC          -9.189e-03  1.709e-02  -0.538    0.592
## tmrcaC         2.531e-02  6.716e-04  37.681   <2e-16 ***
## thetaC:tmrcaC  1.518e+00  7.926e-02  19.154   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000309 on 145 degrees of freedom
## Multiple R-squared:  0.999,  Adjusted R-squared:  0.999
## F-statistic: 3.616e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 8] <- anova.diversity$VarExp[4] * 100
```

1.2.9 Replicate 9

```
rep_9.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_9/rep_9.pi.200kb.bedgraph", hea
rep_9.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_9/sim.tmrca.200k.map", heade

sim.lands.200kb <- as.data.frame(cbind(rep_9.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_9
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 493800, p-value = 0.1366
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.1220935
```

```
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 597682, p-value = 0.4463
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.062593
```

```
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)
```

```r
# for merging:
sim.lands.200kb.rep9 <- sim.lands.200kb
sim.lands.200kb.rep9$Replicate <- 9

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.6818, p-value = 0.022
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.42403, p-value = 0.084
```

```r
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.180e-03 -1.904e-04  1.650e-06  2.078e-04  1.429e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.061e-02  2.834e-05 727.360   <2e-16 ***
## thetaC        1.308e+00  3.854e-03 339.390   <2e-16 ***
## rhoC          6.075e-03  1.879e-02   0.323    0.747
## tmrcaC        2.558e-02  8.061e-04  31.733   <2e-16 ***
## thetaC:tmrcaC 1.713e+00  1.105e-01  15.505   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003441 on 145 degrees of freedom
## Multiple R-squared:  0.9988, Adjusted R-squared:  0.9988
## F-statistic: 3.126e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 9] <- anova.diversity$VarExp[4] * 100
```

1.2.10 Replicate 10

```
rep_10.pi.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_10/rep_10.pi.200kb.bedgraph", h
rep_10.tmrca.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_10/sim.tmrca.200k.map", hea

sim.lands.200kb <- as.data.frame(cbind(rep_10.pi.200kb$pi, sim.theta.200kb$sim, sim.rho.200kb$sim, rep_
names(sim.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 605678, p-value = 0.3502
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.07680806
```

```
cor.test(~theta+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 593676, p-value = 0.5002
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0554712
```

```
cor.test(~rho+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 612862, p-value = 0.2753
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.08958087
```

```
# centering
sim.lands.200kb$thetaC <- sim.lands.200kb$theta - mean(sim.lands.200kb$theta)
sim.lands.200kb$tmrcaC <- sim.lands.200kb$tmrca - mean(sim.lands.200kb$tmrca)
sim.lands.200kb$rhoC <- sim.lands.200kb$rho - mean(sim.lands.200kb$rho)

sim.lands.200kb$bin <- 1:nrow(sim.lands.200kb)
```

```r
# for merging:
sim.lands.200kb.rep50 <- sim.lands.200kb
sim.lands.200kb.rep50$Replicate <- 10

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.200kb)
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 2.1646, p-value = 0.8354
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.41358, p-value = 0.065
```

```r
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.220e-03 -1.345e-04  6.710e-06  1.375e-04  1.100e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.071e-02  2.877e-05 720.015   <2e-16 ***
## thetaC        1.317e+00  4.037e-03 326.195   <2e-16 ***
## rhoC          4.763e-03  1.928e-02   0.247    0.805
## tmrcaC        2.484e-02  6.682e-04  37.182   <2e-16 ***
## thetaC:tmrcaC 1.796e+00  7.919e-02  22.681   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003493 on 145 degrees of freedom
## Multiple R-squared:  0.9987, Adjusted R-squared:  0.9987
## F-statistic: 2.783e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.200kb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.sim.200kb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.sim.200kb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.sim.200kb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.sim.200kb[5, 10] <- anova.diversity$VarExp[4] * 100

r2.sim.200kb$average <- rowMeans(r2.sim.200kb)
r2.sim.200kb <- transform(r2.sim.200kb, sd=apply(r2.sim.200kb, 1, sd, na.rm = TRUE))
```

1.2.11 all replicates:

```
sim.lands.200kb.all <- rbind(sim.lands.200kb.rep5, sim.lands.200kb.rep2, sim.lands.200kb.rep3, sim.lands

m.div.200kb <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)*as.factor(Replicate), data = sim
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.9357, p-value = 0.05828
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.48495, p-value = 0.194
```

```r
hist(resid(m.div.200kb))
```

```r
m.div.200kb.2 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC)*as.factor(Replica
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC + rhoC:tmrcaC)*as

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb   46 -19748.95
## m.div.200kb.2 55 -19741.46
## m.div.200kb.3 64 -19728.72
```

```r
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC) *
##     as.factor(Replicate), data = sim.lands.200kb.all)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.220e-03 -1.705e-04  1.390e-06  1.741e-04  2.099e-03
##
## Coefficients:
##                                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)                           2.056e-02  2.693e-05 763.480  < 2e-16 ***
## thetaC                                1.297e+00  3.664e-03 354.022  < 2e-16 ***
## rhoC                                 -3.118e-02  1.800e-02  -1.733 0.083394 .
## tmrcaC                                2.455e-02  6.646e-04  36.941  < 2e-16 ***
## as.factor(Replicate)3                 8.648e-05  3.808e-05   2.271 0.023308 *
```

```
## as.factor(Replicate)4                -5.611e-05  3.808e-05  -1.473 0.140907
## as.factor(Replicate)5                 5.158e-05  3.298e-05   1.564 0.118060
## as.factor(Replicate)6                -3.139e-06  3.809e-05  -0.082 0.934334
## as.factor(Replicate)7                 6.595e-05  3.833e-05   1.721 0.085525 .
## as.factor(Replicate)8                 1.320e-05  3.816e-05   0.346 0.729361
## as.factor(Replicate)9                 5.572e-05  3.824e-05   1.457 0.145329
## as.factor(Replicate)10                1.554e-04  3.824e-05   4.064 5.09e-05 ***
## thetaC:tmrcaC                         1.643e+00  8.851e-02  18.565  < 2e-16 ***
## thetaC:as.factor(Replicate)3          1.170e-02  5.170e-03   2.263 0.023812 *
## thetaC:as.factor(Replicate)4          4.031e-03  5.165e-03   0.780 0.435225
## thetaC:as.factor(Replicate)5          9.218e-03  4.481e-03   2.057 0.039834 *
## thetaC:as.factor(Replicate)6          2.241e-03  5.272e-03   0.425 0.670855
## thetaC:as.factor(Replicate)7          1.818e-02  5.192e-03   3.502 0.000476 ***
## thetaC:as.factor(Replicate)8          1.225e-02  5.214e-03   2.349 0.018958 *
## thetaC:as.factor(Replicate)9          1.109e-02  5.202e-03   2.131 0.033240 *
## thetaC:as.factor(Replicate)10         1.971e-02  5.286e-03   3.728 0.000200 ***
## rhoC:as.factor(Replicate)3            6.285e-02  2.548e-02   2.467 0.013739 *
## rhoC:as.factor(Replicate)4            3.996e-02  2.545e-02   1.570 0.116550
## rhoC:as.factor(Replicate)5            4.010e-02  2.206e-02   1.818 0.069319 .
## rhoC:as.factor(Replicate)6            3.475e-02  2.543e-02   1.366 0.172044
## rhoC:as.factor(Replicate)7            1.498e-02  2.547e-02   0.588 0.556651
## rhoC:as.factor(Replicate)8            2.199e-02  2.562e-02   0.858 0.390926
## rhoC:as.factor(Replicate)9            3.725e-02  2.546e-02   1.463 0.143599
## rhoC:as.factor(Replicate)10           3.594e-02  2.559e-02   1.404 0.160426
## tmrcaC:as.factor(Replicate)3          3.319e-04  9.312e-04   0.356 0.721606
## tmrcaC:as.factor(Replicate)4         -1.072e-04  9.489e-04  -0.113 0.910045
## tmrcaC:as.factor(Replicate)5         -5.334e-04  8.373e-04  -0.637 0.524142
## tmrcaC:as.factor(Replicate)6         -6.326e-04  9.872e-04  -0.641 0.521738
## tmrcaC:as.factor(Replicate)7         -8.605e-04  9.816e-04  -0.877 0.380793
## tmrcaC:as.factor(Replicate)8          7.566e-04  9.773e-04   0.774 0.438985
## tmrcaC:as.factor(Replicate)9          1.030e-03  1.019e-03   1.011 0.312387
## tmrcaC:as.factor(Replicate)10         2.934e-04  9.162e-04   0.320 0.748831
## thetaC:tmrcaC:as.factor(Replicate)3  -1.709e-01  1.270e-01  -1.346 0.178522
## thetaC:tmrcaC:as.factor(Replicate)4  -1.743e-01  1.240e-01  -1.406 0.160029
## thetaC:tmrcaC:as.factor(Replicate)5  -2.634e-01  1.065e-01  -2.473 0.013515 *
## thetaC:tmrcaC:as.factor(Replicate)6  -1.102e-01  1.182e-01  -0.933 0.351213
## thetaC:tmrcaC:as.factor(Replicate)7  -2.207e-01  1.369e-01  -1.612 0.107174
## thetaC:tmrcaC:as.factor(Replicate)8  -1.250e-01  1.224e-01  -1.021 0.307456
## thetaC:tmrcaC:as.factor(Replicate)9   6.939e-02  1.380e-01   0.503 0.615071
## thetaC:tmrcaC:as.factor(Replicate)10  1.530e-01  1.158e-01   1.321 0.186722
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003297 on 1455 degrees of freedom
## Multiple R-squared:  0.9989, Adjusted R-squared:  0.9989
## F-statistic: 2.96e+04 on 44 and 1455 DF,  p-value: < 2.2e-16
```

1.3 1 Mb scale

```
r2.sim.1Mb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.sim.1Mb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.sim.1Mb) <- reps
```

```r
sim.rho.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/dm.sim.rho.1e+06.txt", header = T)
sim.theta.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/dm.sim.theta.1e+06.txt", header = T)
```

1.3.1 Replicate 1

```r
rep_1.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_1/rep_1.pi.1Mb.bedgraph", header =
rep_1.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_1/sim.tmrca.1M.map", header = T

sim.lands.1Mb <- as.data.frame(cbind(rep_1.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_1.tmrca.1M
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 5134, p-value = 0.452
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.142158
```

```r
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```r
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 4576, p-value = 0.9251
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.01802002
```

```r
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)

# for merging:
```

```r
sim.lands.1Mb.rep5 <- sim.lands.1Mb
sim.lands.1Mb.rep5$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.5295, p-value = 0.09374
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.59412, p-value = 0.751
```

```r
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0005582 -0.0001279  0.0000482  0.0001519  0.0003836
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.067e-02  4.132e-05 500.181  < 2e-16 ***
## thetaC         1.313e+00  9.626e-03 136.418  < 2e-16 ***
## rhoC          -2.472e-03  6.889e-02  -0.036   0.9717
## tmrcaC         2.602e-02  2.881e-03   9.031 2.41e-09 ***
## thetaC:tmrcaC  1.182e+00  6.258e-01   1.889   0.0706 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0002263 on 25 degrees of freedom
## Multiple R-squared:  0.9989, Adjusted R-squared:  0.9988
## F-statistic:  5910 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.sim.1Mb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 1] <- anova.diversity$VarExp[4] * 100
```

1.3.2 Replicate 2

```
rep_2.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_2/rep_2.pi.1Mb.bedgraph", header =
rep_2.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_2/sim.tmrca.1M.map", header = T

sim.lands.1Mb <- as.data.frame(cbind(rep_2.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_2.tmrca.1M
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 3982, p-value = 0.5467
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.1141268
```

```
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 3726, p-value = 0.3645
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.171079
```

```
cor.test(~diversity+theta, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  diversity and theta
```

```
## S = 4053.2, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9927941
```

```
cor.test(~diversity+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  diversity and tmrca
## S = 552924, p-value = 0.8364
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.01698031
```

```
cor.test(~diversity+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  diversity and rho
## S = 596694, p-value = 0.4591
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.06083648
```

```
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)

# for merging:
sim.lands.1Mb.rep2 <- sim.lands.1Mb
sim.lands.1Mb.rep2$Replicate <- 2

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.6531, p-value = 0.9631
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.51119, p-value = 0.514
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.627e-04 -1.567e-04 -1.856e-05  2.165e-04  3.726e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.052e-02  4.606e-05 445.437  < 2e-16 ***
## thetaC         1.285e+00  1.013e-02 126.882  < 2e-16 ***
## rhoC          -4.630e-02  7.505e-02  -0.617 0.542900
## tmrcaC         2.480e-02  2.820e-03   8.796 3.99e-09 ***
## thetaC:tmrcaC  2.351e+00  5.373e-01   4.376 0.000188 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0002496 on 25 degrees of freedom
## Multiple R-squared:  0.9987, Adjusted R-squared:  0.9985
## F-statistic:  4800 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.sim.1Mb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 2] <- anova.diversity$VarExp[4] * 100
```

1.3.3 Replicate 3

```
rep_3.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_3/rep_3.pi.1Mb.bedgraph", header =
rep_3.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_3/sim.tmrca.1M.map", header = 

sim.lands.1Mb <- as.data.frame(cbind(rep_3.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_3.tmrca.1
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 6456, p-value = 0.01674
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.4362625
```

```
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 4500, p-value = 0.9963
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##          rho
## -0.001112347
```

```
cor.test(~diversity+theta, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  diversity and theta
```

```
## S = 4053.2, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9927941
```

```
cor.test(~diversity+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  diversity and tmrca
## S = 552924, p-value = 0.8364
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.01698031
```

```
cor.test(~diversity+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  diversity and rho
## S = 596694, p-value = 0.4591
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.06083648
```

```
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)

# for merging:
sim.lands.1Mb.rep3 <- sim.lands.1Mb
sim.lands.1Mb.rep3$Replicate <- 3

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.3611, p-value = 0.8304
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.69982, p-value = 0.933
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.409e-04 -9.496e-05 -1.057e-05  1.246e-04  2.682e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.070e-02  3.559e-05 581.752  < 2e-16 ***
## thetaC        1.326e+00  7.629e-03 173.822  < 2e-16 ***
## rhoC          6.225e-02  5.347e-02   1.164    0.255
## tmrcaC        3.071e-02  2.054e-03  14.949 5.67e-14 ***
## thetaC:tmrcaC 2.593e+00  5.197e-01   4.990 3.83e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001781 on 25 degrees of freedom
## Multiple R-squared:  0.9993, Adjusted R-squared:  0.9992
## F-statistic:  9011 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.sim.1Mb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 3] <- anova.diversity$VarExp[4] * 100
```

1.3.4 Replicate 4

```
rep_4.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_4/rep_4.pi.1Mb.bedgraph", header =
rep_4.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_4/sim.tmrca.1M.map", header = 

sim.lands.1Mb <- as.data.frame(cbind(rep_4.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_4.tmrca.1
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 4678, p-value = 0.8307
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.0407119
```

```
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 3482, p-value = 0.2302
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.2253615
```

```
cor.test(~diversity+theta, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  diversity and theta
```

```
## S = 4053.2, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9927941
```

```r
cor.test(~diversity+tmrca, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  diversity and tmrca
## S = 552924, p-value = 0.8364
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.01698031
```

```r
cor.test(~diversity+rho, data = sim.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  diversity and rho
## S = 596694, p-value = 0.4591
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.06083648
```

```r
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)

# for merging:
sim.lands.1Mb.rep4 <- sim.lands.1Mb
sim.lands.1Mb.rep4$Replicate <- 4

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.4044, p-value = 0.8773
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.39065, p-value = 0.19
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -5.763e-04 -6.567e-05  2.077e-05  1.210e-04  2.477e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.049e-02  3.521e-05 582.000  < 2e-16 ***
## thetaC        1.290e+00  7.811e-03 165.192  < 2e-16 ***
## rhoC          1.009e-02  5.807e-02   0.174    0.863
## tmrcaC        2.364e-02  1.989e-03  11.884 8.84e-12 ***
## thetaC:tmrcaC 2.028e+00  3.007e-01   6.744 4.55e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001916 on 25 degrees of freedom
## Multiple R-squared:  0.9992, Adjusted R-squared:  0.9991
## F-statistic:  8277 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.sim.1Mb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 4] <- anova.diversity$VarExp[4] * 100
```

1.3.5 Replicate 5

```
rep_5.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_5/rep_5.pi.1Mb.bedgraph", header =
rep_5.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_5/sim.tmrca.1M.map", header = 

sim.lands.1Mb <- as.data.frame(cbind(rep_5.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_5.tmrca.1
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 4438, p-value = 0.9476
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.01268076
```

```
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 4306, p-value = 0.8252
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.04204672
```

```
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)
```

```r
# for merging:
sim.lands.1Mb.rep5 <- sim.lands.1Mb
sim.lands.1Mb.rep5$Replicate <- 5

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.9571, p-value = 0.43
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.35533, p-value = 0.112
```

```r
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -5.859e-04 -7.379e-05 -6.890e-06  1.260e-04  3.018e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.060e-02  3.534e-05 582.791  < 2e-16 ***
## thetaC         1.306e+00  7.849e-03 166.389  < 2e-16 ***
## rhoC          -3.495e-02  5.720e-02  -0.611   0.5467
## tmrcaC         2.514e-02  2.157e-03  11.655 1.34e-11 ***
## thetaC:tmrcaC  1.135e+00  5.053e-01   2.246   0.0338 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001933 on 25 degrees of freedom
## Multiple R-squared:  0.9992, Adjusted R-squared:  0.9991
## F-statistic:  7809 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.sim.1Mb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 5] <- anova.diversity$VarExp[4] * 100
```

1.3.6 Replicate 6

```
rep_6.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_6/rep_6.pi.1Mb.bedgraph", header =
rep_6.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_6/sim.tmrca.1M.map", header = 

sim.lands.1Mb <- as.data.frame(cbind(rep_6.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_6.tmrca.1
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 3592, p-value = 0.2858
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.2008899
```

```
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 4268, p-value = 0.7906
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.05050056
```

```
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)
```

```r
# for merging:
sim.lands.1Mb.rep6 <- sim.lands.1Mb
sim.lands.1Mb.rep6$Replicate <- 6

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.956, p-value = 0.4515
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.62113, p-value = 0.825
```

```r
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.104e-04 -1.172e-04  4.625e-05  1.282e-04  3.696e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.055e-02  3.357e-05 612.052  < 2e-16 ***
## thetaC        1.305e+00  7.159e-03 182.260  < 2e-16 ***
## rhoC          6.955e-02  5.357e-02   1.298  0.20607
## tmrcaC        2.808e-02  2.754e-03  10.193 2.18e-10 ***
## thetaC:tmrcaC 2.260e+00  6.739e-01   3.353  0.00255 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000178 on 25 degrees of freedom
## Multiple R-squared:  0.9993, Adjusted R-squared:  0.9992
## F-statistic:  9504 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.sim.1Mb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 6] <- anova.diversity$VarExp[4] * 100
```

1.3.7 Replicate 7

```
rep_7.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_7/rep_7.pi.1Mb.bedgraph", header =
rep_7.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_7/sim.tmrca.1M.map", header = 

sim.lands.1Mb <- as.data.frame(cbind(rep_7.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_7.tmrca.1
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 5750, p-value = 0.135
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.2791991
```

```
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 3620, p-value = 0.3013
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.1946607
```

```
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)
```

```r
# for merging:
sim.lands.1Mb.rep7 <- sim.lands.1Mb
sim.lands.1Mb.rep7$Replicate <- 7

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.9852, p-value = 0.4718
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.65923, p-value = 0.888
```

```r
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.677e-04 -1.175e-04 -5.270e-06  9.488e-05  3.785e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.060e-02  3.433e-05 599.990  < 2e-16 ***
## thetaC         1.322e+00  7.722e-03 171.218  < 2e-16 ***
## rhoC          -5.037e-02  5.447e-02  -0.925 0.363900
## tmrcaC         2.856e-02  1.841e-03  15.507 2.47e-14 ***
## thetaC:tmrcaC  1.507e+00  3.301e-01   4.565 0.000115 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001836 on 25 degrees of freedom
## Multiple R-squared:  0.9993, Adjusted R-squared:  0.9991
## F-statistic:  8510 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.sim.1Mb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 7] <- anova.diversity$VarExp[4] * 100
```

1.3.8 Replicate 8

```
rep_8.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_8/rep_8.pi.1Mb.bedgraph", header =
rep_8.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_8/sim.tmrca.1M.map", header = T

sim.lands.1Mb <- as.data.frame(cbind(rep_8.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_8.tmrca.1M
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 4082, p-value = 0.628
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.09187987
```

```
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 4888, p-value = 0.6448
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##         rho
## -0.08743048
```

```
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)
```

```r
# for merging:
sim.lands.1Mb.rep8 <- sim.lands.1Mb
sim.lands.1Mb.rep8$Replicate <- 8

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.4886, p-value = 0.9089
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.6984, p-value = 0.942
```

```r
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -6.061e-04 -1.034e-04  2.376e-05  1.402e-04  3.024e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.051e-02  3.685e-05 556.403  < 2e-16 ***
## thetaC        1.294e+00  7.915e-03 163.527  < 2e-16 ***
## rhoC          2.309e-02  5.876e-02   0.393  0.69767
## tmrcaC        2.659e-02  2.276e-03  11.682 1.27e-11 ***
## thetaC:tmrcaC 2.000e+00  5.608e-01   3.567  0.00149 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001996 on 25 degrees of freedom
## Multiple R-squared:  0.9992, Adjusted R-squared:  0.999
## F-statistic:  7462 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.sim.1Mb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 8] <- anova.diversity$VarExp[4] * 100
```

1.3.9 Replicate 9

```
rep_9.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_9/rep_9.pi.1Mb.bedgraph", header =
rep_9.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_9/sim.tmrca.1M.map", header = T

sim.lands.1Mb <- as.data.frame(cbind(rep_9.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_9.tmrca.1M
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 3806, p-value = 0.4171
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.1532814
```

```
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 4996, p-value = 0.5562
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.1114572
```

```
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)
```

```r
# for merging:
sim.lands.1Mb.rep9 <- sim.lands.1Mb
sim.lands.1Mb.rep9$Replicate <- 9

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.2409, p-value = 0.7322
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.5612, p-value = 0.677
```

```r
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0004475 -0.0001548 -0.0000562  0.0001403  0.0005293
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.064e-02  5.033e-05 410.041  < 2e-16 ***
## thetaC        1.316e+00  1.246e-02 105.593  < 2e-16 ***
## rhoC          3.532e-02  8.105e-02   0.436   0.6667
## tmrcaC        2.428e-02  2.655e-03   9.145 1.89e-09 ***
## thetaC:tmrcaC 1.211e+00  5.581e-01   2.171   0.0396 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000263 on 25 degrees of freedom
## Multiple R-squared:  0.9986, Adjusted R-squared:  0.9984
## F-statistic:  4623 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.sim.1Mb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 9] <- anova.diversity$VarExp[4] * 100
```

1.3.10 Replicate 10

```
rep_10.pi.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_10/rep_10.pi.1Mb.bedgraph", head
rep_10.tmrca.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims_2/rep_10/sim.tmrca.1M.map", header =

sim.lands.1Mb <- as.data.frame(cbind(rep_10.pi.1Mb$pi, sim.theta.1Mb$sim, sim.rho.1Mb$sim, rep_10.tmrca
names(sim.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")

cor.test(~theta+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 4496, p-value = 1
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##           rho
## -0.0002224694
```

```
cor.test(~theta+rho, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 5654, p-value = 0.1684
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## -0.257842
```

```
cor.test(~rho+tmrca, data = sim.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 5530, p-value = 0.22
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.2302558
```

```
# centering
sim.lands.1Mb$thetaC <- sim.lands.1Mb$theta - mean(sim.lands.1Mb$theta)
sim.lands.1Mb$tmrcaC <- sim.lands.1Mb$tmrca - mean(sim.lands.1Mb$tmrca)
sim.lands.1Mb$rhoC <- sim.lands.1Mb$rho - mean(sim.lands.1Mb$rho)

sim.lands.1Mb$bin <- 1:nrow(sim.lands.1Mb)
```

```r
# for merging:
sim.lands.1Mb.rep50 <- sim.lands.1Mb
sim.lands.1Mb.rep50$Replicate <- 10

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = sim.lands.1Mb)
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.2351, p-value = 0.7114
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.50343, p-value = 0.488
```

```r
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = sim.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -8.386e-04 -1.396e-04 -5.280e-06  1.214e-04  7.480e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.066e-02  5.729e-05 360.639  < 2e-16 ***
## thetaC         1.308e+00  1.273e-02 102.725  < 2e-16 ***
## rhoC          -9.102e-02  9.370e-02  -0.971  0.34066
## tmrcaC         2.340e-02  3.129e-03   7.479 7.84e-08 ***
## thetaC:tmrcaC  1.965e+00  5.383e-01   3.651  0.00121 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003101 on 25 degrees of freedom
## Multiple R-squared:  0.9979, Adjusted R-squared:  0.9976
## F-statistic:  2967 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.sim.1Mb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```
r2.sim.1Mb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.sim.1Mb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.sim.1Mb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.sim.1Mb[5, 10] <- anova.diversity$VarExp[4] * 100


r2.sim.1Mb$average <- rowMeans(r2.sim.1Mb)
r2.sim.1Mb <- transform(r2.sim.1Mb, sd=apply(r2.sim.1Mb, 1, sd, na.rm = TRUE))
```

1.3.11 all replicates:

```
sim.lands.1Mb.all <- rbind(sim.lands.1Mb.rep5, sim.lands.1Mb.rep2, sim.lands.1Mb.rep3, sim.lands.1Mb.rep

m.div.1Mb <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)*as.factor(Replicate), data = sim.la
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.2764, p-value = 0.9657
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.43169, p-value = 0.041
```

```
hist(resid(m.div.1Mb))
```

```
m.div.1Mb.2 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC)*as.factor(Replicate
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC + rhoC:tmrcaC)*as.fa

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb   46 -4169.467
## m.div.1Mb.2 55 -4178.038
## m.div.1Mb.3 64 -4167.214
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC) *
##     as.factor(Replicate), data = sim.lands.1Mb.all)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -8.386e-04 -1.136e-04 -6.930e-06  1.376e-04  7.480e-04
##
## Coefficients:
##                                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)                           2.052e-02  3.986e-05 514.704  < 2e-16 ***
## thetaC                                1.285e+00  8.766e-03 146.613  < 2e-16 ***
## rhoC                                 -4.630e-02  6.495e-02  -0.713  0.47663
## tmrcaC                                2.480e-02  2.440e-03  10.164  < 2e-16 ***
## as.factor(Replicate)3                 1.876e-04  5.876e-05   3.192  0.00159 **
```

```
## as.factor(Replicate)4                -2.641e-05  5.625e-05  -0.469  0.63918
## as.factor(Replicate)5                 8.122e-05  4.867e-05   1.669  0.09640 .
## as.factor(Replicate)6                 3.169e-05  5.701e-05   0.556  0.57872
## as.factor(Replicate)7                 8.047e-05  5.674e-05   1.418  0.15738
## as.factor(Replicate)8                -1.073e-05  5.640e-05  -0.190  0.84927
## as.factor(Replicate)9                 1.217e-04  5.743e-05   2.119  0.03508 *
## as.factor(Replicate)10                1.433e-04  5.641e-05   2.540  0.01167 *
## thetaC:tmrcaC                         2.351e+00  4.650e-01   5.057 8.14e-07 ***
## thetaC:as.factor(Replicate)3          4.083e-02  1.275e-02   3.203  0.00153 **
## thetaC:as.factor(Replicate)4          5.115e-03  1.243e-02   0.412  0.68094
## thetaC:as.factor(Replicate)5          2.072e-02  1.074e-02   1.929  0.05483 .
## thetaC:as.factor(Replicate)6          1.952e-02  1.234e-02   1.581  0.11512
## thetaC:as.factor(Replicate)7          3.691e-02  1.262e-02   2.924  0.00377 **
## thetaC:as.factor(Replicate)8          9.043e-03  1.226e-02   0.738  0.46136
## thetaC:as.factor(Replicate)9          3.056e-02  1.348e-02   2.268  0.02419 *
## thetaC:as.factor(Replicate)10         2.246e-02  1.247e-02   1.801  0.07285 .
## rhoC:as.factor(Replicate)3            1.085e-01  9.179e-02   1.183  0.23810
## rhoC:as.factor(Replicate)4            5.639e-02  9.222e-02   0.611  0.54146
## rhoC:as.factor(Replicate)5            1.134e-02  7.913e-02   0.143  0.88611
## rhoC:as.factor(Replicate)6            1.158e-01  9.191e-02   1.260  0.20868
## rhoC:as.factor(Replicate)7           -4.076e-03  9.123e-02  -0.045  0.96440
## rhoC:as.factor(Replicate)8            6.939e-02  9.091e-02   0.763  0.44600
## rhoC:as.factor(Replicate)9            8.161e-02  9.301e-02   0.877  0.38106
## rhoC:as.factor(Replicate)10          -4.473e-02  9.209e-02  -0.486  0.62761
## tmrcaC:as.factor(Replicate)3          5.904e-03  3.488e-03   1.693  0.09173 .
## tmrcaC:as.factor(Replicate)4         -1.165e-03  3.314e-03  -0.352  0.72538
## tmrcaC:as.factor(Replicate)5          3.357e-04  2.977e-03   0.113  0.91029
## tmrcaC:as.factor(Replicate)6          3.272e-03  4.139e-03   0.790  0.43002
## tmrcaC:as.factor(Replicate)7          3.752e-03  3.263e-03   1.150  0.25125
## tmrcaC:as.factor(Replicate)8          1.784e-03  3.468e-03   0.514  0.60735
## tmrcaC:as.factor(Replicate)9         -5.270e-04  3.273e-03  -0.161  0.87221
## tmrcaC:as.factor(Replicate)10        -1.399e-03  3.272e-03  -0.428  0.66933
## thetaC:tmrcaC:as.factor(Replicate)3   2.416e-01  7.833e-01   0.308  0.75800
## thetaC:tmrcaC:as.factor(Replicate)4  -3.235e-01  5.754e-01  -0.562  0.57447
## thetaC:tmrcaC:as.factor(Replicate)5  -1.217e+00  6.129e-01  -1.985  0.04820 *
## thetaC:tmrcaC:as.factor(Replicate)6  -9.153e-02  9.410e-01  -0.097  0.92259
## thetaC:tmrcaC:as.factor(Replicate)7  -8.443e-01  6.058e-01  -1.394  0.16463
## thetaC:tmrcaC:as.factor(Replicate)8  -3.510e-01  7.646e-01  -0.459  0.64659
## thetaC:tmrcaC:as.factor(Replicate)9  -1.140e+00  6.530e-01  -1.746  0.08206 .
## thetaC:tmrcaC:as.factor(Replicate)10 -3.859e-01  5.974e-01  -0.646  0.51885
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000216 on 255 degrees of freedom
## Multiple R-squared:  0.999,  Adjusted R-squared:  0.9988
## F-statistic:  5772 on 44 and 255 DF,  p-value: < 2.2e-16
```

2 Drosophila-like neutral simulations of chr 2L (Inferred Landscapes)

First batch of 10 replicate datasets simulated using parameters as inferred from real Drosophila data.

Here we present results using genomic landscapes are inferred by iSMC

2.1 50 kb scale

```
r2.inf.50kb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.inf.50kb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.inf.50kb) <- reps

# sim landscapes
sim.rho.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims/dm.sim.rho.50000.txt")
sim.theta.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims/dm.sim.theta.50000.txt")
sim.lands.50kb <- as.data.frame(cbind(sim.theta.50kb$sim, sim.rho.50kb$sim))
names(sim.lands.50kb ) <- c("theta", "rho")
```

2.1.1 Replicate 1

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_1/rs.pair_1.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

inf.lands.50kb.rep_1 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean, t
names(inf.lands.50kb.rep_1) <- c("diversity", "theta", "rho", "tmrca")

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 781576, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9782895
```

```
# centering
inf.lands.50kb.rep_1$thetaC <- inf.lands.50kb.rep_1$theta - mean(inf.lands.50kb.rep_1$theta)
inf.lands.50kb.rep_1$tmrcaC <- inf.lands.50kb.rep_1$tmrca - mean(inf.lands.50kb.rep_1$tmrca)
inf.lands.50kb.rep_1$rhoC <- inf.lands.50kb.rep_1$rho - mean(inf.lands.50kb.rep_1$rho)

inf.lands.50kb.rep_1$bin <- 1:nrow(inf.lands.50kb.rep_1)

# for merging:
inf.lands.50kb.rep_1$Replicate <- 1

m.diversity.rep_1 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_1

plot(resid(m.diversity.rep_1)~fitted(m.diversity.rep_1))
```

```
dwtest(m.diversity.rep_1)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_1
## DW = 1.4157, p-value = 2.024e-13
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_1)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_1
## HMC = 0.61568, p-value = 1
```

```
hist(resid(m.diversity.rep_1))
```

```
summary(m.diversity.rep_1)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_1)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0029908 -0.0002117  0.0000130  0.0001979  0.0035089
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.043e-02  2.236e-05 913.668   <2e-16 ***
## thetaC         1.096e+00  2.364e-03 463.564   <2e-16 ***
## rhoC          -4.461e-03  1.756e-02  -0.254      0.8
## tmrcaC         2.026e-02  1.640e-04 123.483   <2e-16 ***
## thetaC:tmrcaC  1.100e+00  1.587e-02  69.290   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005436 on 595 degrees of freedom
## Multiple R-squared:  0.9977, Adjusted R-squared:  0.9977
## F-statistic: 6.386e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_1, pred = thetaC, modx= tmrcaC)
```

```
g.rep_1 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_1, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_1)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_1
##         AIC       BIC   logLik
##   -7367.645 -7336.866 3690.822
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3212747
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0204388 0.000031255 653.9355  0.0000
## thetaC        1.0980191 0.002985790 367.7482  0.0000
## tmrcaC        0.0199958 0.000176128 113.5300  0.0000
## rhoC          0.0035408 0.017040672   0.2078  0.8355
## thetaC:tmrcaC 1.0653107 0.017126226  62.2035  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.010
## tmrcaC        -0.018 -0.155
## rhoC          -0.003  0.009  0.012
## thetaC:tmrcaC -0.096 -0.101  0.181  0.042
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -5.45341396 -0.37367657  0.05022105  0.36056906  6.43069153
##
## Residual standard error: 0.000544344
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_1)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.030576      1.054551      1.001941      1.041465
```

```
g.rep_1.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.50kb.rep_1, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_1.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_1
##         AIC       BIC   logLik
##   -5449.253 -5427.268 2729.626
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.5079669
##
## Coefficients:
##                 Value  Std.Error  t-value p-value
## (Intercept) 0.0206321 0.00021240 97.13730  0.0000
## thetaC      1.1816867 0.01703572 69.36522  0.0000
## rhoC        0.0057369 0.08137302  0.07050  0.9438
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.003
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.88116216 -0.45332037  0.04478515  0.48762545  4.98107994
##
## Residual standard error: 0.002969561
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_1)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 1] <- anova.diversity$VarExp[4] * 100
```

2.1.2 Replicate 2

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_2/rs.pair_2.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 557174, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9845229
```

```
inf.lands.50kb.rep_2 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean, 
names(inf.lands.50kb.rep_2) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.50kb.rep_2$thetaC <- inf.lands.50kb.rep_2$theta - mean(inf.lands.50kb.rep_2$theta)
inf.lands.50kb.rep_2$tmrcaC <- inf.lands.50kb.rep_2$tmrca - mean(inf.lands.50kb.rep_2$tmrca)
inf.lands.50kb.rep_2$rhoC <- inf.lands.50kb.rep_2$rho - mean(inf.lands.50kb.rep_2$rho)

inf.lands.50kb.rep_2$bin <- 1:nrow(inf.lands.50kb.rep_2)

# for merging:
inf.lands.50kb.rep_2$Replicate <- 2

m.diversity.rep_2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_2

plot(resid(m.diversity.rep_2)~fitted(m.diversity.rep_2))
```

```
dwtest(m.diversity.rep_2)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_2
## DW = 1.3609, p-value = 1.173e-15
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_2)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_2
## HMC = 0.61478, p-value = 1
```

```
hist(resid(m.diversity.rep_2))
```

```
summary(m.diversity.rep_2)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.097e-03 -2.018e-04  1.307e-05  2.192e-04  2.932e-03
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.043e-02  2.069e-05 987.265   <2e-16 ***
## thetaC         1.087e+00  2.167e-03 501.668   <2e-16 ***
## rhoC          -1.167e-02  1.648e-02  -0.708    0.479
## tmrcaC         1.999e-02  1.568e-04 127.438   <2e-16 ***
## thetaC:tmrcaC  1.063e+00  1.365e-02  77.841   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005043 on 595 degrees of freedom
## Multiple R-squared:  0.998,  Adjusted R-squared:  0.9979
## F-statistic: 7.242e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_2, pred = thetaC, modx= tmrcaC)
```

```
g.rep_2 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_2, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_2)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_2
##         AIC       BIC   logLik
##   -7474.789 -7444.011 3744.395
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3693417
##
## Coefficients:
##                     Value   Std.Error  t-value p-value
## (Intercept)    0.0204377 0.000030697 665.7833  0.0000
## thetaC         1.0884205 0.002800847 388.6041  0.0000
## tmrcaC         0.0196845 0.000163272 120.5625  0.0000
## rhoC          -0.0167357 0.015485360  -1.0807  0.2802
## thetaC:tmrcaC  1.0175254 0.014295984  71.1756  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.001
## tmrcaC        -0.004 -0.116
## rhoC          -0.002  0.001  0.001
## thetaC:tmrcaC -0.073 -0.014  0.052  0.045
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -5.967387960 -0.392529054  0.008556582  0.424116285  5.938323518
##
## Residual standard error: 0.0005073101
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_2)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.013673      1.016230      1.002036      1.004808
```

```
g.rep_2.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.50kb.rep_2, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_2.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_2
##         AIC       BIC   logLik
##   -5390.022 -5368.037 2700.011
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3415288
##
## Coefficients:
##                  Value  Std.Error   t-value p-value
## (Intercept)  0.0205980 0.00016691 123.40674  0.0000
## thetaC       1.1278261 0.01545460  72.97670  0.0000
## rhoC        -0.0736804 0.08844426  -0.83307  0.4051
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.003
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -3.532949546 -0.502347038  0.001576082  0.522732374  6.220744946
##
## Residual standard error: 0.002859674
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_2)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 2] <- anova.diversity$VarExp[4] * 100
```

2.1.3 Replicate 3

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_3/rs.pair_3.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 789617, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9780661
```

```
inf.lands.50kb.rep_3 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean,
names(inf.lands.50kb.rep_3) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.50kb.rep_3$thetaC <- inf.lands.50kb.rep_3$theta - mean(inf.lands.50kb.rep_3$theta)
inf.lands.50kb.rep_3$tmrcaC <- inf.lands.50kb.rep_3$tmrca - mean(inf.lands.50kb.rep_3$tmrca)
inf.lands.50kb.rep_3$rhoC <- inf.lands.50kb.rep_3$rho - mean(inf.lands.50kb.rep_3$rho)

inf.lands.50kb.rep_3$bin <- 1:nrow(inf.lands.50kb.rep_3)

# for merging:
inf.lands.50kb.rep_3$Replicate <- 3

m.diversity.rep_3 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_3

plot(resid(m.diversity.rep_3)~fitted(m.diversity.rep_3))
```

```
dwtest(m.diversity.rep_3)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_3
## DW = 1.4619, p-value = 1.133e-11
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_3)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_3
## HMC = 0.52153, p-value = 0.76
```

```
hist(resid(m.diversity.rep_3))
```

```
summary(m.diversity.rep_3)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.464e-03 -1.891e-04  1.021e-05  2.163e-04  2.022e-03
##
## Coefficients:
##                 Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    0.0205593  0.0000198 1038.471   <2e-16 ***
## thetaC         1.0769710  0.0020324  529.902   <2e-16 ***
## rhoC          -0.0138881  0.0157297   -0.883    0.378
## tmrcaC         0.0202409  0.0001408  143.768   <2e-16 ***
## thetaC:tmrcaC  1.0174596  0.0120993   84.092   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004822 on 595 degrees of freedom
## Multiple R-squared:  0.9982, Adjusted R-squared:  0.9982
## F-statistic: 8.432e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_3, pred = thetaC, modx= tmrcaC)
```

```
g.rep_3 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_3, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_3
##         AIC      BIC   logLik
##   -7500.229 -7469.45 3757.114
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.2838818
##
## Coefficients:
##                     Value   Std.Error  t-value p-value
## (Intercept)    0.0205629 0.000026507 775.7537  0.0000
## thetaC         1.0781251 0.002506188 430.1853  0.0000
## tmrcaC         0.0201223 0.000150857 133.3870  0.0000
## rhoC          -0.0178210 0.015324110  -1.1629  0.2453
## thetaC:tmrcaC  0.9985038 0.013270574  75.2419  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.007
## tmrcaC        -0.006 -0.122
## rhoC          -0.004  0.009  0.029
## thetaC:tmrcaC -0.087 -0.075  0.058  0.056
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -5.02048902 -0.38805559  0.03745702  0.44738982  4.48985459
##
## Residual standard error: 0.0004813794
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_3)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.020199      1.018281      1.004142      1.011269
```

```
g.rep_3.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.50kb.rep_3, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_3.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_3
##         AIC       BIC   logLik
##   -5335.691 -5313.707 2672.846
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.4625776
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0207420 0.00021384 96.99954  0.0000
## thetaC       1.1398975 0.01768070 64.47130  0.0000
## rhoC        -0.1017491 0.09007408 -1.12962  0.2591
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.011
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.53093269 -0.44179911 -0.01021147  0.44918244  5.33441211
##
## Residual standard error: 0.003171671
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_3)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 3] <- anova.diversity$VarExp[4] * 100
```

2.1.4 Replicate 4

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_4/rs.pair_4.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 807402, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9775721
```

```
inf.lands.50kb.rep_4 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean,
names(inf.lands.50kb.rep_4) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.50kb.rep_4$thetaC <- inf.lands.50kb.rep_4$theta - mean(inf.lands.50kb.rep_4$theta)
inf.lands.50kb.rep_4$tmrcaC <- inf.lands.50kb.rep_4$tmrca - mean(inf.lands.50kb.rep_4$tmrca)
inf.lands.50kb.rep_4$rhoC <- inf.lands.50kb.rep_4$rho - mean(inf.lands.50kb.rep_4$rho)

inf.lands.50kb.rep_4$bin <- 1:nrow(inf.lands.50kb.rep_4)

# for merging:
inf.lands.50kb.rep_4$Replicate <- 4

m.diversity.rep_4 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_4

plot(resid(m.diversity.rep_4)~fitted(m.diversity.rep_4))
```

```
dwtest(m.diversity.rep_4)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_4
## DW = 0.96213, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_4)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_4
## HMC = 0.70321, p-value = 1
```

```
hist(resid(m.diversity.rep_4))
```

```
summary(m.diversity.rep_4)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0028517 -0.0002016 -0.0000049  0.0001933  0.0041675
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.061e-02  2.304e-05 894.744   <2e-16 ***
## thetaC         1.124e+00  2.418e-03 464.707   <2e-16 ***
## rhoC          -1.112e-02  1.845e-02  -0.603    0.547
## tmrcaC         1.960e-02  1.576e-04 124.378   <2e-16 ***
## thetaC:tmrcaC  1.055e+00  1.499e-02  70.356   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005637 on 595 degrees of freedom
## Multiple R-squared:  0.9975, Adjusted R-squared:  0.9975
## F-statistic: 5.993e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_4, pred = thetaC, modx= tmrcaC)
```

```
g.rep_4 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_4, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_4)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_4
##         AIC       BIC   logLik
##   -7464.493 -7433.715 3739.247
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.5423451
##
## Coefficients:
##                     Value   Std.Error  t-value p-value
## (Intercept)    0.0206153 0.000042516 484.8841  0.0000
## thetaC         1.1223093 0.003267770 343.4480  0.0000
## tmrcaC         0.0195383 0.000157223 124.2717  0.0000
## rhoC          -0.0073321 0.015064681  -0.4867  0.6266
## thetaC:tmrcaC  1.0070650 0.014339665  70.2293  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.001
## tmrcaC        -0.001 -0.069
## rhoC           0.000 -0.009  0.081
## thetaC:tmrcaC -0.024 -0.065 -0.008  0.006
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -5.20520625 -0.37375535 -0.02131367  0.35668752  7.22457562
##
## Residual standard error: 0.0005658744
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_4)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.009184      1.011566      1.006630      1.004416
```

```
g.rep_4.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.50kb.rep_4, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_4.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_4
##         AIC       BIC   logLik
##   -5330.615 -5308.631 2670.308
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.4710929
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0206892 0.00021819 94.82308  0.0000
## thetaC       1.1585830 0.01825400 63.47010  0.0000
## rhoC        -0.1715986 0.09086145 -1.88857  0.0594
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.000  0.003
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -5.45522216 -0.42729625  0.05682276  0.47857205  5.82296892
##
## Residual standard error: 0.00320132
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_4)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 4] <- anova.diversity$VarExp[4] * 100
```

2.1.5 Replicate 5

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_5/rs.pair_5.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 655949, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9817791
```

```
inf.lands.50kb.rep_5 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean,
names(inf.lands.50kb.rep_5) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.50kb.rep_5$thetaC <- inf.lands.50kb.rep_5$theta - mean(inf.lands.50kb.rep_5$theta)
inf.lands.50kb.rep_5$tmrcaC <- inf.lands.50kb.rep_5$tmrca - mean(inf.lands.50kb.rep_5$tmrca)
inf.lands.50kb.rep_5$rhoC <- inf.lands.50kb.rep_5$rho - mean(inf.lands.50kb.rep_5$rho)

inf.lands.50kb.rep_5$bin <- 1:nrow(inf.lands.50kb.rep_5)

# for merging:
inf.lands.50kb.rep_5$Replicate <- 5

m.diversity.rep_5 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_5

plot(resid(m.diversity.rep_5)~fitted(m.diversity.rep_5))
```

```
dwtest(m.diversity.rep_5)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_5
## DW = 1.4275, p-value = 5.886e-13
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_5)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_5
## HMC = 0.55418, p-value = 0.976
```

```
hist(resid(m.diversity.rep_5))
```

```
summary(m.diversity.rep_5)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_5)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0020455 -0.0002004  0.0000140  0.0002247  0.0033998
##
## Coefficients:
##                 Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    0.0204679  0.0000189 1082.787   <2e-16 ***
## thetaC         1.1078321  0.0019892  556.934   <2e-16 ***
## rhoC          -0.0069343  0.0156805   -0.442    0.658
## tmrcaC         0.0198380  0.0001375  144.318   <2e-16 ***
## thetaC:tmrcaC  1.0621447  0.0118730   89.459   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004621 on 595 degrees of freedom
## Multiple R-squared:  0.9983, Adjusted R-squared:  0.9983
## F-statistic: 8.634e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_5, pred = thetaC, modx= tmrcaC)
```

```
g.rep_5 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_5, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_5)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_5
##         AIC       BIC   logLik
##   -7570.892 -7540.114 3792.446
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3687999
##
## Coefficients:
##                     Value   Std.Error  t-value p-value
## (Intercept)    0.0204733 0.000028264 724.3584  0.0000
## thetaC         1.1067703 0.002603042 425.1835  0.0000
## tmrcaC         0.0195333 0.000148186 131.8159  0.0000
## rhoC          -0.0117411 0.014941182  -0.7858  0.4323
## thetaC:tmrcaC  1.0119684 0.012760307  79.3060  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.002
## tmrcaC         0.002 -0.094
## rhoC          -0.001  0.002  0.058
## thetaC:tmrcaC -0.045  0.037 -0.040  0.028
##
## Standardized residuals:
##           Min            Q1           Med            Q3           Max
## -4.3488500199 -0.4439991675  0.0002670608  0.4775851241  7.1344633188
##
## Residual standard error: 0.0004681579
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_5)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.010139      1.013921      1.004356      1.003709
```

```
g.rep_5.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.50kb.rep_5, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_5.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_5
##         AIC       BIC  logLik
##   -5329.279 -5307.294 2669.64
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.4006009
##
## Coefficients:
##                  Value  Std.Error   t-value p-value
## (Intercept)  0.0205709 0.00019283 106.68153  0.0000
## thetaC       1.1335447 0.01727411  65.62102  0.0000
## rhoC        -0.1554325 0.09617437  -1.61615  0.1066
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.007
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.31295794 -0.44285975  0.05270967  0.47530087  5.66926609
##
## Residual standard error: 0.003085565
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_5)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 5] <- anova.diversity$VarExp[4] * 100
```

2.1.6 Replicate 6

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_6/rs.pair_6.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 576502, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.983986
```

```
inf.lands.50kb.rep_6 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean, 
names(inf.lands.50kb.rep_6) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.50kb.rep_6$thetaC <- inf.lands.50kb.rep_6$theta - mean(inf.lands.50kb.rep_6$theta)
inf.lands.50kb.rep_6$tmrcaC <- inf.lands.50kb.rep_6$tmrca - mean(inf.lands.50kb.rep_6$tmrca)
inf.lands.50kb.rep_6$rhoC <- inf.lands.50kb.rep_6$rho - mean(inf.lands.50kb.rep_6$rho)

inf.lands.50kb.rep_6$bin <- 1:nrow(inf.lands.50kb.rep_6)

# for merging:
inf.lands.50kb.rep_6$Replicate <- 6

m.diversity.rep_6 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_6

plot(resid(m.diversity.rep_6)~fitted(m.diversity.rep_6))
```

```
dwtest(m.diversity.rep_6)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_6
## DW = 1.4152, p-value = 1.956e-13
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_6)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_6
## HMC = 0.5501, p-value = 0.966
```

```
hist(resid(m.diversity.rep_6))
```

```
summary(m.diversity.rep_6)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0025487 -0.0002073 -0.0000008  0.0002258  0.0049661
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.046e-02  2.183e-05 937.468   <2e-16 ***
## thetaC        1.086e+00  2.249e-03 482.709   <2e-16 ***
## rhoC          9.344e-03  1.682e-02   0.555    0.579
## tmrcaC        2.016e-02  1.691e-04 119.259   <2e-16 ***
## thetaC:tmrcaC 1.050e+00  1.470e-02  71.423   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005333 on 595 degrees of freedom
## Multiple R-squared:  0.9977, Adjusted R-squared:  0.9977
## F-statistic: 6.529e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_6, pred = thetaC, modx= tmrcaC)
```

```
g.rep_6 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_6, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_6)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_6
##        AIC       BIC  logLik
##   -7391.16 -7360.382 3702.58
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3219522
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0204681 0.000030586 669.1903  0.0000
## thetaC        1.0858244 0.002832026 383.4091  0.0000
## tmrcaC        0.0199376 0.000176983 112.6524  0.0000
## rhoC          0.0002293 0.016283078   0.0141  0.9888
## thetaC:tmrcaC 1.0151180 0.015909552  63.8056  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.000
## tmrcaC        -0.006 -0.082
## rhoC           0.000  0.014  0.075
## thetaC:tmrcaC -0.056  0.002  0.082  0.013
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -4.664990405 -0.377078774 -0.001461415  0.411936942  9.302435018
##
## Residual standard error: 0.0005339103
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_6)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.007216      1.019473      1.006158      1.006869
```

```
g.rep_6.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.50kb.rep_6, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_6.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_6
##         AIC       BIC   logLik
##   -5414.912 -5392.927 2712.456
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.4107098
##
## Coefficients:
##                  Value  Std.Error   t-value p-value
## (Intercept)  0.0205846 0.00018261 112.72153  0.0000
## thetaC       1.1129166 0.01585204  70.20654  0.0000
## rhoC        -0.1062152 0.08342243  -1.27322  0.2034
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.014
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -4.246615321 -0.484784795 -0.004012756  0.502253352  5.520109924
##
## Residual standard error: 0.002887164
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_6)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 6] <- anova.diversity$VarExp[4] * 100
```

2.1.7 Replicate 7

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_7/rs.pair_7.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 715696, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9801195
```

```
inf.lands.50kb.rep_7 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean, 
names(inf.lands.50kb.rep_7) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.50kb.rep_7$thetaC <- inf.lands.50kb.rep_7$theta - mean(inf.lands.50kb.rep_7$theta)
inf.lands.50kb.rep_7$tmrcaC <- inf.lands.50kb.rep_7$tmrca - mean(inf.lands.50kb.rep_7$tmrca)
inf.lands.50kb.rep_7$rhoC <- inf.lands.50kb.rep_7$rho - mean(inf.lands.50kb.rep_7$rho)

inf.lands.50kb.rep_7$bin <- 1:nrow(inf.lands.50kb.rep_7)

# for merging:
inf.lands.50kb.rep_7$Replicate <- 7

m.diversity.rep_7 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_7

plot(resid(m.diversity.rep_7)~fitted(m.diversity.rep_7))
```

```
dwtest(m.diversity.rep_7)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_7
## DW = 1.2084, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_7)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_7
## HMC = 0.58586, p-value = 1
```

```
hist(resid(m.diversity.rep_7))
```

```
summary(m.diversity.rep_7)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_7)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0023413 -0.0001931  0.0000183  0.0002113  0.0034898
##
## Coefficients:
##                 Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    2.046e-02  1.906e-05 1073.378  < 2e-16 ***
## thetaC         1.093e+00  1.994e-03  548.323  < 2e-16 ***
## rhoC          -3.984e-02  1.517e-02   -2.627  0.00885 **
## tmrcaC         2.021e-02  1.462e-04  138.265  < 2e-16 ***
## thetaC:tmrcaC  1.056e+00  1.344e-02   78.551  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004657 on 595 degrees of freedom
## Multiple R-squared:  0.9982, Adjusted R-squared:  0.9982
## F-statistic: 8.313e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_7, pred = thetaC, modx= tmrcaC)
```

```
g.rep_7 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_7, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_7)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_7
##        AIC       BIC   logLik
##   -7617.29 -7586.512 3815.645
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.4690905
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)    0.0204701 0.000032313 633.4967  0.0000
## thetaC         1.0964625 0.002720215 403.0793  0.0000
## tmrcaC         0.0197764 0.000149360 132.4076  0.0000
## rhoC          -0.0310546 0.013470566  -2.3054  0.0215
## thetaC:tmrcaC  0.9959236 0.014228586  69.9946  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.001
## tmrcaC        -0.006 -0.099
## rhoC           0.002  0.005 -0.022
## thetaC:tmrcaC -0.047  0.014  0.100 -0.014
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.79375373 -0.41801231  0.05547631  0.45600293  7.65648237
##
## Residual standard error: 0.0004740043
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_7)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.010511      1.020984      1.000614      1.010927
```

```
g.rep_7.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.50kb.rep_7, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_7.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_7
##         AIC       BIC   logLik
##   -5472.449 -5450.464 2741.224
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##      Phi
## 0.469828
##
## Coefficients:
##                 Value  Std.Error   t-value p-value
## (Intercept) 0.0205801 0.00019340 106.41030  0.0000
## thetaC      1.1283864 0.01620181  69.64571  0.0000
## rhoC        0.0166423 0.08056264   0.20658  0.8364
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.003
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.47464746 -0.42695326  0.04565358  0.50179301  5.14163869
##
## Residual standard error: 0.002842278
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_7)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 7] <- anova.diversity$VarExp[4] * 100
```

2.1.8 Replicate 8

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_8/rs.pair_8.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 682855, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9810318
```

```
inf.lands.50kb.rep_8 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean,
names(inf.lands.50kb.rep_8) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.50kb.rep_8$thetaC <- inf.lands.50kb.rep_8$theta - mean(inf.lands.50kb.rep_8$theta)
inf.lands.50kb.rep_8$tmrcaC <- inf.lands.50kb.rep_8$tmrca - mean(inf.lands.50kb.rep_8$tmrca)
inf.lands.50kb.rep_8$rhoC <- inf.lands.50kb.rep_8$rho - mean(inf.lands.50kb.rep_8$rho)

inf.lands.50kb.rep_8$bin <- 1:nrow(inf.lands.50kb.rep_8)

# for merging:
inf.lands.50kb.rep_8$Replicate <- 8

m.diversity.rep_8 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_8

plot(resid(m.diversity.rep_8)~fitted(m.diversity.rep_8))
```

```
dwtest(m.diversity.rep_8)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_8
## DW = 1.3796, p-value = 6.877e-15
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_8)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_8
## HMC = 0.62278, p-value = 1
```

```
hist(resid(m.diversity.rep_8))
```

```
summary(m.diversity.rep_8)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_8)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.362e-03 -1.893e-04  3.500e-07  2.086e-04  2.138e-03
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.070e-02  1.938e-05 1067.98   <2e-16 ***
## thetaC         1.107e+00  2.023e-03  547.27   <2e-16 ***
## rhoC          -6.776e-03  1.539e-02   -0.44     0.66
## tmrcaC         2.013e-02  1.393e-04  144.50   <2e-16 ***
## thetaC:tmrcaC  1.071e+00  1.222e-02   87.61   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004737 on 595 degrees of freedom
## Multiple R-squared:  0.9982, Adjusted R-squared:  0.9982
## F-statistic: 8.292e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_8, pred = thetaC, modx= tmrcaC)
```

```
g.rep_8 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_8, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_8)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_8
##         AIC       BIC   logLik
##   -7548.618 -7517.839 3781.309
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3796214
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0207031 0.000029299 706.6176  0.0000
## thetaC        1.1082548 0.002641688 419.5253  0.0000
## tmrcaC        0.0198182 0.000148295 133.6401  0.0000
## rhoC          0.0047060 0.014548476   0.3235  0.7465
## thetaC:tmrcaC 1.0223058 0.013460362  75.9494  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.001
## tmrcaC        -0.004 -0.069
## rhoC           0.001  0.012  0.067
## thetaC:tmrcaC -0.048  0.026  0.075 -0.015
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.79143530 -0.40162824  0.00873902  0.44671770  4.52028563
##
## Residual standard error: 0.0004791766
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_8)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.006009      1.015515      1.005156      1.006994
```

```
g.rep_8.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.50kb.rep_8, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_8.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_8
##         AIC       BIC  logLik
##   -5369.959 -5347.975 2689.98
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.4676478
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0208143 0.00020979 99.21612  0.0000
## thetaC       1.1333322 0.01746884 64.87738  0.0000
## rhoC        -0.0900602 0.08770082 -1.02690  0.3049
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.000  0.012
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -7.41084900 -0.48902470 -0.01613762  0.49491223  5.48392430
##
## Residual standard error: 0.003091654
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_8)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 8] <- anova.diversity$VarExp[4] * 100
```

2.1.9 Replicate 9

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_9/rs.pair_9.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 609380, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9830727
```

```
inf.lands.50kb.rep_9 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean, 
names(inf.lands.50kb.rep_9) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.50kb.rep_9$thetaC <- inf.lands.50kb.rep_9$theta - mean(inf.lands.50kb.rep_9$theta)
inf.lands.50kb.rep_9$tmrcaC <- inf.lands.50kb.rep_9$tmrca - mean(inf.lands.50kb.rep_9$tmrca)
inf.lands.50kb.rep_9$rhoC <- inf.lands.50kb.rep_9$rho - mean(inf.lands.50kb.rep_9$rho)

inf.lands.50kb.rep_9$bin <- 1:nrow(inf.lands.50kb.rep_9)

# for merging:
inf.lands.50kb.rep_9$Replicate <- 9

m.diversity.rep_9 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_9

plot(resid(m.diversity.rep_9)~fitted(m.diversity.rep_9))
```

```
dwtest(m.diversity.rep_9)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_9
## DW = 1.3455, p-value = 2.541e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_9)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_9
## HMC = 0.52947, p-value = 0.858
```

```
hist(resid(m.diversity.rep_9))
```

```
summary(m.diversity.rep_9)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_9)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0028280 -0.0002274  0.0000078  0.0002235  0.0033670
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.037e-02  2.069e-05 984.748   <2e-16 ***
## thetaC        1.086e+00  2.147e-03 505.918   <2e-16 ***
## rhoC          1.171e-02  1.718e-02   0.682    0.496
## tmrcaC        1.990e-02  1.617e-04 123.090   <2e-16 ***
## thetaC:tmrcaC 1.045e+00  1.408e-02  74.239   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005043 on 595 degrees of freedom
## Multiple R-squared:  0.998,  Adjusted R-squared:  0.998
## F-statistic: 7.306e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_9, pred = thetaC, modx= tmrcaC)
```

```
g.rep_9 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_9, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_9)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_9
##         AIC       BIC   logLik
##   -7487.186 -7456.408 3750.593
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##      Phi
## 0.412105
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0203835 0.000032571 625.8123  0.0000
## thetaC        1.0864837 0.002863160 379.4701  0.0000
## tmrcaC        0.0194098 0.000164202 118.2072  0.0000
## rhoC          0.0014542 0.015946130   0.0912  0.9274
## thetaC:tmrcaC 0.9860233 0.014806396  66.5944  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.003
## tmrcaC         0.001 -0.125
## rhoC          -0.004  0.017  0.041
## thetaC:tmrcaC -0.067 -0.040 -0.014  0.076
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.86419788 -0.43551103  0.02063926  0.43146966  6.60989138
##
## Residual standard error: 0.0005120851
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_9)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.018252      1.018249      1.008251      1.008039
```

```
g.rep_9.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.50kb.rep_9, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_9.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_9
##         AIC       BIC   logLik
##   -5406.246 -5384.261 2708.123
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3833499
##
## Coefficients:
##                  Value  Std.Error   t-value p-value
## (Intercept)  0.0205299 0.00017580 116.77810  0.0000
## thetaC       1.1341848 0.01570929  72.19836  0.0000
## rhoC        -0.1665026 0.09061323  -1.83751  0.0666
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.026
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.88951873 -0.46241592  0.05059189  0.47484598  6.24257165
##
## Residual standard error: 0.002870881
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_9)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 9] <- anova.diversity$VarExp[4] * 100
```

2.1.10 Replicate 10

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_10/rs.pair_10.", sep = "")
pi.50kb <- read.table(paste(p, "diversity.50kb.bedgraph", sep = ""), header = T)
pi.50kb$avg <- apply(pi.50kb[4:ncol(pi.50kb)], 1, mean)
tmrca.50kb <- read.table(paste(p, "TMRCA.50kb.bedgraph", sep = ""), header = T)
tmrca.50kb$avg <- apply(tmrca.50kb[4:ncol(tmrca.50kb)], 1, mean)
rho.50kb <- read.table(paste(p, "rho.50kb.bedgraph", sep = ""), header = T)
theta.50kb <- read.table(paste(p, "theta.50kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.50kb$sim, theta.50kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.50kb$sim and theta.50kb$sample_mean
## S = 558257, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9844928
```

```
inf.lands.50kb.rep_10 <- as.data.frame(cbind(pi.50kb$avg, theta.50kb$sample_mean, rho.50kb$sample_mean,
names(inf.lands.50kb.rep_10) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.50kb.rep_10$thetaC <- inf.lands.50kb.rep_10$theta - mean(inf.lands.50kb.rep_10$theta)
inf.lands.50kb.rep_10$tmrcaC <- inf.lands.50kb.rep_10$tmrca - mean(inf.lands.50kb.rep_10$tmrca)
inf.lands.50kb.rep_10$rhoC <- inf.lands.50kb.rep_10$rho - mean(inf.lands.50kb.rep_10$rho)

inf.lands.50kb.rep_10$bin <- 1:nrow(inf.lands.50kb.rep_10)

# for merging:
inf.lands.50kb.rep_10$Replicate <- 10

m.diversity.rep_10 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.50kb.rep_

plot(resid(m.diversity.rep_10)~fitted(m.diversity.rep_10))
```

```
dwtest(m.diversity.rep_10)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_10
## DW = 1.3449, p-value = 2.3e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_10)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_10
## HMC = 0.67092, p-value = 1
```

```
hist(resid(m.diversity.rep_10))
```

```
summary(m.diversity.rep_10)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.50kb.rep_10)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0021127 -0.0002121 -0.0000085  0.0001956  0.0054007
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.048e-02  2.313e-05 885.243   <2e-16 ***
## thetaC         1.074e+00  2.407e-03 446.391   <2e-16 ***
## rhoC          -2.139e-02  1.771e-02  -1.208    0.227
## tmrcaC         2.070e-02  1.767e-04 117.152   <2e-16 ***
## thetaC:tmrcaC  1.096e+00  1.562e-02  70.181   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000561 on 595 degrees of freedom
## Multiple R-squared:  0.9975, Adjusted R-squared:  0.9975
## F-statistic: 6.053e+04 on 4 and 595 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_10, pred = thetaC, modx= tmrcaC)
```

```
g.rep_10 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.50kb.rep_1, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_10)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep_1
##         AIC       BIC   logLik
##   -7367.645 -7336.866 3690.822
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3212747
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0204388 0.000031255 653.9355  0.0000
## thetaC        1.0980191 0.002985790 367.7482  0.0000
## tmrcaC        0.0199958 0.000176128 113.5300  0.0000
## rhoC          0.0035408 0.017040672   0.2078  0.8355
## thetaC:tmrcaC 1.0653107 0.017126226  62.2035  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.010
## tmrcaC        -0.018 -0.155
## rhoC          -0.003  0.009  0.012
## thetaC:tmrcaC -0.096 -0.101  0.181  0.042
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -5.45341396 -0.37367657  0.05022105  0.36056906  6.43069153
##
## Residual standard error: 0.000544344
## Degrees of freedom: 600 total; 595 residual
```

```
vif(g.rep_10)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.030576      1.054551      1.001941      1.041465
```

```
g.rep_10.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                         data = inf.lands.50kb.rep_10, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_10.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.50kb.rep_10
##         AIC       BIC   logLik
##   -5437.978 -5415.994 2723.989
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.4549714
##
## Coefficients:
##                  Value  Std.Error   t-value p-value
## (Intercept)  0.0207082 0.00019363 106.94544  0.0000
## thetaC       1.1606422 0.01631116  71.15634  0.0000
## rhoC        -0.1660896 0.08076784  -2.05638  0.0402
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.000
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.86544315 -0.50573567  0.01449108  0.52601610  6.72401413
##
## Residual standard error: 0.002899699
## Degrees of freedom: 600 total; 597 residual
```

```
anova.diversity <- Anova(m.diversity.rep_10)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.50kb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.50kb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.inf.50kb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.inf.50kb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.inf.50kb[5, 10] <- anova.diversity$VarExp[4] * 100


r2.inf.50kb$average <- rowMeans(r2.inf.50kb)
r2.inf.50kb <- transform(r2.inf.50kb, sd=apply(r2.inf.50kb, 1, sd, na.rm = TRUE))
```

2.1.11 all replicates:

```
inf.lands.50kb.all <- rbind(inf.lands.50kb.rep_1, inf.lands.50kb.rep_2, inf.lands.50kb.rep_3, inf.lands

m.div.50kb <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)*as.factor(Replicate), data = inf.l
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.3249, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.50366, p-value = 0.651
```

```r
hist(resid(m.div.50kb))
```

```r
m.div.50kb.2 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC)*as.factor(Replica
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC + rhoC:tmrcaC)*as.

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb   51 -73877.32
## m.div.50kb.2 61 -73868.89
## m.div.50kb.3 71 -73992.94
```

```r
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC) *
##     as.factor(Replicate), data = inf.lands.50kb.all)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0030973 -0.0002036  0.0000076  0.0002133  0.0054007
##
## Coefficients:
##                                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)                           2.043e-02  2.101e-05 972.548  < 2e-16 ***
## thetaC                                1.096e+00  2.221e-03 493.438  < 2e-16 ***
## rhoC                                 -4.461e-03  1.650e-02  -0.270 0.786909
## tmrcaC                                2.026e-02  1.541e-04 131.441  < 2e-16 ***
## as.factor(Replicate)2                -2.363e-06  2.968e-05  -0.080 0.936530
```

```
## as.factor(Replicate)3                 1.266e-04  2.968e-05   4.266 2.02e-05 ***
## as.factor(Replicate)4                 1.793e-04  2.961e-05   6.054 1.50e-09 ***
## as.factor(Replicate)5                 3.524e-05  2.963e-05   1.189 0.234326
## as.factor(Replicate)6                 3.133e-05  2.964e-05   1.057 0.290526
## as.factor(Replicate)7                 3.036e-05  2.964e-05   1.024 0.305728
## as.factor(Replicate)8                 2.651e-04  2.963e-05   8.947  < 2e-16 ***
## as.factor(Replicate)9                -5.825e-05  2.967e-05  -1.963 0.049696 *
## as.factor(Replicate)10                4.479e-05  2.975e-05   1.506 0.132197
## thetaC:tmrcaC                         1.100e+00  1.491e-02  73.755  < 2e-16 ***
## thetaC:as.factor(Replicate)2         -8.685e-03  3.122e-03  -2.782 0.005420 **
## thetaC:as.factor(Replicate)3         -1.871e-02  3.093e-03  -6.049 1.54e-09 ***
## thetaC:as.factor(Replicate)4          2.796e-02  3.119e-03   8.963  < 2e-16 ***
## thetaC:as.factor(Replicate)5          1.215e-02  3.125e-03   3.889 0.000102 ***
## thetaC:as.factor(Replicate)6         -1.002e-02  3.094e-03  -3.240 0.001202 **
## thetaC:as.factor(Replicate)7         -2.283e-03  3.117e-03  -0.732 0.463905
## thetaC:as.factor(Replicate)8          1.151e-02  3.113e-03   3.697 0.000220 ***
## thetaC:as.factor(Replicate)9         -9.252e-03  3.108e-03  -2.977 0.002925 **
## thetaC:as.factor(Replicate)10        -2.131e-02  3.120e-03  -6.832 9.20e-12 ***
## rhoC:as.factor(Replicate)2           -7.206e-03  2.347e-02  -0.307 0.758849
## rhoC:as.factor(Replicate)3           -9.427e-03  2.345e-02  -0.402 0.687674
## rhoC:as.factor(Replicate)4           -6.662e-03  2.349e-02  -0.284 0.776732
## rhoC:as.factor(Replicate)5           -2.473e-03  2.393e-02  -0.103 0.917680
## rhoC:as.factor(Replicate)6            1.381e-02  2.306e-02   0.599 0.549449
## rhoC:as.factor(Replicate)7           -3.538e-02  2.343e-02  -1.510 0.131105
## rhoC:as.factor(Replicate)8           -2.315e-03  2.340e-02  -0.099 0.921183
## rhoC:as.factor(Replicate)9            1.617e-02  2.398e-02   0.674 0.500029
## rhoC:as.factor(Replicate)10          -1.693e-02  2.307e-02  -0.734 0.462966
## tmrcaC:as.factor(Replicate)2         -2.689e-04  2.213e-04  -1.215 0.224342
## tmrcaC:as.factor(Replicate)3         -1.546e-05  2.144e-04  -0.072 0.942513
## tmrcaC:as.factor(Replicate)4         -6.536e-04  2.101e-04  -3.111 0.001874 **
## tmrcaC:as.factor(Replicate)5         -4.184e-04  2.164e-04  -1.933 0.053249 .
## tmrcaC:as.factor(Replicate)6         -9.362e-05  2.235e-04  -0.419 0.675348
## tmrcaC:as.factor(Replicate)7         -4.557e-05  2.224e-04  -0.205 0.837659
## tmrcaC:as.factor(Replicate)8         -1.258e-04  2.152e-04  -0.584 0.558918
## tmrcaC:as.factor(Replicate)9         -3.572e-04  2.248e-04  -1.589 0.112145
## tmrcaC:as.factor(Replicate)10         4.449e-04  2.228e-04   1.997 0.045880 *
## thetaC:tmrcaC:as.factor(Replicate)2  -3.725e-02  2.033e-02  -1.832 0.067009 .
## thetaC:tmrcaC:as.factor(Replicate)3  -8.236e-02  1.966e-02  -4.189 2.84e-05 ***
## thetaC:tmrcaC:as.factor(Replicate)4  -4.519e-02  2.017e-02  -2.241 0.025093 *
## thetaC:tmrcaC:as.factor(Replicate)5  -3.768e-02  1.986e-02  -1.897 0.057874 .
## thetaC:tmrcaC:as.factor(Replicate)6  -4.980e-02  2.051e-02  -2.428 0.015191 *
## thetaC:tmrcaC:as.factor(Replicate)7  -4.418e-02  2.097e-02  -2.107 0.035148 *
## thetaC:tmrcaC:as.factor(Replicate)8  -2.915e-02  1.990e-02  -1.465 0.142953
## thetaC:tmrcaC:as.factor(Replicate)9  -5.477e-02  2.063e-02  -2.655 0.007954 **
## thetaC:tmrcaC:as.factor(Replicate)10 -3.768e-03  2.060e-02  -0.183 0.854904
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005107 on 5950 degrees of freedom
## Multiple R-squared:  0.9979, Adjusted R-squared:  0.9979
## F-statistic: 5.859e+04 on 49 and 5950 DF,  p-value: < 2.2e-16
```

2.1.12 Plots

```
scale.3d <- function(x) sprintf("%.3f", x)

rho.plot <- as.data.frame(cbind(inf.lands.50kb.rep_1$bin,
                                sim.rho.50kb$sim,
                                inf.lands.50kb.rep_1$rho,
                                inf.lands.50kb.rep_2$rho,
                                inf.lands.50kb.rep_3$rho,
                                inf.lands.50kb.rep_4$rho,
                                inf.lands.50kb.rep_5$rho,
                                inf.lands.50kb.rep_6$rho,
                                inf.lands.50kb.rep_7$rho,
                                inf.lands.50kb.rep_8$rho,
                                inf.lands.50kb.rep_9$rho,
                                inf.lands.50kb.rep_10$rho))

rho.plot[,3:ncol(rho.plot)] <- 1.53 * rho.plot[,3:ncol(rho.plot)] # scaling because of SCRM N0 vs Ne

names(rho.plot) <- c("bin", "sim", reps)
molten.rho <- melt(rho.plot, id.vars = "bin")
rho.map.50kb <- ggplot(data = molten.rho, aes(x = bin * 50, y = value, colour = variable)) + theme_bw()
rho.map.50kb <- rho.map.50kb + geom_line(data = molten.rho, aes(size = variable)) + scale_size_manual(va
rho.map.50kb <- rho.map.50kb + scale_color_manual(values = c("black", brewer.pal(n = 9, name = "Blues")
rho.map.50kb <- rho.map.50kb + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks
rho.map.50kb <- rho.map.50kb + labs(title = NULL, x = NULL, y = expression(rho))
rho.map.50kb <- rho.map.50kb + theme(text = element_text(size = 20), axis.title.x = element_text(size =

theta.plot <- as.data.frame(cbind(inf.lands.50kb.rep_1$bin,
                                sim.theta.50kb$sim,
                                inf.lands.50kb.rep_1$theta,
                                inf.lands.50kb.rep_2$theta,
                                inf.lands.50kb.rep_3$theta,
                                inf.lands.50kb.rep_4$theta,
                                inf.lands.50kb.rep_5$theta,
                                inf.lands.50kb.rep_6$theta,
                                inf.lands.50kb.rep_7$theta,
                                inf.lands.50kb.rep_8$theta,
                                inf.lands.50kb.rep_9$theta,
                                inf.lands.50kb.rep_10$theta))

names(theta.plot) <- c("bin", "sim", reps)
molten.theta <- melt(theta.plot, id.vars = "bin")
theta.map.50kb <- ggplot(data = molten.theta, aes(x = bin * 50, y = value, colour = variable)) + theme_b
theta.map.50kb <- theta.map.50kb + geom_line(data = molten.theta, aes(size = variable)) + scale_size_man
theta.map.50kb <- theta.map.50kb + scale_color_manual(values = c("black", brewer.pal(n = 9, name = "Reds
theta.map.50kb <- theta.map.50kb + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(bre
theta.map.50kb <- theta.map.50kb + labs(title = NULL, x = NULL, y = expression(theta))
theta.map.50kb <- theta.map.50kb + theme(text = element_text(size = 20), axis.title.x = element_text(si

plot_grid(rho.map.50kb, theta.map.50kb, nrow = 2, ncol = 1)
```

2.2 200 kb scale

```
r2.inf.200kb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.inf.200kb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.inf.200kb) <- reps

# sim landscapes
sim.rho.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims/dm.sim.rho.2e+05.txt")
sim.theta.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims/dm.sim.theta.2e+05.txt")
sim.lands.200kb  <- as.data.frame(cbind(sim.theta.200kb$sim, sim.rho.200kb$sim))
names(sim.lands.200kb ) <- c("theta", "rho")
```

2.2.1 Replicate 1

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_1/rs.pair_1.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
```

```
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 8264.3, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9853072
```

```
inf.lands.200kb.rep_1 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_mea
names(inf.lands.200kb.rep_1) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_1$thetaC <- inf.lands.200kb.rep_1$theta - mean(inf.lands.200kb.rep_1$theta)
inf.lands.200kb.rep_1$tmrcaC <- inf.lands.200kb.rep_1$tmrca - mean(inf.lands.200kb.rep_1$tmrca)
inf.lands.200kb.rep_1$rhoC <- inf.lands.200kb.rep_1$rho - mean(inf.lands.200kb.rep_1$rho)

inf.lands.200kb.rep_1$bin <- 1:nrow(inf.lands.200kb.rep_1)

# for merging:
inf.lands.200kb.rep_1$Replicate <- 1

m.diversity.rep_1 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep_

plot(resid(m.diversity.rep_1)~fitted(m.diversity.rep_1))
```

```
dwtest(m.diversity.rep_1)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_1
## DW = 1.8429, p-value = 0.1407
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.diversity.rep_1)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_1
## HMC = 0.49224, p-value = 0.441
```

```r
hist(resid(m.diversity.rep_1))
```

```r
summary(m.diversity.rep_1)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_1)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.575e-03 -2.359e-04 -1.035e-05  1.834e-04  2.399e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.049e-02  4.583e-05 447.178   <2e-16 ***
## thetaC        1.096e+00  5.418e-03 202.263   <2e-16 ***
## rhoC          1.426e-02  6.350e-02   0.225    0.823
## tmrcaC        2.103e-02  5.010e-04  41.966   <2e-16 ***
## thetaC:tmrcaC 1.224e+00  5.186e-02  23.595   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.0005573 on 145 degrees of freedom
## Multiple R-squared:  0.9968, Adjusted R-squared:  0.9967
## F-statistic: 1.131e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_1, pred = thetaC, modx= tmrcaC)
```

```
g.rep_1 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_1, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_1)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_1
##         AIC       BIC   logLik
##   -1814.128 -1793.054 914.0641
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##        Phi
## 0.08562895
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0204958 0.00004989 410.8349  0.0000
## thetaC        1.0967285 0.00562895 194.8371  0.0000
```

```
## tmrcaC        0.0210125 0.00050621  41.5092  0.0000
## rhoC          0.0219683 0.06340360   0.3465  0.7295
## thetaC:tmrcaC 1.2110957 0.05322577  22.7539  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
## thetaC        -0.001
## tmrcaC        -0.043 -0.110
## rhoC          -0.010 -0.009  0.248
## thetaC:tmrcaC -0.113  0.010  0.375  0.090
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.94983846 -0.42035716 -0.01013693  0.33115763  4.40016170
##
## Residual standard error: 0.0005481049
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_1)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.015853      1.249427      1.066075      1.167458
```

```
g.rep_1.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.200kb.rep_1, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_1.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_1
##         AIC       BIC   logLik
##   -1432.216 -1417.163 721.1082
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.2090144
##
## Coefficients:
##                  Value  Std.Error   t-value p-value
## (Intercept)  0.0206297 0.00020572 100.27847  0.0000
## thetaC       1.1244227 0.02111497  53.25239  0.0000
## rhoC        -0.5711032 0.21962397  -2.60037  0.0103
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.015
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.46691260 -0.42444799  0.03262027  0.46216734  3.67858710
##
```

```
## Residual standard error: 0.002021045
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_1)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.inf.200kb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
r2.inf.200kb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 1] <- anova.diversity$VarExp[4] * 100
```

2.2.2 Replicate 2

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_2/rs.pair_2.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 6148.2, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9890693
```

```
inf.lands.200kb.rep_2 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_mea
names(inf.lands.200kb.rep_2) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_2$thetaC <- inf.lands.200kb.rep_2$theta - mean(inf.lands.200kb.rep_2$theta)
inf.lands.200kb.rep_2$tmrcaC <- inf.lands.200kb.rep_2$tmrca - mean(inf.lands.200kb.rep_2$tmrca)
inf.lands.200kb.rep_2$rhoC <- inf.lands.200kb.rep_2$rho - mean(inf.lands.200kb.rep_2$rho)

inf.lands.200kb.rep_2$bin <- 1:nrow(inf.lands.200kb.rep_2)

# for merging:
inf.lands.200kb.rep_2$Replicate <- 2

m.diversity.rep_2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep_2

plot(resid(m.diversity.rep_2)~fitted(m.diversity.rep_2))
```

```
dwtest(m.diversity.rep_2)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_2
## DW = 1.7156, p-value = 0.03159
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_2)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_2
## HMC = 0.50531, p-value = 0.527
```

```
hist(resid(m.diversity.rep_2))
```

```
summary(m.diversity.rep_2)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.457e-03 -1.632e-04  2.314e-05  1.700e-04  2.044e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.046e-02  4.515e-05 453.271   <2e-16 ***
## thetaC        1.085e+00  5.297e-03 204.913   <2e-16 ***
## rhoC          9.251e-03  6.537e-02   0.142    0.888
## tmrcaC        2.050e-02  5.178e-04  39.582   <2e-16 ***
## thetaC:tmrcaC 1.161e+00  4.967e-02  23.381   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005487 on 145 degrees of freedom
## Multiple R-squared:  0.9969, Adjusted R-squared:  0.9968
## F-statistic: 1.164e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_2, pred = thetaC, modx= tmrcaC)
```

```
g.rep_2 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_2, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_2)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_2
##         AIC       BIC   logLik
##   -1821.325 -1800.251 917.6627
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1654346
##
## Coefficients:
##                     Value  Std.Error  t-value p-value
## (Intercept)    0.0204681 0.00005329 384.0912  0.0000
## thetaC         1.0868465 0.00566891 191.7207  0.0000
## tmrcaC         0.0204136 0.00053121  38.4284  0.0000
## rhoC          -0.0060738 0.06368644  -0.0954  0.9242
## thetaC:tmrcaC  1.1274440 0.05080579  22.1912  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.005
## tmrcaC        -0.054 -0.150
## rhoC          -0.013 -0.022  0.171
## thetaC:tmrcaC -0.108 -0.049  0.491  0.125
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.58086608 -0.34056969  0.06376542  0.29976087  3.77133875
##
## Residual standard error: 0.0005405587
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_2)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.023970      1.366753      1.032349      1.320968
```

```
g.rep_2.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.200kb.rep_2, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_2.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_2
##         AIC       BIC   logLik
##   -1465.219 -1450.166 737.6095
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3137584
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0206008 0.00021211 97.12366  0.0000
## thetaC       1.1262748 0.01942595 57.97785  0.0000
## rhoC        -0.3984795 0.20018010 -1.99061  0.0484
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.002  0.001
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.31916797 -0.44159324  0.03266396  0.49987953  3.39180105
##
## Residual standard error: 0.001864311
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_2)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.200kb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
r2.inf.200kb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 2] <- anova.diversity$VarExp[4] * 100
```

2.2.3 Replicate 3

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_3/rs.pair_3.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 9791.4, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9825924
```

```
inf.lands.200kb.rep_3 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_mea
names(inf.lands.200kb.rep_3) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_3$thetaC <- inf.lands.200kb.rep_3$theta - mean(inf.lands.200kb.rep_3$theta)
inf.lands.200kb.rep_3$tmrcaC <- inf.lands.200kb.rep_3$tmrca - mean(inf.lands.200kb.rep_3$tmrca)
inf.lands.200kb.rep_3$rhoC <- inf.lands.200kb.rep_3$rho - mean(inf.lands.200kb.rep_3$rho)

inf.lands.200kb.rep_3$bin <- 1:nrow(inf.lands.200kb.rep_3)

# for merging:
inf.lands.200kb.rep_3$Replicate <- 3

m.diversity.rep_3 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep_3

plot(resid(m.diversity.rep_3)~fitted(m.diversity.rep_3))
```

```
dwtest(m.diversity.rep_3)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_3
## DW = 1.7271, p-value = 0.03626
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_3)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_3
## HMC = 0.49269, p-value = 0.46
```

```
hist(resid(m.diversity.rep_3))
```

```
summary(m.diversity.rep_3)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.723e-03 -1.975e-04 -2.090e-06  2.053e-04  2.035e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.060e-02  4.271e-05 482.307   <2e-16 ***
## thetaC        1.079e+00  4.918e-03 219.494   <2e-16 ***
## rhoC          3.612e-02  6.152e-02   0.587    0.558
## tmrcaC        2.055e-02  4.262e-04  48.218   <2e-16 ***
## thetaC:tmrcaC 1.092e+00  3.816e-02  28.629   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005196 on 145 degrees of freedom
## Multiple R-squared:  0.9974, Adjusted R-squared:  0.9973
## F-statistic: 1.387e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_3, pred = thetaC, modx= tmrcaC)
```

```
g.rep_3 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_3, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_3
##         AIC       BIC   logLik
##   -1837.259 -1816.184 925.6294
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1534496
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0206000 0.00004979 413.7691  0.0000
## thetaC        1.0808335 0.00523730 206.3723  0.0000
## tmrcaC        0.0205864 0.00043609  47.2067  0.0000
## rhoC          0.0431677 0.06142122   0.7028  0.4833
## thetaC:tmrcaC 1.0712178 0.03947149  27.1390  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.006
## tmrcaC        -0.020 -0.153
## rhoC          -0.009 -0.025  0.167
## thetaC:tmrcaC -0.101 -0.061  0.181  0.089
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.38419912 -0.35874981 -0.00849962  0.40321853  4.03279383
##
## Residual standard error: 0.0005116005
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_3)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.025025      1.081264      1.032608      1.039063
```

```
g.rep_3.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.200kb.rep_3, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_3.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_3
##         AIC       BIC   logLik
##   -1403.367 -1388.314 706.6834
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.2774686
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0207429 0.00024772 83.73477   0.000
## thetaC       1.1273552 0.02313319 48.73324   0.000
## rhoC        -0.5122430 0.25571026 -2.00322   0.047
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.001  0.003
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.57656146 -0.41478403  0.08100049  0.42358182  3.69818275
##
## Residual standard error: 0.002264567
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_3)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.200kb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.200kb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 3] <- anova.diversity$VarExp[4] * 100
```

2.2.4 Replicate 4

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_4/rs.pair_4.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 9866.4, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.982459
```

```
inf.lands.200kb.rep_4 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_mea
names(inf.lands.200kb.rep_4) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_4$thetaC <- inf.lands.200kb.rep_4$theta - mean(inf.lands.200kb.rep_4$theta)
inf.lands.200kb.rep_4$tmrcaC <- inf.lands.200kb.rep_4$tmrca - mean(inf.lands.200kb.rep_4$tmrca)
inf.lands.200kb.rep_4$rhoC <- inf.lands.200kb.rep_4$rho - mean(inf.lands.200kb.rep_4$rho)

inf.lands.200kb.rep_4$bin <- 1:nrow(inf.lands.200kb.rep_4)

# for merging:
inf.lands.200kb.rep_4$Replicate <- 4

m.diversity.rep_4 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep_4

plot(resid(m.diversity.rep_4)~fitted(m.diversity.rep_4))
```

```
dwtest(m.diversity.rep_4)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_4
## DW = 1.6283, p-value = 0.00823
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_4)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_4
## HMC = 0.54628, p-value = 0.787
```

```
hist(resid(m.diversity.rep_4))
```

```
summary(m.diversity.rep_4)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.768e-03 -2.282e-04 -4.140e-06  1.588e-04  2.724e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.063e-02  4.965e-05 415.535   <2e-16 ***
## thetaC        1.125e+00  5.855e-03 192.138   <2e-16 ***
## rhoC          6.411e-02  7.349e-02   0.872    0.384
## tmrcaC        1.975e-02  4.667e-04  42.319   <2e-16 ***
## thetaC:tmrcaC 1.146e+00  5.168e-02  22.178   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0006075 on 145 degrees of freedom
## Multiple R-squared:  0.9963, Adjusted R-squared:  0.9962
## F-statistic:  9799 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_4, pred = thetaC, modx= tmrcaC)
```

```
g.rep_4 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_4, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_4)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_4
##         AIC       BIC   logLik
##   -1792.908 -1771.834 903.4542
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1996946
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0206346 0.00006076 339.6014  0.0000
## thetaC        1.1223743 0.00626603 179.1204  0.0000
## tmrcaC        0.0198270 0.00047759  41.5143  0.0000
## rhoC          0.0706769 0.07071306   0.9995  0.3192
## thetaC:tmrcaC 1.1216837 0.05373238  20.8754  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.003
## tmrcaC        -0.012 -0.039
## rhoC          -0.002  0.018  0.182
## thetaC:tmrcaC -0.040  0.064  0.264  0.027
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.63571446 -0.38925930 -0.02298232  0.24825431  4.56473653
##
## Residual standard error: 0.0005980989
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_4)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.008269      1.115533      1.035414      1.081806
```

```
g.rep_4.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.200kb.rep_4, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_4.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_4
##         AIC       BIC   logLik
##   -1406.204 -1391.151 708.1021
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.2791937
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0206898 0.00024597 84.11461  0.0000
## thetaC       1.1309856 0.02340816 48.31587  0.0000
## rhoC        -0.4082832 0.24994231 -1.63351  0.1045
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.00
## rhoC   0.00   0.02
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -4.9242575 -0.4591282  0.0538119  0.4694468  4.4837294
##
## Residual standard error: 0.002244411
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_4)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.200kb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
r2.inf.200kb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 4] <- anova.diversity$VarExp[4] * 100
```

2.2.5 Replicate 5

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_5/rs.pair_5.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 7360.3, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9869145
```

```
inf.lands.200kb.rep_5 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_mea
names(inf.lands.200kb.rep_5) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_5$thetaC <- inf.lands.200kb.rep_5$theta - mean(inf.lands.200kb.rep_5$theta)
inf.lands.200kb.rep_5$tmrcaC <- inf.lands.200kb.rep_5$tmrca - mean(inf.lands.200kb.rep_5$tmrca)
inf.lands.200kb.rep_5$rhoC <- inf.lands.200kb.rep_5$rho - mean(inf.lands.200kb.rep_5$rho)

inf.lands.200kb.rep_5$bin <- 1:nrow(inf.lands.200kb.rep_5)

# for merging:
inf.lands.200kb.rep_5$Replicate <- 5

m.diversity.rep_5 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep_5

plot(resid(m.diversity.rep_5)~fitted(m.diversity.rep_5))
```

```
dwtest(m.diversity.rep_5)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_5
## DW = 1.473, p-value = 0.0003904
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_5)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_5
## HMC = 0.42356, p-value = 0.102
```

```
hist(resid(m.diversity.rep_5))
```

```
summary(m.diversity.rep_5)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_5)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0023718 -0.0002435  0.0000000  0.0001819  0.0019428
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.051e-02  4.143e-05 495.020   <2e-16 ***
## thetaC        1.109e+00  4.896e-03 226.607   <2e-16 ***
## rhoC          8.631e-02  6.272e-02   1.376    0.171
## tmrcaC        2.045e-02  4.162e-04  49.148   <2e-16 ***
## thetaC:tmrcaC 1.199e+00  3.896e-02  30.786   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005067 on 145 degrees of freedom
## Multiple R-squared:  0.9974, Adjusted R-squared:  0.9973
## F-statistic: 1.37e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_5, pred = thetaC, modx= tmrcaC)
```

```
g.rep_5 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_5, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_5)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_5
##         AIC       BIC   logLik
##   -1853.508 -1832.434 933.7542
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.2890306
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0205091 0.00005581 367.4517  0.0000
## thetaC        1.1068824 0.00529802 208.9238  0.0000
## tmrcaC        0.0206087 0.00042265  48.7607  0.0000
## rhoC          0.0912157 0.06030229   1.5126  0.1325
## thetaC:tmrcaC 1.1670320 0.03885841  30.0329  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.001
## tmrcaC        -0.009 -0.098
## rhoC          -0.006 -0.045  0.198
## thetaC:tmrcaC -0.038  0.025  0.256  0.166
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.69081849 -0.45570753 -0.04325154  0.33563272  4.08749773
##
## Residual standard error: 0.0005001316
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_5)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.013524      1.111015      1.057373      1.089764
```

```
g.rep_5.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.200kb.rep_5, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_5.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_5
##         AIC       BIC   logLik
##   -1410.762 -1395.709 710.3812
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.2658747
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0205716 0.00023791 86.46709  0.0000
## thetaC       1.1258771 0.02302967 48.88811  0.0000
## rhoC        -0.6236980 0.25913685 -2.40683  0.0173
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.000
## rhoC    0.000 -0.032
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -3.2105114 -0.4827096  0.0393999  0.4735940  3.9755306
##
## Residual standard error: 0.002201973
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_5)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.200kb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
r2.inf.200kb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 5] <- anova.diversity$VarExp[4] * 100
```

2.2.6 Replicate 6

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_6/rs.pair_6.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 7436.3, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9867794
```

```
inf.lands.200kb.rep_6 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_mea
names(inf.lands.200kb.rep_6) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_6$thetaC <- inf.lands.200kb.rep_6$theta - mean(inf.lands.200kb.rep_6$theta)
inf.lands.200kb.rep_6$tmrcaC <- inf.lands.200kb.rep_6$tmrca - mean(inf.lands.200kb.rep_6$tmrca)
inf.lands.200kb.rep_6$rhoC <- inf.lands.200kb.rep_6$rho - mean(inf.lands.200kb.rep_6$rho)

inf.lands.200kb.rep_6$bin <- 1:nrow(inf.lands.200kb.rep_6)

# for merging:
inf.lands.200kb.rep_6$Replicate <- 6

m.diversity.rep_6 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep_

plot(resid(m.diversity.rep_6)~fitted(m.diversity.rep_6))
```

```
dwtest(m.diversity.rep_6)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_6
## DW = 1.6828, p-value = 0.01928
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_6)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_6
## HMC = 0.48734, p-value = 0.426
```

```
hist(resid(m.diversity.rep_6))
```

```
summary(m.diversity.rep_6)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.837e-03 -2.547e-04  3.600e-07  2.011e-04  2.684e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.048e-02  4.657e-05 439.670   <2e-16 ***
## thetaC        1.085e+00  5.404e-03 200.771   <2e-16 ***
## rhoC          4.915e-02  6.439e-02   0.763    0.447
## tmrcaC        2.024e-02  5.329e-04  37.991   <2e-16 ***
## thetaC:tmrcaC 1.146e+00  4.835e-02  23.708   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000568 on 145 degrees of freedom
## Multiple R-squared:  0.9967, Adjusted R-squared:  0.9966
## F-statistic: 1.106e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_6, pred = thetaC, modx= tmrcaC)
```

```
g.rep_6 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_6, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_6)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_6
##        AIC       BIC   logLik
##   -1811.57 -1790.496 912.7851
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1730188
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0204792 0.00005540 369.6328  0.0000
## thetaC        1.0855830 0.00577130 188.1002  0.0000
## tmrcaC        0.0203230 0.00054805  37.0823  0.0000
## rhoC          0.0441005 0.06334022   0.6962  0.4874
## thetaC:tmrcaC 1.1199378 0.04982885  22.4757  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.002
## tmrcaC        -0.023 -0.109
## rhoC          -0.008 -0.005  0.340
## thetaC:tmrcaC -0.080 -0.023  0.278  0.116
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -5.126866241 -0.430648507 -0.002042255  0.325479464  4.863421007
##
## Residual standard error: 0.0005591586
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_6)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.013289      1.224420      1.132487      1.084133
```

```
g.rep_6.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.200kb.rep_6, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_6.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_6
##         AIC       BIC   logLik
##   -1450.868 -1435.815 730.4341
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3062371
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0205840 0.00022012 93.51378   0e+00
## thetaC       1.1123598 0.02006909 55.42651   0e+00
## rhoC        -0.7401946 0.19645270 -3.76780   2e-04
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.002  0.031
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.19333658 -0.46017960  0.02519769  0.43441985  3.88632753
##
## Residual standard error: 0.001950655
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_6)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.200kb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
r2.inf.200kb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 6] <- anova.diversity$VarExp[4] * 100
```

2.2.7 Replicate 7

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_7/rs.pair_7.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
## Warning in cor.test.default(sim.theta.200kb$sim, theta.200kb$sample_mean, :
## Cannot compute exact p-value with ties
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 7134.3, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9873163
```

```
inf.lands.200kb.rep_7 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_mea
names(inf.lands.200kb.rep_7) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_7$thetaC <- inf.lands.200kb.rep_7$theta - mean(inf.lands.200kb.rep_7$theta)
inf.lands.200kb.rep_7$tmrcaC <- inf.lands.200kb.rep_7$tmrca - mean(inf.lands.200kb.rep_7$tmrca)
inf.lands.200kb.rep_7$rhoC <- inf.lands.200kb.rep_7$rho - mean(inf.lands.200kb.rep_7$rho)

inf.lands.200kb.rep_7$bin <- 1:nrow(inf.lands.200kb.rep_7)

# for merging:
inf.lands.200kb.rep_7$Replicate <- 7

m.diversity.rep_7 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep_

plot(resid(m.diversity.rep_7)~fitted(m.diversity.rep_7))
```

```
dwtest(m.diversity.rep_7)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_7
## DW = 1.6673, p-value = 0.01553
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_7)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_7
## HMC = 0.43605, p-value = 0.126
```

```
hist(resid(m.diversity.rep_7))
```

```
summary(m.diversity.rep_7)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_7)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.255e-03 -2.038e-04 -2.189e-05  1.931e-04  2.342e-03
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0204967  0.0000427 480.036   <2e-16 ***
## thetaC         1.0913523  0.0049822 219.050   <2e-16 ***
## rhoC          -0.0194028  0.0611569  -0.317    0.752
## tmrcaC         0.0204993  0.0004700  43.612   <2e-16 ***
## thetaC:tmrcaC  1.1606406  0.0451718  25.694   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005216 on 145 degrees of freedom
## Multiple R-squared:  0.9972, Adjusted R-squared:  0.9971
## F-statistic: 1.281e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_7, pred = thetaC, modx= tmrcaC)
```

```
g.rep_7 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_7, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_7)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_7
##         AIC       BIC   logLik
##   -1837.873 -1816.799 925.9367
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1935696
##
## Coefficients:
##                     Value  Std.Error  t-value p-value
## (Intercept)    0.0204997 0.00005196 394.5197  0.0000
## thetaC         1.0931140 0.00535150 204.2630  0.0000
## tmrcaC         0.0203603 0.00047964  42.4489  0.0000
## rhoC          -0.0138547 0.06038474  -0.2294  0.8189
## thetaC:tmrcaC  1.1234814 0.04563517  24.6188  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.002
## tmrcaC        -0.021 -0.082
## rhoC          -0.003 -0.022  0.167
## thetaC:tmrcaC -0.060  0.039  0.329  0.078
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.36585180 -0.42077188 -0.01421357  0.38381511  4.57664531
##
## Residual standard error: 0.0005142079
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_7)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.011928      1.158318      1.029312      1.127670
```

```
g.rep_7.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.200kb.rep_7, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_7.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_7
##         AIC       BIC   logLik
##   -1444.025 -1428.971 727.0123
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.2825213
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0205796 0.00021783 94.47441  0.0000
## thetaC       1.1122060 0.02050319 54.24550  0.0000
## rhoC        -0.4602880 0.22027764 -2.08958  0.0384
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.000
## rhoC    0.002 -0.007
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -3.4561092 -0.4625831  0.1353834  0.5540699  4.0615520
##
## Residual standard error: 0.001980566
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_7)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.200kb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.200kb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 7] <- anova.diversity$VarExp[4] * 100
```

2.2.8 Replicate 8

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_8/rs.pair_8.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 6883.3, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9877626
```

```
inf.lands.200kb.rep_8 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_mea
names(inf.lands.200kb.rep_8) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_8$thetaC <- inf.lands.200kb.rep_8$theta - mean(inf.lands.200kb.rep_8$theta)
inf.lands.200kb.rep_8$tmrcaC <- inf.lands.200kb.rep_8$tmrca - mean(inf.lands.200kb.rep_8$tmrca)
inf.lands.200kb.rep_8$rhoC <- inf.lands.200kb.rep_8$rho - mean(inf.lands.200kb.rep_8$rho)

inf.lands.200kb.rep_8$bin <- 1:nrow(inf.lands.200kb.rep_8)

# for merging:
inf.lands.200kb.rep_8$Replicate <- 8

m.diversity.rep_8 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep_8

plot(resid(m.diversity.rep_8)~fitted(m.diversity.rep_8))
```

```
dwtest(m.diversity.rep_8)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_8
## DW = 1.5909, p-value = 0.004238
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_8)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_8
## HMC = 0.45552, p-value = 0.201
```

```
hist(resid(m.diversity.rep_8))
```

```
summary(m.diversity.rep_8)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_8)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.112e-03 -1.846e-04  7.690e-06  1.961e-04  1.739e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.070e-02  4.624e-05  447.65   <2e-16 ***
## thetaC        1.100e+00  5.401e-03  203.69   <2e-16 ***
## rhoC          6.479e-02  6.610e-02    0.98    0.329
## tmrcaC        2.076e-02  4.865e-04   42.68   <2e-16 ***
## thetaC:tmrcaC 1.181e+00  4.604e-02   25.64   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005639 on 145 degrees of freedom
## Multiple R-squared:  0.9968, Adjusted R-squared:  0.9967
## F-statistic: 1.124e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_8, pred = thetaC, modx= tmrcaC)
```

```
g.rep_8 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_8, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_8)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_8
##         AIC       BIC   logLik
##   -1816.857 -1795.783 915.4287
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.2325221
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0207042 0.00005859 353.3963  0.0000
## thetaC        1.0990491 0.00584866 187.9147  0.0000
## tmrcaC        0.0207188 0.00049940  41.4878  0.0000
## rhoC          0.0549165 0.06405597   0.8573  0.3927
## thetaC:tmrcaC 1.1364878 0.04655992  24.4091  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.001
## tmrcaC        -0.026 -0.121
## rhoC          -0.003 -0.003  0.185
## thetaC:tmrcaC -0.074  0.014  0.351  0.043
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.78620355 -0.37183043 -0.03601008  0.33263464  3.19188937
##
## Residual standard error: 0.0005563072
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_8)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.019264      1.202188      1.036381      1.145973
```

```
g.rep_8.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.200kb.rep_8, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_8.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_8
##         AIC       BIC   logLik
##   -1429.786 -1414.733 719.8932
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3033588
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0208111 0.00023518 88.49129  0.0000
## thetaC       1.1297636 0.02165099 52.18069  0.0000
## rhoC        -0.4075453 0.22773180 -1.78958  0.0756
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.000  0.018
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.25604299 -0.41587461  0.04359114  0.48798012  3.54428535
##
## Residual standard error: 0.002090655
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_8)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.200kb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
r2.inf.200kb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 8] <- anova.diversity$VarExp[4] * 100
```

2.2.9 Replicate 9

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_9/rs.pair_9.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 7233.3, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9871403
```

```
inf.lands.200kb.rep_9 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_mea
names(inf.lands.200kb.rep_9) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_9$thetaC <- inf.lands.200kb.rep_9$theta - mean(inf.lands.200kb.rep_9$theta)
inf.lands.200kb.rep_9$tmrcaC <- inf.lands.200kb.rep_9$tmrca - mean(inf.lands.200kb.rep_9$tmrca)
inf.lands.200kb.rep_9$rhoC <- inf.lands.200kb.rep_9$rho - mean(inf.lands.200kb.rep_9$rho)

inf.lands.200kb.rep_9$bin <- 1:nrow(inf.lands.200kb.rep_9)

# for merging:
inf.lands.200kb.rep_9$Replicate <- 9

m.diversity.rep_9 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep_9

plot(resid(m.diversity.rep_9)~fitted(m.diversity.rep_9))
```

```
dwtest(m.diversity.rep_9)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_9
## DW = 1.6772, p-value = 0.01787
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_9)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_9
## HMC = 0.38369, p-value = 0.022
```

```
hist(resid(m.diversity.rep_9))
```

```
summary(m.diversity.rep_9)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_9)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0018926 -0.0002394 -0.0000149  0.0001830  0.0030102
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.044e-02  4.788e-05 426.832   <2e-16 ***
## thetaC        1.091e+00  5.584e-03 195.371   <2e-16 ***
## rhoC          1.218e-01  7.409e-02   1.643    0.102
## tmrcaC        2.096e-02  5.758e-04  36.402   <2e-16 ***
## thetaC:tmrcaC 1.200e+00  5.571e-02  21.530   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000584 on 145 degrees of freedom
## Multiple R-squared:  0.9965, Adjusted R-squared:  0.9964
## F-statistic: 1.02e+04 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_9, pred = thetaC, modx= tmrcaC)
```

```
g.rep_9 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_9, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_9)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_9
##         AIC       BIC   logLik
##   -1804.101 -1783.026 909.0504
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.2058788
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0204394 0.00005911 345.7760  0.0000
## thetaC        1.0918715 0.00599966 181.9888  0.0000
## tmrcaC        0.0210780 0.00059395  35.4877  0.0000
## rhoC          0.1568335 0.07270121   2.1572  0.0326
## thetaC:tmrcaC 1.1467223 0.05773633  19.8614  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.006
## tmrcaC        -0.029 -0.067
## rhoC          -0.008  0.011  0.270
## thetaC:tmrcaC -0.076  0.077  0.351  0.110
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.22090386 -0.40412235 -0.03092443  0.30494856  5.19848696
##
## Residual standard error: 0.000576945
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_9)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.017102      1.228533      1.079801      1.153832
```

```
g.rep_9.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.200kb.rep_9, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_9.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_9
##         AIC       BIC  logLik
##   -1464.018 -1448.965 737.009
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3448494
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0205379 0.00022295 92.11977  0.0000
## thetaC       1.1066746 0.01936981 57.13401  0.0000
## rhoC        -0.4865608 0.21452090 -2.26813  0.0248
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.000
## rhoC   0.002  0.028
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.01675167 -0.35192121 -0.05519082  0.39502917  4.12501538
##
## Residual standard error: 0.00189327
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_9)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.200kb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
r2.inf.200kb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 9] <- anova.diversity$VarExp[4] * 100
```

2.2.10 Replicate 10

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_10/rs.pair_10.", sep = "")
pi.200kb <- read.table(paste(p, "diversity.200kb.bedgraph", sep = ""), header = T)
pi.200kb$avg <- apply(pi.200kb[4:ncol(pi.200kb)], 1, mean)
tmrca.200kb <- read.table(paste(p, "TMRCA.200kb.bedgraph", sep = ""), header = T)
tmrca.200kb$avg <- apply(tmrca.200kb[4:ncol(tmrca.200kb)], 1, mean)
rho.200kb <- read.table(paste(p, "rho.200kb.bedgraph", sep = ""), header = T)
theta.200kb <- read.table(paste(p, "theta.200kb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.200kb$sim, theta.200kb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.200kb$sim and theta.200kb$sample_mean
## S = 6266.2, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9888595
```

```
inf.lands.200kb.rep_10 <- as.data.frame(cbind(pi.200kb$avg, theta.200kb$sample_mean, rho.200kb$sample_m
names(inf.lands.200kb.rep_10) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.200kb.rep_10$thetaC <- inf.lands.200kb.rep_10$theta - mean(inf.lands.200kb.rep_10$theta)
inf.lands.200kb.rep_10$tmrcaC <- inf.lands.200kb.rep_10$tmrca - mean(inf.lands.200kb.rep_10$tmrca)
inf.lands.200kb.rep_10$rhoC <- inf.lands.200kb.rep_10$rho - mean(inf.lands.200kb.rep_10$rho)

inf.lands.200kb.rep_10$bin <- 1:nrow(inf.lands.200kb.rep_10)

# for merging:
inf.lands.200kb.rep_10$Replicate <- 10

m.diversity.rep_10 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.200kb.rep

plot(resid(m.diversity.rep_10)~fitted(m.diversity.rep_10))
```

```
dwtest(m.diversity.rep_10)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_10
## DW = 1.777, p-value = 0.06912
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_10)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_10
## HMC = 0.6193, p-value = 0.979
```

```
hist(resid(m.diversity.rep_10))
```

```
summary(m.diversity.rep_10)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.200kb.rep_10)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.938e-03 -1.748e-04 -4.971e-05  1.828e-04  2.971e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.053e-02  5.021e-05 408.838   <2e-16 ***
## thetaC        1.071e+00  5.803e-03 184.625   <2e-16 ***
## rhoC          1.995e-02  7.054e-02   0.283    0.778
## tmrcaC        2.204e-02  5.689e-04  38.738   <2e-16 ***
## thetaC:tmrcaC 1.283e+00  5.870e-02  21.866   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0006067 on 145 degrees of freedom
## Multiple R-squared:  0.9963, Adjusted R-squared:  0.9962
## F-statistic:  9827 on 4 and 145 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_10, pred = thetaC, modx= tmrcaC)
```

```
g.rep_10 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.200kb.rep_1, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_10)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep_1
##         AIC       BIC   logLik
##   -1814.128 -1793.054 914.0641
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##        Phi
## 0.08562895
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0204958 0.00004989 410.8349  0.0000
## thetaC        1.0967285 0.00562895 194.8371  0.0000
## tmrcaC        0.0210125 0.00050621  41.5092  0.0000
## rhoC          0.0219683 0.06340360   0.3465  0.7295
## thetaC:tmrcaC 1.2110957 0.05322577  22.7539  0.0000
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.001
## tmrcaC        -0.043 -0.110
## rhoC          -0.010 -0.009  0.248
## thetaC:tmrcaC -0.113  0.010  0.375  0.090
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.94983846 -0.42035716 -0.01013693  0.33115763  4.40016170
##
## Residual standard error: 0.0005481049
## Degrees of freedom: 150 total; 145 residual
```

```
vif(g.rep_10)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.015853      1.249427      1.066075      1.167458
```

```
g.rep_10.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                         data = inf.lands.200kb.rep_10, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_10.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.200kb.rep_10
##         AIC      BIC   logLik
##   -1442.284 -1427.23 726.1418
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3304014
##
## Coefficients:
##                  Value  Std.Error  t-value p-value
## (Intercept)  0.0207131 0.00023459 88.29634  0.0000
## thetaC       1.1352063 0.02089711 54.32360  0.0000
## rhoC        -0.5008572 0.21144100 -2.36878  0.0191
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.000
## rhoC    0.002 -0.030
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.75464197 -0.53844834  0.02521541  0.52547583  4.31476850
##
## Residual standard error: 0.002024422
## Degrees of freedom: 150 total; 147 residual
```

```
anova.diversity <- Anova(m.diversity.rep_10)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.200kb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
r2.inf.200kb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.inf.200kb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.inf.200kb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.inf.200kb[5, 10] <- anova.diversity$VarExp[4] * 100

r2.inf.200kb$average <- rowMeans(r2.inf.200kb)
r2.inf.200kb <- transform(r2.inf.200kb, sd=apply(r2.inf.200kb, 1, sd, na.rm = TRUE))
```

2.2.11 all replicates:

```
inf.lands.200kb.all <- rbind(inf.lands.200kb.rep_1, inf.lands.200kb.rep_2, inf.lands.200kb.rep_3, inf.la

m.div.200kb <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)*as.factor(Replicate), data = inf
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.6825, p-value = 8.31e-12
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.48172, p-value = 0.158
```

```r
hist(resid(m.div.200kb))
```

```r
m.div.200kb.2 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC)*as.factor(Replica
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC + rhoC:tmrcaC)*as

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb   51 -18158.03
## m.div.200kb.2 61 -18155.65
## m.div.200kb.3 71 -18154.31
```

```r
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC) *
##     as.factor(Replicate), data = inf.lands.200kb.all)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.837e-03 -2.188e-04 -6.580e-06  1.836e-04  3.010e-03
##
## Coefficients:
##                                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)                           2.049e-02  4.600e-05 445.509  < 2e-16 ***
## thetaC                                1.096e+00  5.439e-03 201.508  < 2e-16 ***
## rhoC                                  1.426e-02  6.373e-02   0.224 0.822950
## tmrcaC                                2.103e-02  5.029e-04  41.809  < 2e-16 ***
## as.factor(Replicate)2                -3.060e-05  6.507e-05  -0.470 0.638292
```

```
## as.factor(Replicate)3                 1.025e-04  6.504e-05   1.576 0.115287
## as.factor(Replicate)4                 1.385e-04  6.486e-05   2.136 0.032875 *
## as.factor(Replicate)5                 1.180e-05  6.487e-05   0.182 0.855637
## as.factor(Replicate)6                -1.818e-05  6.496e-05  -0.280 0.779594
## as.factor(Replicate)7                 2.154e-06  6.491e-05   0.033 0.973535
## as.factor(Replicate)8                 2.053e-04  6.496e-05   3.160 0.001611 **
## as.factor(Replicate)9                -5.912e-05  6.496e-05  -0.910 0.362922
## as.factor(Replicate)10                3.230e-05  6.526e-05   0.495 0.620766
## thetaC:tmrcaC                         1.224e+00  5.205e-02  23.507  < 2e-16 ***
## thetaC:as.factor(Replicate)2         -1.053e-02  7.664e-03  -1.374 0.169687
## thetaC:as.factor(Replicate)3         -1.650e-02  7.590e-03  -2.174 0.029886 *
## thetaC:as.factor(Replicate)4          2.909e-02  7.659e-03   3.799 0.000151 ***
## thetaC:as.factor(Replicate)5          1.351e-02  7.668e-03   1.762 0.078268 .
## thetaC:as.factor(Replicate)6         -1.096e-02  7.610e-03  -1.441 0.149866
## thetaC:as.factor(Replicate)7         -4.595e-03  7.624e-03  -0.603 0.546846
## thetaC:as.factor(Replicate)8          4.142e-03  7.634e-03   0.543 0.587471
## thetaC:as.factor(Replicate)9         -4.972e-03  7.628e-03  -0.652 0.514662
## thetaC:as.factor(Replicate)10        -2.460e-02  7.629e-03  -3.225 0.001289 **
## rhoC:as.factor(Replicate)2           -5.012e-03  9.221e-02  -0.054 0.956658
## rhoC:as.factor(Replicate)3            2.185e-02  9.191e-02   0.238 0.812089
## rhoC:as.factor(Replicate)4            4.984e-02  9.296e-02   0.536 0.591924
## rhoC:as.factor(Replicate)5            7.204e-02  9.411e-02   0.766 0.444095
## rhoC:as.factor(Replicate)6            3.489e-02  8.991e-02   0.388 0.698042
## rhoC:as.factor(Replicate)7           -3.367e-02  9.146e-02  -0.368 0.712840
## rhoC:as.factor(Replicate)8            5.053e-02  9.144e-02   0.553 0.580613
## rhoC:as.factor(Replicate)9            1.075e-01  9.538e-02   1.127 0.259957
## rhoC:as.factor(Replicate)10           5.683e-03  9.106e-02   0.062 0.950247
## tmrcaC:as.factor(Replicate)2         -5.309e-04  7.291e-04  -0.728 0.466627
## tmrcaC:as.factor(Replicate)3         -4.744e-04  6.808e-04  -0.697 0.486026
## tmrcaC:as.factor(Replicate)4         -1.276e-03  6.615e-04  -1.930 0.053851 .
## tmrcaC:as.factor(Replicate)5         -5.725e-04  6.812e-04  -0.840 0.400788
## tmrcaC:as.factor(Replicate)6         -7.813e-04  7.269e-04  -1.075 0.282629
## tmrcaC:as.factor(Replicate)7         -5.267e-04  7.121e-04  -0.740 0.459573
## tmrcaC:as.factor(Replicate)8         -2.644e-04  6.970e-04  -0.379 0.704532
## tmrcaC:as.factor(Replicate)9         -6.463e-05  7.464e-04  -0.087 0.931011
## tmrcaC:as.factor(Replicate)10         1.011e-03  7.267e-04   1.391 0.164432
## thetaC:tmrcaC:as.factor(Replicate)2  -6.224e-02  7.262e-02  -0.857 0.391526
## thetaC:tmrcaC:as.factor(Replicate)3  -1.311e-01  6.631e-02  -1.978 0.048172 *
## thetaC:tmrcaC:as.factor(Replicate)4  -7.738e-02  7.053e-02  -1.097 0.272721
## thetaC:tmrcaC:as.factor(Replicate)5  -2.416e-02  6.752e-02  -0.358 0.720504
## thetaC:tmrcaC:as.factor(Replicate)6  -7.740e-02  7.054e-02  -1.097 0.272752
## thetaC:tmrcaC:as.factor(Replicate)7  -6.291e-02  7.111e-02  -0.885 0.376449
## thetaC:tmrcaC:as.factor(Replicate)8  -4.279e-02  6.925e-02  -0.618 0.536691
## thetaC:tmrcaC:as.factor(Replicate)9  -2.404e-02  7.454e-02  -0.322 0.747141
## thetaC:tmrcaC:as.factor(Replicate)10  5.990e-02  7.509e-02   0.798 0.425121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005594 on 1450 degrees of freedom
## Multiple R-squared:  0.9968, Adjusted R-squared:  0.9967
## F-statistic:  9292 on 49 and 1450 DF,  p-value: < 2.2e-16
```

2.2.12 Plots

```
scale.3d <- function(x) sprintf("%.3f", x)

rho.plot <- as.data.frame(cbind(inf.lands.200kb.rep_1$bin,
                                sim.rho.200kb$sim,
                                inf.lands.200kb.rep_1$rho,
                                inf.lands.200kb.rep_2$rho,
                                inf.lands.200kb.rep_3$rho,
                                inf.lands.200kb.rep_4$rho,
                                inf.lands.200kb.rep_5$rho,
                                inf.lands.200kb.rep_6$rho,
                                inf.lands.200kb.rep_7$rho,
                                inf.lands.200kb.rep_8$rho,
                                inf.lands.200kb.rep_9$rho,
                                inf.lands.200kb.rep_10$rho))

rho.plot[,3:ncol(rho.plot)] <- 1.53 * rho.plot[,3:ncol(rho.plot)] # scaling because of SCRM N0 vs Ne

names(rho.plot) <- c("bin", "sim", reps)
molten.rho <- melt(rho.plot, id.vars = "bin")
rho.map.200kb <- ggplot(data = molten.rho, aes(x = bin * 200, y = value, colour = variable)) + theme_bw
rho.map.200kb <- rho.map.200kb + geom_line(data = molten.rho, aes(size = variable)) + scale_size_manual
rho.map.200kb <- rho.map.200kb + scale_color_manual(values = c("black", brewer.pal(n = 9, name = "Blues
rho.map.200kb <- rho.map.200kb + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(brea
rho.map.200kb <- rho.map.200kb + labs(title = NULL, x = NULL, y = expression(rho))
rho.map.200kb <- rho.map.200kb + theme(text = element_text(size = 20), axis.title.x = element_text(size

theta.plot <- as.data.frame(cbind(inf.lands.200kb.rep_1$bin,
                                sim.theta.200kb$sim,
                                inf.lands.200kb.rep_1$theta,
                                inf.lands.200kb.rep_2$theta,
                                inf.lands.200kb.rep_3$theta,
                                inf.lands.200kb.rep_4$theta,
                                inf.lands.200kb.rep_5$theta,
                                inf.lands.200kb.rep_6$theta,
                                inf.lands.200kb.rep_7$theta,
                                inf.lands.200kb.rep_8$theta,
                                inf.lands.200kb.rep_9$theta,
                                inf.lands.200kb.rep_10$theta))

names(theta.plot) <- c("bin", "sim", reps)
molten.theta <- melt(theta.plot, id.vars = "bin")
theta.map.200kb <- ggplot(data = molten.theta, aes(x = bin * 200, y = value, colour = variable)) + theme
theta.map.200kb <- theta.map.200kb + geom_line(data = molten.theta, aes(size = variable)) + scale_size_m
theta.map.200kb <- theta.map.200kb + scale_color_manual(values = c("black", brewer.pal(n = 9, name = "Re
theta.map.200kb <- theta.map.200kb + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(b
theta.map.200kb <- theta.map.200kb + labs(title = NULL, x = NULL, y = expression(theta))
theta.map.200kb <- theta.map.200kb + theme(text = element_text(size = 20), axis.title.x = element_text(s

plot_grid(rho.map.200kb, theta.map.200kb, nrow = 2, ncol = 1)
```

2.3 1 Mb scale

```
r2.inf.1Mb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.inf.1Mb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.inf.1Mb) <- reps

# sim landscapes
sim.rho.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims/dm.sim.rho.1e+06.txt")
sim.theta.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_coal_sims/dm.sim.theta.1e+06.txt")
sim.lands.1Mb <- as.data.frame(cbind(sim.theta.1Mb$sim, sim.rho.1Mb$sim))
names(sim.lands.1Mb) <- c("theta", "rho")
```

2.3.1 Replicate 1

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_1/rs.pair_1.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
```

```
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 94, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9790879
```

```
inf.lands.1Mb.rep_1 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmrca
names(inf.lands.1Mb.rep_1) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_1$thetaC <- inf.lands.1Mb.rep_1$theta - mean(inf.lands.1Mb.rep_1$theta)
inf.lands.1Mb.rep_1$tmrcaC <- inf.lands.1Mb.rep_1$tmrca - mean(inf.lands.1Mb.rep_1$tmrca)
inf.lands.1Mb.rep_1$rhoC <- inf.lands.1Mb.rep_1$rho - mean(inf.lands.1Mb.rep_1$rho)

inf.lands.1Mb.rep_1$bin <- 1:nrow(inf.lands.1Mb.rep_1)

# for merging:
inf.lands.1Mb.rep_1$Replicate <- 1

m.diversity.rep_1 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_1)

plot(resid(m.diversity.rep_1)~fitted(m.diversity.rep_1))
```

```
dwtest(m.diversity.rep_1)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_1
## DW = 1.9667, p-value = 0.4252
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.diversity.rep_1)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_1
## HMC = 0.46606, p-value = 0.417
```

```r
hist(resid(m.diversity.rep_1))
```

```r
summary(m.diversity.rep_1)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_1)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0009583 -0.0003251 -0.0000061  0.0002355  0.0017816
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.020625   0.000115 179.401  < 2e-16 ***
## thetaC        1.092123   0.021125  51.697  < 2e-16 ***
## rhoC          0.003099   0.351525   0.009    0.993
## tmrcaC        0.016704   0.002985   5.596 8.05e-06 ***
## thetaC:tmrcaC 0.819285   0.498125   1.645    0.113
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.0006297 on 25 degrees of freedom
## Multiple R-squared:  0.9917, Adjusted R-squared:  0.9904
## F-statistic: 749.1 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_1, pred = thetaC, modx= tmrcaC)
```

```
g.rep_1 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_1, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_1)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_1
##       AIC       BIC logLik
##   -348.56 -338.7516 181.28
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.0188108
##
## Coefficients:
##                   Value Std.Error   t-value p-value
## (Intercept)   0.0206248 0.0001171 176.16584  0.0000
## thetaC        1.0920063 0.0211898  51.53453  0.0000
```

```
## tmrcaC        0.0168279 0.0029670   5.67164  0.0000
## rhoC          0.0007886 0.3523958   0.00224  0.9982
## thetaC:tmrcaC 0.8389718 0.4973907   1.68675  0.1041
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
## thetaC        -0.001
## tmrcaC         0.001 -0.171
## rhoC           0.000 -0.021  0.376
## thetaC:tmrcaC  0.004 -0.283  0.575 -0.003
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -1.67095844 -0.56781078 -0.01108315  0.40839933  3.09944720
##
## Residual standard error: 0.0005748314
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_1)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.087501      1.898249      1.271394      1.717045
```

```
g.rep_1.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.1Mb.rep_1, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_1.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_1
##         AIC       BIC  logLik
##   -329.2361 -322.2301 169.618
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##        Phi
## -0.3222673
##
## Coefficients:
##                  Value Std.Error   t-value p-value
## (Intercept)  0.0206090 0.0001242 165.95945  0.0000
## thetaC       1.0935838 0.0263685  41.47310  0.0000
## rhoC        -0.9523384 0.3871228  -2.46004  0.0206
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.002
## rhoC   -0.001 -0.056
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.3284963 -0.5352289  0.1792210  0.6290505  1.9096098
##
```

```
## Residual standard error: 0.0008939388
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_1)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.inf.1Mb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.inf.1Mb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 1] <- anova.diversity$VarExp[4] * 100
```

2.3.2 Replicate 2

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_2/rs.pair_2.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 114, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9746385
```

```
inf.lands.1Mb.rep_2 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmrca
names(inf.lands.1Mb.rep_2) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_2$thetaC <- inf.lands.1Mb.rep_2$theta - mean(inf.lands.1Mb.rep_2$theta)
inf.lands.1Mb.rep_2$tmrcaC <- inf.lands.1Mb.rep_2$tmrca - mean(inf.lands.1Mb.rep_2$tmrca)
inf.lands.1Mb.rep_2$rhoC <- inf.lands.1Mb.rep_2$rho - mean(inf.lands.1Mb.rep_2$rho)

inf.lands.1Mb.rep_2$bin <- 1:nrow(inf.lands.1Mb.rep_2)

# for merging:
inf.lands.1Mb.rep_2$Replicate <- 2

m.diversity.rep_2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_2)

plot(resid(m.diversity.rep_2)~fitted(m.diversity.rep_2))
```

```
dwtest(m.diversity.rep_2)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_2
## DW = 1.966, p-value = 0.4212
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_2)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_2
## HMC = 0.4866, p-value = 0.461
```

```
hist(resid(m.diversity.rep_2))
```

```
summary(m.diversity.rep_2)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.091e-03 -3.697e-04 -6.640e-06  2.531e-04  1.138e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0206424  0.0001134 182.048  < 2e-16 ***
## thetaC        1.0774623  0.0197421  54.577  < 2e-16 ***
## rhoC          0.3474698  0.4535646   0.766   0.4508
## tmrcaC        0.0195388  0.0026090   7.489 7.66e-08 ***
## thetaC:tmrcaC 1.3891871  0.5019294   2.768   0.0105 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0006138 on 25 degrees of freedom
## Multiple R-squared:  0.9919, Adjusted R-squared:  0.9906
## F-statistic: 765.1 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_2, pred = thetaC, modx= tmrcaC)
```

```
g.rep_2 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_2, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_2)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_2
##         AIC       BIC   logLik
##   -350.0905 -340.2821 182.0452
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##        Phi
## 0.01325283
##
## Coefficients:
##                   Value Std.Error   t-value p-value
## (Intercept)   0.0206428 0.0001148 179.80497  0.0000
## thetaC        1.0773596 0.0197550  54.53607  0.0000
## tmrcaC        0.0195384 0.0026060   7.49753  0.0000
## rhoC          0.3372081 0.4548652   0.74134  0.4654
## thetaC:tmrcaC 1.3998089 0.5015497   2.79097  0.0099
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.015
## tmrcaC         0.078  0.012
## rhoC          -0.056  0.004  0.165
## thetaC:tmrcaC  0.150 -0.097  0.527 -0.369
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -1.95006224 -0.66519213 -0.01401553  0.44798577  2.03312007
##
## Residual standard error: 0.0005603058
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_2)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.021369      1.764385      1.468410      2.007938
```

```
g.rep_2.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.1Mb.rep_2, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_2.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_2
##         AIC       BIC   logLik
##   -318.0822 -311.0762 164.0411
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##         Phi
## -0.01297054
##
## Coefficients:
##                  Value Std.Error   t-value p-value
## (Intercept)  0.0205937 0.0001941 106.12154  0.0000
## thetaC       1.0716533 0.0343149  31.23000  0.0000
## rhoC        -0.5592219 0.6546032  -0.85429  0.4005
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.000
## rhoC    0.000 -0.081
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.0512845 -0.6085269  0.1849871  0.5460123  1.6511980
##
## Residual standard error: 0.00102108
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_2)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.1Mb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.inf.1Mb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 2] <- anova.diversity$VarExp[4] * 100
```

2.3.3 Replicate 3

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_3/rs.pair_3.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 86, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9808676
```

```
inf.lands.1Mb.rep_3 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmrca
names(inf.lands.1Mb.rep_3) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_3$thetaC <- inf.lands.1Mb.rep_3$theta - mean(inf.lands.1Mb.rep_3$theta)
inf.lands.1Mb.rep_3$tmrcaC <- inf.lands.1Mb.rep_3$tmrca - mean(inf.lands.1Mb.rep_3$tmrca)
inf.lands.1Mb.rep_3$rhoC <- inf.lands.1Mb.rep_3$rho - mean(inf.lands.1Mb.rep_3$rho)

inf.lands.1Mb.rep_3$bin <- 1:nrow(inf.lands.1Mb.rep_3)

# for merging:
inf.lands.1Mb.rep_3$Replicate <- 3

m.diversity.rep_3 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_3)

plot(resid(m.diversity.rep_3)~fitted(m.diversity.rep_3))
```

```
dwtest(m.diversity.rep_3)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_3
## DW = 1.3502, p-value = 0.02515
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_3)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_3
## HMC = 0.51975, p-value = 0.575
```

```
hist(resid(m.diversity.rep_3))
```

```
summary(m.diversity.rep_3)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.317e-03 -3.626e-04  1.232e-05  2.530e-04  1.587e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0206899  0.0001206 171.565  < 2e-16 ***
## thetaC        1.0723605  0.0213849  50.146  < 2e-16 ***
## rhoC          0.1501119  0.3861502   0.389  0.70076
## tmrcaC        0.0209333  0.0025600   8.177 1.57e-08 ***
## thetaC:tmrcaC 1.8224106  0.5068011   3.596  0.00139 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0006567 on 25 degrees of freedom
## Multiple R-squared:  0.9915, Adjusted R-squared:  0.9901
## F-statistic: 728.3 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_3, pred = thetaC, modx= tmrcaC)
```

```
g.rep_3 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_3, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_3
##         AIC       BIC   logLik
##   -349.3777 -339.5693 181.6889
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.3277574
##
## Coefficients:
##                    Value Std.Error   t-value p-value
## (Intercept)   0.0206863 0.0001662 124.43077  0.0000
## thetaC        1.0693355 0.0206663  51.74291  0.0000
## tmrcaC        0.0216118 0.0022595   9.56486  0.0000
## rhoC          0.0916173 0.3604146   0.25420  0.8014
## thetaC:tmrcaC 1.9813808 0.4596404   4.31072  0.0002
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.024
## tmrcaC        -0.051 -0.272
## rhoC           0.026  0.090  0.062
## thetaC:tmrcaC -0.081 -0.284  0.478 -0.312
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.06120732 -0.59725161  0.02282195  0.36381078  2.61148857
##
## Residual standard error: 0.0005989662
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_3)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.119139      1.425723      1.185890      1.564342
```

```
g.rep_3.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.1Mb.rep_3, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_3.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_3
##         AIC       BIC   logLik
##   -311.0899 -304.0839 160.5449
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##         Phi
## -0.06226692
##
## Coefficients:
##                  Value Std.Error  t-value p-value
## (Intercept)  0.0207330 0.0002082 99.56794  0.0000
## thetaC       1.1012864 0.0377084 29.20532  0.0000
## rhoC        -0.3177321 0.6160273 -0.51578  0.6102
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.000
## rhoC   -0.001 -0.036
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.20029757 -0.55512347  0.03085907  0.50826091  1.89284077
##
## Residual standard error: 0.00114935
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_3)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.1Mb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.inf.1Mb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 3] <- anova.diversity$VarExp[4] * 100
```

2.3.4 Replicate 4

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_4/rs.pair_4.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 80, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9822024
```

```
inf.lands.1Mb.rep_4 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmrca
names(inf.lands.1Mb.rep_4) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_4$thetaC <- inf.lands.1Mb.rep_4$theta - mean(inf.lands.1Mb.rep_4$theta)
inf.lands.1Mb.rep_4$tmrcaC <- inf.lands.1Mb.rep_4$tmrca - mean(inf.lands.1Mb.rep_4$tmrca)
inf.lands.1Mb.rep_4$rhoC <- inf.lands.1Mb.rep_4$rho - mean(inf.lands.1Mb.rep_4$rho)

inf.lands.1Mb.rep_4$bin <- 1:nrow(inf.lands.1Mb.rep_4)

# for merging:
inf.lands.1Mb.rep_4$Replicate <- 4

m.diversity.rep_4 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_4)

plot(resid(m.diversity.rep_4)~fitted(m.diversity.rep_4))
```

```
dwtest(m.diversity.rep_4)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_4
## DW = 2.0871, p-value = 0.5863
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_4)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_4
## HMC = 0.42593, p-value = 0.305
```

```
hist(resid(m.diversity.rep_4))
```

```
summary(m.diversity.rep_4)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.950e-03 -3.397e-04 -2.417e-05  4.190e-04  1.645e-03
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0206840  0.0001284 161.149  < 2e-16 ***
## thetaC         1.1134126  0.0224842  49.520  < 2e-16 ***
## rhoC          -0.0339570  0.4581591  -0.074    0.942
## tmrcaC         0.0178545  0.0033294   5.363 1.46e-05 ***
## thetaC:tmrcaC  1.0646819  0.6921958   1.538    0.137
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000703 on 25 degrees of freedom
## Multiple R-squared:  0.9901, Adjusted R-squared:  0.9886
## F-statistic: 628.1 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_4, pred = thetaC, modx= tmrcaC)
```

```
g.rep_4 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_4, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_4)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_4
##         AIC       BIC   logLik
##   -342.0193 -332.2109 178.0096
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##         Phi
## -0.05853672
##
## Coefficients:
##                     Value Std.Error   t-value p-value
## (Intercept)    0.0206839 0.0001213 170.49504  0.0000
## thetaC         1.1137846 0.0222646  50.02482  0.0000
## tmrcaC         0.0173804 0.0033848   5.13488  0.0000
## rhoC          -0.0467801 0.4513716  -0.10364  0.9183
## thetaC:tmrcaC  0.9716766 0.6991874   1.38972  0.1769
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.001
## tmrcaC        -0.003 -0.036
## rhoC           0.002  0.016  0.215
## thetaC:tmrcaC -0.006 -0.054  0.653 -0.181
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.03132272 -0.49951188 -0.03148496  0.64957883  2.58514901
##
## Residual standard error: 0.0006419926
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_4)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.002978      2.181651      1.293335      2.154008
```

```
g.rep_4.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.1Mb.rep_4, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_4.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_4
##         AIC       BIC   logLik
##   -323.3904 -316.3844 166.6952
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##        Phi
## -0.3430849
##
## Coefficients:
##                  Value Std.Error   t-value p-value
## (Intercept)  0.0206798 0.0001348 153.42630  0.0000
## thetaC       1.1175389 0.0287817  38.82807  0.0000
## rhoC        -1.1557791 0.4865494  -2.37546  0.0249
##
##  Correlation:
##        (Intr) thetaC
## thetaC 0.004
## rhoC   0.003  0.002
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.3370913 -0.2969692  0.1671986  0.5302828  2.0808719
##
## Residual standard error: 0.0009928647
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_4)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.1Mb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.inf.1Mb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 4] <- anova.diversity$VarExp[4] * 100
```

2.3.5 Replicate 5

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_5/rs.pair_5.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 76, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9830923
```

```
inf.lands.1Mb.rep_5 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmrca
names(inf.lands.1Mb.rep_5) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_5$thetaC <- inf.lands.1Mb.rep_5$theta - mean(inf.lands.1Mb.rep_5$theta)
inf.lands.1Mb.rep_5$tmrcaC <- inf.lands.1Mb.rep_5$tmrca - mean(inf.lands.1Mb.rep_5$tmrca)
inf.lands.1Mb.rep_5$rhoC <- inf.lands.1Mb.rep_5$rho - mean(inf.lands.1Mb.rep_5$rho)

inf.lands.1Mb.rep_5$bin <- 1:nrow(inf.lands.1Mb.rep_5)

# for merging:
inf.lands.1Mb.rep_5$Replicate <- 5

m.diversity.rep_5 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_5)

plot(resid(m.diversity.rep_5)~fitted(m.diversity.rep_5))
```

```
dwtest(m.diversity.rep_5)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_5
## DW = 1.8381, p-value = 0.2796
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_5)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_5
## HMC = 0.49429, p-value = 0.482
```

```
hist(resid(m.diversity.rep_5))
```

```
summary(m.diversity.rep_5)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_5)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.367e-03 -3.078e-04  4.773e-05  2.499e-04  1.494e-03
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0205901  0.0001215 169.475  < 2e-16 ***
## thetaC         1.1075044  0.0217775  50.855  < 2e-16 ***
## rhoC          -0.1291914  0.4158082  -0.311   0.7586
## tmrcaC         0.0174135  0.0033337   5.224 2.09e-05 ***
## thetaC:tmrcaC  1.4750759  0.6476973   2.277   0.0316 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0006642 on 25 degrees of freedom
## Multiple R-squared:  0.9905, Adjusted R-squared:  0.989
## F-statistic: 650.9 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_5, pred = thetaC, modx= tmrcaC)
```

```
g.rep_5 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_5, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_5)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_5
##         AIC      BIC   logLik
##   -345.5754 -335.767 179.7877
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##        Phi
## 0.09742974
##
## Coefficients:
##                     Value Std.Error   t-value p-value
## (Intercept)    0.0205933 0.0001336 154.16894  0.0000
## thetaC         1.1072954 0.0218882  50.58862  0.0000
## tmrcaC         0.0179569 0.0032649   5.50000  0.0000
## rhoC          -0.1181976 0.4172436  -0.28328  0.7793
## thetaC:tmrcaC  1.6258812 0.6398768   2.54093  0.0176
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.002
## tmrcaC         0.033  0.055
## rhoC          -0.002  0.016  0.365
## thetaC:tmrcaC  0.058  0.056  0.609 -0.028
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.2002408 -0.4979867  0.0723577  0.3822209  2.4009759
##
## Residual standard error: 0.0006068364
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_5)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.003874      2.071251      1.303671      1.797134
```

```
g.rep_5.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.1Mb.rep_5, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_5.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_5
##         AIC       BIC   logLik
##   -327.0266 -320.0206 168.5133
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## -0.151117
##
## Coefficients:
##                  Value Std.Error   t-value p-value
## (Intercept)  0.0205697 0.0001476 139.31706  0.0000
## thetaC       1.0982569 0.0293493  37.42021  0.0000
## rhoC        -1.2150813 0.4706686  -2.58161  0.0156
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.001
## rhoC   -0.002 -0.018
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.9762372 -0.5677260  0.1592677  0.6160406  2.4429343
##
## Residual standard error: 0.0008894714
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_5)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.1Mb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.inf.1Mb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 5] <- anova.diversity$VarExp[4] * 100
```

2.3.6 Replicate 6

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_6/rs.pair_6.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 84, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9813126
```

```
inf.lands.1Mb.rep_6 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmrca
names(inf.lands.1Mb.rep_6) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_6$thetaC <- inf.lands.1Mb.rep_6$theta - mean(inf.lands.1Mb.rep_6$theta)
inf.lands.1Mb.rep_6$tmrcaC <- inf.lands.1Mb.rep_6$tmrca - mean(inf.lands.1Mb.rep_6$tmrca)
inf.lands.1Mb.rep_6$rhoC <- inf.lands.1Mb.rep_6$rho - mean(inf.lands.1Mb.rep_6$rho)

inf.lands.1Mb.rep_6$bin <- 1:nrow(inf.lands.1Mb.rep_6)

# for merging:
inf.lands.1Mb.rep_6$Replicate <- 6

m.diversity.rep_6 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_6)

plot(resid(m.diversity.rep_6)~fitted(m.diversity.rep_6))
```

```
dwtest(m.diversity.rep_6)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_6
## DW = 1.7678, p-value = 0.2162
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_6)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_6
## HMC = 0.38206, p-value = 0.207
```

```
hist(resid(m.diversity.rep_6))
```

```
summary(m.diversity.rep_6)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.277e-03 -4.527e-04  7.960e-06  3.734e-04  1.290e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0205897  0.0001193 172.524  < 2e-16 ***
## thetaC        1.0872234  0.0212487  51.167  < 2e-16 ***
## rhoC          0.3998838  0.3697035   1.082   0.2897
## tmrcaC        0.0196732  0.0028089   7.004 2.43e-07 ***
## thetaC:tmrcaC 1.1711909  0.4979919   2.352   0.0269 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0006531 on 25 degrees of freedom
## Multiple R-squared:  0.9911, Adjusted R-squared:  0.9897
## F-statistic: 694.4 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_6, pred = thetaC, modx= tmrcaC)
```

```
g.rep_6 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_6, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_6)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_6
##         AIC       BIC   logLik
##   -346.8212 -337.0128 180.4106
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1358209
##
## Coefficients:
##                    Value Std.Error   t-value p-value
## (Intercept)   0.0205924 0.0001363 151.13216  0.0000
## thetaC        1.0860890 0.0213756  50.80966  0.0000
## tmrcaC        0.0198458 0.0027501   7.21651  0.0000
## rhoC          0.3466959 0.3730361   0.92939  0.3616
## thetaC:tmrcaC 1.3137006 0.4935342   2.66182  0.0134
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.007
## tmrcaC         0.000 -0.065
## rhoC          -0.010  0.040  0.485
## thetaC:tmrcaC  0.033 -0.186  0.336 -0.157
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.09674866 -0.76764130  0.02681106  0.64010637  2.02153252
##
## Residual standard error: 0.0005969812
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_6)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.036142      1.691913      1.539768      1.360141
```

```
g.rep_6.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.1Mb.rep_6, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_6.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_6
##         AIC       BIC   logLik
##   -317.8306 -310.8246 163.9153
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##         Phi
## -0.04598612
##
## Coefficients:
##                  Value Std.Error   t-value p-value
## (Intercept)  0.0205747 0.0001889 108.91122  0.0000
## thetaC       1.0881740 0.0343243  31.70273  0.0000
## rhoC        -0.8461274 0.4782870  -1.76908  0.0882
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.000
## rhoC   -0.001 -0.001
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.7855991 -0.6320726  0.1751760  0.5604975  2.9148885
##
## Residual standard error: 0.001026338
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_6)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.1Mb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.inf.1Mb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 6] <- anova.diversity$VarExp[4] * 100
```

2.3.7 Replicate 7

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_7/rs.pair_7.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 112, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9750834
```

```
inf.lands.1Mb.rep_7 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmrca
names(inf.lands.1Mb.rep_7) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_7$thetaC <- inf.lands.1Mb.rep_7$theta - mean(inf.lands.1Mb.rep_7$theta)
inf.lands.1Mb.rep_7$tmrcaC <- inf.lands.1Mb.rep_7$tmrca - mean(inf.lands.1Mb.rep_7$tmrca)
inf.lands.1Mb.rep_7$rhoC <- inf.lands.1Mb.rep_7$rho - mean(inf.lands.1Mb.rep_7$rho)

inf.lands.1Mb.rep_7$bin <- 1:nrow(inf.lands.1Mb.rep_7)

# for merging:
inf.lands.1Mb.rep_7$Replicate <- 7

m.diversity.rep_7 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_7)

plot(resid(m.diversity.rep_7)~fitted(m.diversity.rep_7))
```

```
dwtest(m.diversity.rep_7)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_7
## DW = 1.6938, p-value = 0.1634
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_7)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_7
## HMC = 0.40789, p-value = 0.216
```

```
hist(resid(m.diversity.rep_7))
```

```
summary(m.diversity.rep_7)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_7)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0010431 -0.0003022 -0.0000846  0.0002404  0.0009636
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0206041  0.0001001 205.827  < 2e-16 ***
## thetaC        1.0776129  0.0174765  61.661  < 2e-16 ***
## rhoC          0.0680221  0.3366597   0.202 0.841513
## tmrcaC        0.0238615  0.0027851   8.568 6.58e-09 ***
## thetaC:tmrcaC 2.0280640  0.4774338   4.248 0.000262 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0005471 on 25 degrees of freedom
## Multiple R-squared:  0.9937, Adjusted R-squared:  0.9927
## F-statistic: 992.7 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_7, pred = thetaC, modx= tmrcaC)
```

```
g.rep_7 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_7, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_7)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_7
##         AIC       BIC   logLik
##   -357.7791 -347.9708 185.8896
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1726746
##
## Coefficients:
##                     Value Std.Error   t-value p-value
## (Intercept)    0.0206041 0.0001184 174.02851  0.0000
## thetaC         1.0780510 0.0175352  61.47931  0.0000
## tmrcaC         0.0233774 0.0026923   8.68312  0.0000
## rhoC          -0.0447197 0.3354118  -0.13333  0.8950
## thetaC:tmrcaC  2.0331470 0.4572788   4.44619  0.0002
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.005
## tmrcaC         0.022 -0.097
## rhoC          -0.009 -0.074  0.228
## thetaC:tmrcaC  0.046 -0.103  0.586 -0.198
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.1497675 -0.5506932 -0.1283266  0.4700014  1.9723133
##
## Residual standard error: 0.0005001377
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_7)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.020214      1.873531      1.290005      1.862508
```

```
g.rep_7.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.1Mb.rep_7, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_7.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_7
##         AIC       BIC  logLik
##   -319.9699 -312.9639 164.985
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1179429
##
## Coefficients:
##                  Value Std.Error  t-value p-value
## (Intercept)  0.0205818 0.0002149 95.79213  0.0000
## thetaC       1.0897560 0.0336214 32.41258  0.0000
## rhoC        -0.8110199 0.5672255 -1.42980  0.1642
##
##  Correlation:
##        (Intr) thetaC
## thetaC -0.001
## rhoC    0.002 -0.097
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.37431230 -0.29863999  0.01070772  0.55844318  2.30950358
##
## Residual standard error: 0.0009960966
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_7)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.1Mb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.inf.1Mb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 7] <- anova.diversity$VarExp[4] * 100
```

2.3.8 Replicate 8

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_8/rs.pair_8.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 80, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9822024
```

```
inf.lands.1Mb.rep_8 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmrca
names(inf.lands.1Mb.rep_8) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_8$thetaC <- inf.lands.1Mb.rep_8$theta - mean(inf.lands.1Mb.rep_8$theta)
inf.lands.1Mb.rep_8$tmrcaC <- inf.lands.1Mb.rep_8$tmrca - mean(inf.lands.1Mb.rep_8$tmrca)
inf.lands.1Mb.rep_8$rhoC <- inf.lands.1Mb.rep_8$rho - mean(inf.lands.1Mb.rep_8$rho)

inf.lands.1Mb.rep_8$bin <- 1:nrow(inf.lands.1Mb.rep_8)

# for merging:
inf.lands.1Mb.rep_8$Replicate <- 8

m.diversity.rep_8 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_8)

plot(resid(m.diversity.rep_8)~fitted(m.diversity.rep_8))
```

```
dwtest(m.diversity.rep_8)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_8
## DW = 1.6816, p-value = 0.17
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_8)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_8
## HMC = 0.45613, p-value = 0.356
```

```
hist(resid(m.diversity.rep_8))
```

```
summary(m.diversity.rep_8)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_8)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.226e-03 -3.903e-04  2.222e-05  3.255e-04  1.477e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0208365  0.0001224 170.170  < 2e-16 ***
## thetaC        1.0871732  0.0214774  50.620  < 2e-16 ***
## rhoC          0.3643104  0.4059648   0.897   0.3781
## tmrcaC        0.0197018  0.0033310   5.915 3.58e-06 ***
## thetaC:tmrcaC 1.2143037  0.5213795   2.329   0.0282 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0006676 on 25 degrees of freedom
## Multiple R-squared:  0.9906, Adjusted R-squared:  0.9891
## F-statistic: 660.6 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_8, pred = thetaC, modx= tmrcaC)
```

```
g.rep_8 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_8, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_8)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_8
##         AIC       BIC   logLik
##   -345.8378 -336.0294 179.9189
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1655662
##
## Coefficients:
##                    Value Std.Error   t-value p-value
## (Intercept)   0.0208387 0.0001435 145.21451  0.0000
## thetaC        1.0841156 0.0213728  50.72412  0.0000
## tmrcaC        0.0201746 0.0030896   6.52983  0.0000
## rhoC          0.3427329 0.4152276   0.82541  0.4169
## thetaC:tmrcaC 1.3177433 0.4885467   2.69727  0.0123
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.012
## tmrcaC         0.046 -0.065
## rhoC          -0.024  0.082  0.056
## thetaC:tmrcaC  0.071 -0.140  0.721 -0.279
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -1.96269837 -0.63215676  0.03038078  0.52605973  2.37801009
##
## Residual standard error: 0.0006095428
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_8)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.023443      2.452808      1.275892      2.678721
```

```
g.rep_8.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.1Mb.rep_8, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_8.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_8
##         AIC       BIC   logLik
##   -319.0793 -312.0733 164.5397
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##        Phi
## -0.1415906
##
## Coefficients:
##                  Value Std.Error   t-value p-value
## (Intercept)  0.0208035 0.0001699 122.42339  0.0000
## thetaC       1.0823875 0.0328438  32.95562  0.0000
## rhoC        -0.5688595 0.5260339  -1.08141  0.2891
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.001
## rhoC    0.001 -0.012
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -1.87467372 -0.74418645  0.07635072  0.53300754  2.76001890
##
## Residual standard error: 0.001014048
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_8)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.1Mb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.inf.1Mb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 8] <- anova.diversity$VarExp[4] * 100
```

2.3.9 Replicate 9

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_9/rs.pair_9.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 84, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9813126
```

```
inf.lands.1Mb.rep_9 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmrca
names(inf.lands.1Mb.rep_9) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_9$thetaC <- inf.lands.1Mb.rep_9$theta - mean(inf.lands.1Mb.rep_9$theta)
inf.lands.1Mb.rep_9$tmrcaC <- inf.lands.1Mb.rep_9$tmrca - mean(inf.lands.1Mb.rep_9$tmrca)
inf.lands.1Mb.rep_9$rhoC <- inf.lands.1Mb.rep_9$rho - mean(inf.lands.1Mb.rep_9$rho)

inf.lands.1Mb.rep_9$bin <- 1:nrow(inf.lands.1Mb.rep_9)

# for merging:
inf.lands.1Mb.rep_9$Replicate <- 9

m.diversity.rep_9 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_9)

plot(resid(m.diversity.rep_9)~fitted(m.diversity.rep_9))
```

```
dwtest(m.diversity.rep_9)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_9
## DW = 1.686, p-value = 0.1501
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_9)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_9
## HMC = 0.49123, p-value = 0.476
```

```
hist(resid(m.diversity.rep_9))
```

```
summary(m.diversity.rep_9)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_9)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.302e-03 -3.910e-04 -8.900e-07  3.164e-04  1.317e-03
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0205233  0.0001107 185.445  < 2e-16 ***
## thetaC        1.0737054  0.0195206  55.004  < 2e-16 ***
## rhoC          0.1883059  0.3756234   0.501   0.6205
## tmrcaC        0.0183079  0.0027965   6.547 7.39e-07 ***
## thetaC:tmrcaC 0.9925142  0.4598859   2.158   0.0407 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0006061 on 25 degrees of freedom
## Multiple R-squared:  0.9922, Adjusted R-squared:  0.991
## F-statistic: 799.9 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_9, pred = thetaC, modx= tmrcaC)
```

```
g.rep_9 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_9, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_9)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_9
##         AIC       BIC   logLik
##   -351.6229 -341.8146 182.8115
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.1633414
##
## Coefficients:
##                    Value Std.Error   t-value p-value
## (Intercept)   0.0205240 0.0001297 158.24435  0.0000
## thetaC        1.0696841 0.0194352  55.03845  0.0000
## tmrcaC        0.0185521 0.0026328   7.04662  0.0000
## rhoC          0.2186371 0.3777124   0.57885  0.5679
## thetaC:tmrcaC 1.0045124 0.4613387   2.17739  0.0391
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC         0.004
## tmrcaC        -0.021 -0.112
## rhoC           0.002  0.012  0.321
## thetaC:tmrcaC -0.022 -0.171  0.296 -0.217
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -2.277428606 -0.752076800 -0.002716583  0.581067342  2.423831154
##
## Residual standard error: 0.0005533147
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_9)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.034584      1.327310      1.266322      1.268614
```

```
g.rep_9.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                        data = inf.lands.1Mb.rep_9, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_9.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_9
##         AIC       BIC   logLik
##   -325.8718 -318.8658 167.9359
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##        Phi
## -0.2006686
##
## Coefficients:
##                  Value Std.Error   t-value p-value
## (Intercept)  0.0205144 0.0001444 142.03128  0.0000
## thetaC       1.0817758 0.0289649  37.34776  0.0000
## rhoC        -0.9433213 0.4654231  -2.02680  0.0527
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.001
## rhoC   -0.004 -0.041
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.57578335 -0.68638049 -0.02735148  0.65238571  1.96492871
##
## Residual standard error: 0.0009146795
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_9)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.1Mb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.inf.1Mb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 9] <- anova.diversity$VarExp[4] * 100
```

2.3.10 Replicate 10

```
p <- paste("dm_data/dm_chr_maps/2L/dm_coal_sims/rep_10/rs.pair_10.", sep = "")
pi.1Mb <- read.table(paste(p, "diversity.1Mb.bedgraph", sep = ""), header = T)
pi.1Mb$avg <- apply(pi.1Mb[4:ncol(pi.1Mb)], 1, mean)
tmrca.1Mb <- read.table(paste(p, "TMRCA.1Mb.bedgraph", sep = ""), header = T)
tmrca.1Mb$avg <- apply(tmrca.1Mb[4:ncol(tmrca.1Mb)], 1, mean)
rho.1Mb <- read.table(paste(p, "rho.1Mb.bedgraph", sep = ""), header = T)
theta.1Mb <- read.table(paste(p, "theta.1Mb.bedgraph", sep = ""), header = T)

cor.test(sim.theta.1Mb$sim, theta.1Mb$sample_mean, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  sim.theta.1Mb$sim and theta.1Mb$sample_mean
## S = 70, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9844271
```

```
inf.lands.1Mb.rep_10 <- as.data.frame(cbind(pi.1Mb$avg, theta.1Mb$sample_mean, rho.1Mb$sample_mean, tmr
names(inf.lands.1Mb.rep_10) <- c("diversity", "theta", "rho", "tmrca")

# centering
inf.lands.1Mb.rep_10$thetaC <- inf.lands.1Mb.rep_10$theta - mean(inf.lands.1Mb.rep_10$theta)
inf.lands.1Mb.rep_10$tmrcaC <- inf.lands.1Mb.rep_10$tmrca - mean(inf.lands.1Mb.rep_10$tmrca)
inf.lands.1Mb.rep_10$rhoC <- inf.lands.1Mb.rep_10$rho - mean(inf.lands.1Mb.rep_10$rho)

inf.lands.1Mb.rep_10$bin <- 1:nrow(inf.lands.1Mb.rep_10)

# for merging:
inf.lands.1Mb.rep_10$Replicate <- 10

m.diversity.rep_10 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC*tmrcaC, data = inf.lands.1Mb.rep_1

plot(resid(m.diversity.rep_10)~fitted(m.diversity.rep_10))
```

```
dwtest(m.diversity.rep_10)
```

```
##
##  Durbin-Watson test
##
## data:  m.diversity.rep_10
## DW = 1.5209, p-value = 0.06939
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.diversity.rep_10)
```

```
##
##  Harrison-McCabe test
##
## data:  m.diversity.rep_10
## HMC = 0.59687, p-value = 0.77
```

```
hist(resid(m.diversity.rep_10))
```

```
summary(m.diversity.rep_10)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC * tmrcaC,
##     data = inf.lands.1Mb.rep_10)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0011734 -0.0004527  0.0000019  0.0001794  0.0017898
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0207164  0.0001292 160.343  < 2e-16 ***
## thetaC        1.0619903  0.0218331  48.641  < 2e-16 ***
## rhoC          0.0656389  0.4323685   0.152   0.8806
## tmrcaC        0.0189583  0.0031915   5.940 3.35e-06 ***
## thetaC:tmrcaC 1.0582401  0.5049627   2.096   0.0464 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0007071 on 25 degrees of freedom
## Multiple R-squared:  0.9898, Adjusted R-squared:  0.9882
## F-statistic: 605.8 on 4 and 25 DF,  p-value: < 2.2e-16
```

```
interact_plot(m.diversity.rep_10, pred = thetaC, modx= tmrcaC)
```

```
g.rep_10 <- gls(diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC,
               data = inf.lands.1Mb.rep_1, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_10)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + tmrcaC + rhoC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep_1
##       AIC       BIC logLik
##   -348.56 -338.7516 181.28
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
## 0.0188108
##
## Coefficients:
##                   Value Std.Error   t-value p-value
## (Intercept)   0.0206248 0.0001171 176.16584  0.0000
## thetaC        1.0920063 0.0211898  51.53453  0.0000
## tmrcaC        0.0168279 0.0029670   5.67164  0.0000
## rhoC          0.0007886 0.3523958   0.00224  0.9982
## thetaC:tmrcaC 0.8389718 0.4973907   1.68675  0.1041
##
##  Correlation:
##               (Intr) thetaC tmrcaC rhoC
```

```
## thetaC        -0.001
## tmrcaC         0.001 -0.171
## rhoC           0.000 -0.021  0.376
## thetaC:tmrcaC  0.004 -0.283  0.575 -0.003
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -1.67095844 -0.56781078 -0.01108315  0.40839933  3.09944720
##
## Residual standard error: 0.0005748314
## Degrees of freedom: 30 total; 25 residual
```

```
vif(g.rep_10)
```

```
##        thetaC        tmrcaC          rhoC thetaC:tmrcaC
##      1.087501      1.898249      1.271394      1.717045
```

```
g.rep_10.no.tmrca <- gls(diversity ~ thetaC + rhoC,
                         data = inf.lands.1Mb.rep_10, cor = corAR1(0, ~bin), method = "ML")

summary(g.rep_10.no.tmrca)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC
##   Data: inf.lands.1Mb.rep_10
##         AIC       BIC   logLik
##   -316.9544 -309.9484 163.4772
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##         Phi
## -0.01181332
##
## Coefficients:
##                 Value Std.Error   t-value p-value
## (Intercept)  0.020705 0.0001980 104.59296  0.0000
## thetaC       1.062802 0.0337999  31.44392  0.0000
## rhoC        -0.727976 0.5917022  -1.23031  0.2292
##
##  Correlation:
##        (Intr) thetaC
## thetaC  0.000
## rhoC    0.000 -0.106
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.05588562 -0.66664272 -0.06688445  0.53908458  1.87562405
##
## Residual standard error: 0.001040441
## Degrees of freedom: 30 total; 27 residual
```

```
anova.diversity <- Anova(m.diversity.rep_10)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.inf.1Mb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.inf.1Mb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.inf.1Mb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.inf.1Mb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.inf.1Mb[5, 10] <- anova.diversity$VarExp[4] * 100

r2.inf.1Mb$average <- rowMeans(r2.inf.1Mb)
r2.inf.1Mb <- transform(r2.inf.1Mb, sd=apply(r2.inf.1Mb, 1, sd, na.rm = TRUE))
```

2.3.11 all replicates:

```
inf.lands.1Mb.all <- rbind(inf.lands.1Mb.rep_1, inf.lands.1Mb.rep_2, inf.lands.1Mb.rep_3, inf.lands.1Mb

m.div.1Mb <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)*as.factor(Replicate), data = inf.la
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.7602, p-value = 0.001541
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.51208, p-value = 0.606
```

```r
hist(resid(m.div.1Mb))
```

```r
m.div.1Mb.2 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC)*as.factor(Replicate
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC + thetaC:rhoC + rhoC:tmrcaC)*as.fa

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb   51 -3507.758
## m.div.1Mb.2 61 -3522.377
## m.div.1Mb.3 71 -3504.645
```

```r
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC) *
##     as.factor(Replicate), data = inf.lands.1Mb.all)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0019505 -0.0003895  0.0000113  0.0002755  0.0017898
##
## Coefficients:
##                                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)                           2.062e-02  1.180e-04 174.750  < 2e-16 ***
## thetaC                                1.092e+00  2.169e-02  50.357  < 2e-16 ***
## rhoC                                  3.099e-03  3.609e-01   0.009    0.993
## tmrcaC                                1.670e-02  3.065e-03   5.451  1.2e-07 ***
## as.factor(Replicate)2                 1.783e-05  1.679e-04   0.106    0.915
```

```
## as.factor(Replicate)3                 6.531e-05  1.674e-04   0.390    0.697
## as.factor(Replicate)4                 5.942e-05  1.669e-04   0.356    0.722
## as.factor(Replicate)5                -3.449e-05  1.671e-04  -0.206    0.837
## as.factor(Replicate)6                -3.480e-05  1.670e-04  -0.208    0.835
## as.factor(Replicate)7                -2.047e-05  1.671e-04  -0.123    0.903
## as.factor(Replicate)8                 2.120e-04  1.673e-04   1.267    0.206
## as.factor(Replicate)9                -1.013e-04  1.669e-04  -0.607    0.545
## as.factor(Replicate)10                9.189e-05  1.670e-04   0.550    0.583
## thetaC:tmrcaC                         8.193e-01  5.114e-01   1.602    0.110
## thetaC:as.factor(Replicate)2         -1.466e-02  3.004e-02  -0.488    0.626
## thetaC:as.factor(Replicate)3         -1.976e-02  3.022e-02  -0.654    0.514
## thetaC:as.factor(Replicate)4          2.129e-02  2.996e-02   0.711    0.478
## thetaC:as.factor(Replicate)5          1.538e-02  3.032e-02   0.507    0.612
## thetaC:as.factor(Replicate)6         -4.900e-03  3.021e-02  -0.162    0.871
## thetaC:as.factor(Replicate)7         -1.451e-02  2.995e-02  -0.485    0.628
## thetaC:as.factor(Replicate)8         -4.950e-03  3.005e-02  -0.165    0.869
## thetaC:as.factor(Replicate)9         -1.842e-02  3.006e-02  -0.613    0.541
## thetaC:as.factor(Replicate)10        -3.013e-02  2.947e-02  -1.022    0.308
## rhoC:as.factor(Replicate)2            3.444e-01  5.987e-01   0.575    0.566
## rhoC:as.factor(Replicate)3            1.470e-01  5.241e-01   0.280    0.779
## rhoC:as.factor(Replicate)4           -3.706e-02  5.547e-01  -0.067    0.947
## rhoC:as.factor(Replicate)5           -1.323e-01  5.422e-01  -0.244    0.807
## rhoC:as.factor(Replicate)6            3.968e-01  5.139e-01   0.772    0.441
## rhoC:as.factor(Replicate)7            6.492e-02  5.371e-01   0.121    0.904
## rhoC:as.factor(Replicate)8            3.612e-01  5.336e-01   0.677    0.499
## rhoC:as.factor(Replicate)9            1.852e-01  5.392e-01   0.343    0.732
## rhoC:as.factor(Replicate)10           6.254e-02  5.352e-01   0.117    0.907
## tmrcaC:as.factor(Replicate)2          2.835e-03  4.116e-03   0.689    0.492
## tmrcaC:as.factor(Replicate)3          4.229e-03  3.968e-03   1.066    0.287
## tmrcaC:as.factor(Replicate)4          1.151e-03  4.332e-03   0.266    0.791
## tmrcaC:as.factor(Replicate)5          7.096e-04  4.463e-03   0.159    0.874
## tmrcaC:as.factor(Replicate)6          2.969e-03  4.138e-03   0.718    0.474
## tmrcaC:as.factor(Replicate)7          7.158e-03  4.497e-03   1.592    0.113
## tmrcaC:as.factor(Replicate)8          2.998e-03  4.449e-03   0.674    0.501
## tmrcaC:as.factor(Replicate)9          1.604e-03  4.277e-03   0.375    0.708
## tmrcaC:as.factor(Replicate)10         2.254e-03  4.231e-03   0.533    0.595
## thetaC:tmrcaC:as.factor(Replicate)2   5.699e-01  7.355e-01   0.775    0.439
## thetaC:tmrcaC:as.factor(Replicate)3   1.003e+00  7.144e-01   1.404    0.162
## thetaC:tmrcaC:as.factor(Replicate)4   2.454e-01  8.165e-01   0.301    0.764
## thetaC:tmrcaC:as.factor(Replicate)5   6.558e-01  8.117e-01   0.808    0.420
## thetaC:tmrcaC:as.factor(Replicate)6   3.519e-01  7.102e-01   0.495    0.621
## thetaC:tmrcaC:as.factor(Replicate)7   1.209e+00  7.614e-01   1.588    0.114
## thetaC:tmrcaC:as.factor(Replicate)8   3.950e-01  7.186e-01   0.550    0.583
## thetaC:tmrcaC:as.factor(Replicate)9   1.732e-01  7.086e-01   0.244    0.807
## thetaC:tmrcaC:as.factor(Replicate)10  2.390e-01  6.889e-01   0.347    0.729
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0006464 on 250 degrees of freedom
## Multiple R-squared:  0.9913, Adjusted R-squared:  0.9896
## F-statistic: 582.5 on 49 and 250 DF,  p-value: < 2.2e-16
```

2.3.12 Plots

```
scale.3d <- function(x) sprintf("%.3f", x)

rho.plot <- as.data.frame(cbind(inf.lands.1Mb.rep_1$bin,
                                sim.rho.1Mb$sim,
                                inf.lands.1Mb.rep_1$rho,
                                inf.lands.1Mb.rep_2$rho,
                                inf.lands.1Mb.rep_3$rho,
                                inf.lands.1Mb.rep_4$rho,
                                inf.lands.1Mb.rep_5$rho,
                                inf.lands.1Mb.rep_6$rho,
                                inf.lands.1Mb.rep_7$rho,
                                inf.lands.1Mb.rep_8$rho,
                                inf.lands.1Mb.rep_9$rho,
                                inf.lands.1Mb.rep_10$rho))

rho.plot[,3:ncol(rho.plot)] <- 1.53 * rho.plot[,3:ncol(rho.plot)] # scaling because of SCRM N0 vs Ne

names(rho.plot) <- c("bin", "sim", reps)
molten.rho <- melt(rho.plot, id.vars = "bin")
rho.map.1Mb <- ggplot(data = molten.rho, aes(x = bin * 1000, y = value, colour = variable)) + theme_bw(
rho.map.1Mb <- rho.map.1Mb + geom_line(data = molten.rho, aes(size = variable)) + scale_size_manual(valu
rho.map.1Mb <- rho.map.1Mb + scale_color_manual(values = c("black", brewer.pal(n = 9, name = "Blues"),
rho.map.1Mb <- rho.map.1Mb + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks =
rho.map.1Mb <- rho.map.1Mb + labs(title = NULL, x = NULL, y = expression(rho))
rho.map.1Mb <- rho.map.1Mb + theme(text = element_text(size = 20), axis.title.x = element_text(size = 20

theta.plot <- as.data.frame(cbind(inf.lands.1Mb.rep_1$bin,
                                sim.theta.1Mb$sim,
                                inf.lands.1Mb.rep_1$theta,
                                inf.lands.1Mb.rep_2$theta,
                                inf.lands.1Mb.rep_3$theta,
                                inf.lands.1Mb.rep_4$theta,
                                inf.lands.1Mb.rep_5$theta,
                                inf.lands.1Mb.rep_6$theta,
                                inf.lands.1Mb.rep_7$theta,
                                inf.lands.1Mb.rep_8$theta,
                                inf.lands.1Mb.rep_9$theta,
                                inf.lands.1Mb.rep_10$theta))

names(theta.plot) <- c("bin", "sim", reps)
molten.theta <- melt(theta.plot, id.vars = "bin")
theta.map.1Mb <- ggplot(data = molten.theta, aes(x = bin * 1000, y = value, colour = variable)) + theme_
theta.map.1Mb <- theta.map.1Mb + geom_line(data = molten.theta, aes(size = variable)) + scale_size_manua
theta.map.1Mb <- theta.map.1Mb + scale_color_manual(values = c("black", brewer.pal(n = 9, name = "Reds"
theta.map.1Mb <- theta.map.1Mb + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breal
theta.map.1Mb <- theta.map.1Mb + labs(title = NULL, x = "Position (kb)", y = expression(theta))
theta.map.1Mb <- theta.map.1Mb + theme(text = element_text(size = 20), axis.title.x = element_text(size

plot_grid(rho.map.1Mb, theta.map.1Mb, nrow = 2, ncol = 1)
```

```
theta.maps.coal <- plot_grid(theta.map.50kb, theta.map.200kb, theta.map.1Mb, nrow = 3, ncol = 1)
theta.maps.coal
```

3 Real Drosophila data.

We first focus on chr 2L for comparing R^2 with simulations

```
r2.dm.tab <- as.data.frame(matrix(ncol = 6, nrow = 3))
colnames(r2.dm.tab) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA", "bin.size")

# 50kb
# recombination landscapes
rho.dm.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.rho.50kb.bedgraph", header = T)

# diversity
diversity.dm.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.diversity.50kb.bedgraph", header = T)
diversity.dm.50kb$avg <- apply(diversity.dm.50kb[4:ncol(diversity.dm.50kb)], 1, mean)

# mutation landscapes
theta.dm.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.theta.50kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.TMRCA.50kb.bedgraph", header = T)
tmrca.dm.50kb$sample_mean <- apply(tmrca.dm.50kb[4:ncol(tmrca.dm.50kb)], 1, mean)

# missing data
missing.prop.50kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.missing.prop.50kb.bedgraph", header =
intersect.50kb <- apply(missing.prop.50kb[4:ncol(missing.prop.50kb)], 1, function(x) any(x > 0.1))
```

```
dm.lands.50kb <- as.data.frame(cbind(diversity.dm.50kb$avg,
                                     theta.dm.50kb$sample_mean,
                                     rho.dm.50kb$sample_mean,
                                     tmrca.dm.50kb$sample_mean))

names(dm.lands.50kb) <- c("diversity", "theta", "rho", "tmrca")
dm.lands.50kb$bin <- 1:nrow(dm.lands.50kb)

# filters based on missing data
dm.lands.50kb <- dm.lands.50kb[which(intersect.50kb == F),]

# OLS

# centering
dm.lands.50kb$thetaC <- dm.lands.50kb$theta - mean(dm.lands.50kb$theta)
dm.lands.50kb$tmrcaC <- dm.lands.50kb$tmrca - mean(dm.lands.50kb$tmrca)
dm.lands.50kb$rhoC <- dm.lands.50kb$rho - mean(dm.lands.50kb$rho)

m.diversity <- lm(diversity ~ thetaC + rhoC + tmrcaC + tmrcaC*thetaC, data = dm.lands.50kb)

# type 2 ANOVA
anova.diversity <- Anova(m.diversity)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.dm.tab[1,] <- c((anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
                    anova.diversity$VarExp[1] * 100,
                    anova.diversity$VarExp[2] * 100,
                    anova.diversity$VarExp[3] * 100,
                    anova.diversity$VarExp[4] * 100, 50)

# 200kb
# recombination landscapes
rho.dm.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.rho.200kb.bedgraph", header = T)

# diversity
diversity.dm.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.diversity.200kb.bedgraph", header = T
diversity.dm.200kb$avg <- apply(diversity.dm.200kb[4:ncol(diversity.dm.200kb)], 1, mean)

# mutation landscapes
theta.dm.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.theta.200kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.TMRCA.200kb.bedgraph", header = T)
tmrca.dm.200kb$sample_mean <- apply(tmrca.dm.200kb[4:ncol(tmrca.dm.200kb)], 1, mean)

# missing data
missing.prop.200kb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.missing.prop.200kb.bedgraph", header
intersect.200kb <- apply(missing.prop.200kb[4:ncol(missing.prop.200kb)], 1, function(x) any(x > 0.1))

dm.lands.200kb <- as.data.frame(cbind(diversity.dm.200kb$avg,
                                     theta.dm.200kb$sample_mean,
                                     rho.dm.200kb$sample_mean,
```

```
                                    tmrca.dm.200kb$sample_mean))

names(dm.lands.200kb) <- c("diversity", "theta", "rho", "tmrca")
dm.lands.200kb$bin <- 1:nrow(dm.lands.200kb)
# filters based on missing data
dm.lands.200kb <- dm.lands.200kb[which(intersect.200kb == F),]

# OLS

dm.lands.200kb$thetaC <- dm.lands.200kb$theta - mean(dm.lands.200kb$theta)
dm.lands.200kb$tmrcaC <- dm.lands.200kb$tmrca - mean(dm.lands.200kb$tmrca)
dm.lands.200kb$rhoC <- dm.lands.200kb$rho - mean(dm.lands.200kb$rho)

m.diversity <- lm(diversity ~ thetaC + rhoC + tmrcaC + tmrcaC*thetaC, data = dm.lands.200kb)

# type 2 ANOVA
anova.diversity <- Anova(m.diversity)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.dm.tab[2,] <- c((anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
                    anova.diversity$VarExp[1] * 100,
                    anova.diversity$VarExp[2] * 100,
                    anova.diversity$VarExp[3] * 100,
                    anova.diversity$VarExp[4] * 100, 200)

# 1Mb
# recombination landscapes
rho.dm.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.rho.1Mb.bedgraph", header = T)

# diversity
diversity.dm.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.diversity.1Mb.bedgraph", header = T)
diversity.dm.1Mb$avg <- apply(diversity.dm.1Mb[4:ncol(diversity.dm.1Mb)], 1, mean)

# mutation landscapes
theta.dm.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.theta.1Mb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.TMRCA.1Mb.bedgraph", header = T)
tmrca.dm.1Mb$sample_mean <- apply(tmrca.dm.1Mb[4:ncol(tmrca.dm.1Mb)], 1, mean)

# missing data
missing.prop.1Mb <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.missing.prop.1Mb.bedgraph", header = T
intersect.1Mb <- apply(missing.prop.1Mb[4:ncol(missing.prop.1Mb)], 1, function(x) any(x > 0.1))

dm.lands.1Mb <- as.data.frame(cbind(diversity.dm.1Mb$avg,
                                    theta.dm.1Mb$sample_mean,
                                    rho.dm.1Mb$sample_mean,
                                    tmrca.dm.1Mb$sample_mean))

names(dm.lands.1Mb) <- c("diversity", "theta", "rho", "tmrca")
dm.lands.1Mb$bin <- 1:nrow(dm.lands.1Mb)
# filters based on missing data
```

```
dm.lands.1Mb <- dm.lands.1Mb[which(intersect.1Mb == F),]

# OLS
dm.lands.1Mb$thetaC <- dm.lands.1Mb$theta - mean(dm.lands.1Mb$theta)
dm.lands.1Mb$tmrcaC <- dm.lands.1Mb$tmrca - mean(dm.lands.1Mb$tmrca)
dm.lands.1Mb$rhoC <- dm.lands.1Mb$rho - mean(dm.lands.1Mb$rho)

m.diversity <- lm(diversity ~ thetaC + rhoC + tmrcaC + tmrcaC*thetaC, data = dm.lands.1Mb)

# type 2 ANOVA
anova.diversity <- Anova(m.diversity)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)


r2.dm.tab[3,] <- c((anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
                    anova.diversity$VarExp[4]) * 100,
                    anova.diversity$VarExp[1] * 100,
                    anova.diversity$VarExp[2] * 100,
                    anova.diversity$VarExp[3] * 100,
                    anova.diversity$VarExp[4] * 100, 1000)
```

We now move on to whole-genome analyses

3.1 50 kb windows

```
# Chr 2L

# recombination landscapes
rho.dm.50kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.rho.50kb.bedgraph", header = T)

# diversity
diversity.dm.50kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.diversity.50kb.bedgraph", header =
diversity.dm.50kb.2L$avg <- apply(diversity.dm.50kb.2L[4:ncol(diversity.dm.50kb.2L)], 1, mean)

# mutation landscapes
theta.dm.50kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.theta.50kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.50kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.TMRCA.50kb.bedgraph", header = T)
tmrca.dm.50kb.2L$sample_mean <- apply(tmrca.dm.50kb.2L[4:ncol(tmrca.dm.50kb.2L)], 1, mean)

# missing data
missing.prop.50kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.missing.prop.50kb.bedgraph", heade
intersect.50kb.2L <- apply(missing.prop.50kb.2L[4:ncol(missing.prop.50kb.2L)], 1, function(x) any(x > 0

dm.lands.50kb.2L <- as.data.frame(cbind(diversity.dm.50kb.2L$chromStart,
                                        diversity.dm.50kb.2L$chromEnd,
                                        diversity.dm.50kb.2L$avg,
                                        theta.dm.50kb.2L$sample_mean,
                                        rho.dm.50kb.2L$sample_mean,
                                        tmrca.dm.50kb.2L$sample_mean))
names(dm.lands.50kb.2L) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
```

```r
dm.lands.50kb.2L$bin <- 1:nrow(dm.lands.50kb.2L)

# filters based on missing data
dm.lands.50kb.2L <- dm.lands.50kb.2L[which(intersect.50kb.2L == F),]

dm.lands.50kb.2L$chr <- "2L"

dm.lands.50kb.2L$thetaC <- dm.lands.50kb.2L$theta - mean(dm.lands.50kb.2L$theta)
dm.lands.50kb.2L$tmrcaC <- dm.lands.50kb.2L$tmrca - mean(dm.lands.50kb.2L$tmrca)
dm.lands.50kb.2L$rhoC <- dm.lands.50kb.2L$rho - mean(dm.lands.50kb.2L$rho)

g.div.dm.50kb.2L <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                        data = dm.lands.50kb.2L, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), m

summary(g.div.dm.50kb.2L)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.50kb.2L
##         AIC       BIC  logLik
##   -4690.441 -4660.219 2353.22
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~bin
##  Parameter estimate(s):
##      Phi1
## 0.1318436
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##      power
## 0.02239814
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0097130 0.00001087 893.9585  0.0000
## thetaC        0.9874554 0.00504783 195.6197  0.0000
## rhoC          0.0017748 0.00146095   1.2148  0.2253
## tmrcaC        0.0126819 0.00021188  59.8534  0.0000
## thetaC:tmrcaC 1.2080584 0.04507126  26.8033  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.016
## rhoC           0.027  0.026
## tmrcaC        -0.132 -0.379 -0.456
## thetaC:tmrcaC -0.306  0.061 -0.088  0.402
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.60864294 -0.52688547 -0.05949911  0.51229899  4.16571201
##
## Residual standard error: 0.0001811126
```

```
## Degrees of freedom: 323 total; 318 residual
```

```r
# Chr 2R

# recombination landscapes
rho.dm.50kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.rho.50kb.bedgraph", header = T)

# diversity
diversity.dm.50kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.diversity.50kb.bedgraph", header =
diversity.dm.50kb.2R$avg <- apply(diversity.dm.50kb.2R[4:ncol(diversity.dm.50kb.2R)], 1, mean)

# mutation landscapes
theta.dm.50kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.theta.50kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.50kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.TMRCA.50kb.bedgraph", header = T)
tmrca.dm.50kb.2R$sample_mean <- apply(tmrca.dm.50kb.2R[4:ncol(tmrca.dm.50kb.2R)], 1, mean)

# missing data
missing.prop.50kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.missing.prop.50kb.bedgraph", heade
intersect.50kb.2R <- apply(missing.prop.50kb.2R[4:ncol(missing.prop.50kb.2R)], 1, function(x) any(x > 0

dm.lands.50kb.2R <- as.data.frame(cbind(diversity.dm.50kb.2R$chromStart,
                                        diversity.dm.50kb.2R$chromEnd,
                                        diversity.dm.50kb.2R$avg,
                                        theta.dm.50kb.2R$sample_mean,
                                        rho.dm.50kb.2R$sample_mean,
                                        tmrca.dm.50kb.2R$sample_mean))
names(dm.lands.50kb.2R) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.50kb.2R$bin <- 1:nrow(dm.lands.50kb.2R)

# filters based on missing data
dm.lands.50kb.2R <- dm.lands.50kb.2R[which(intersect.50kb.2R == F),]

dm.lands.50kb.2R$chr <- "2R"

dm.lands.50kb.2R$thetaC <- dm.lands.50kb.2R$theta- mean(dm.lands.50kb.2R$theta)
dm.lands.50kb.2R$tmrcaC <- dm.lands.50kb.2R$tmrca - mean(dm.lands.50kb.2R$tmrca)
dm.lands.50kb.2R$rhoC <- dm.lands.50kb.2R$rho - mean(dm.lands.50kb.2R$rho)

g.div.dm.50kb.2R <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                        data = dm.lands.50kb.2R, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin),

summary(g.div.dm.50kb.2R)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.50kb.2R
##         AIC       BIC   logLik
##   -4473.266 -4443.583 2244.633
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~bin
##  Parameter estimate(s):
##      Phi1
```

```
## 0.1094586
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##      power
## 0.05919674
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0085652 0.00000938 913.0793  0.0000
## thetaC        0.9709055 0.00361092 268.8801  0.0000
## rhoC          0.0001728 0.00149052   0.1160  0.9078
## tmrcaC        0.0116837 0.00019594  59.6303  0.0000
## thetaC:tmrcaC 1.0680644 0.04933969  21.6472  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.044
## rhoC          -0.003  0.084
## tmrcaC        -0.103 -0.360 -0.381
## thetaC:tmrcaC -0.260 -0.114  0.022  0.341
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.56718360 -0.56905223  0.01512333  0.64333847  3.32352563
##
## Residual standard error: 0.0001778837
## Degrees of freedom: 302 total; 297 residual
```

```
# recombination landscapes
rho.dm.50kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.rho.50kb.bedgraph", header = T)

# diversity
diversity.dm.50kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.diversity.50kb.bedgraph", header =
diversity.dm.50kb.3L$avg <- apply(diversity.dm.50kb.3L[4:ncol(diversity.dm.50kb.3L)], 1, mean)

# mutation landscapes
theta.dm.50kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.theta.50kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.50kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.TMRCA.50kb.bedgraph", header = T)
tmrca.dm.50kb.3L$sample_mean <- apply(tmrca.dm.50kb.3L[4:ncol(tmrca.dm.50kb.3L)], 1, mean)

# missing data
missing.prop.50kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.missing.prop.50kb.bedgraph", header
intersect.50kb.3L <- apply(missing.prop.50kb.3L[4:ncol(missing.prop.50kb.3L)], 1, function(x) any(x > 0

dm.lands.50kb.3L <- as.data.frame(cbind(diversity.dm.50kb.3L$chromStart,
                                        diversity.dm.50kb.3L$chromEnd,
                                        diversity.dm.50kb.3L$avg,
                                        theta.dm.50kb.3L$sample_mean,
                                        rho.dm.50kb.3L$sample_mean,
                                        tmrca.dm.50kb.3L$sample_mean))
```

```r
names(dm.lands.50kb.3L) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.50kb.3L$bin <- 1:nrow(dm.lands.50kb.3L)

# filters based on missing data
dm.lands.50kb.3L <- dm.lands.50kb.3L[which(intersect.50kb.3L == F),]

dm.lands.50kb.3L$chr <- "3L"

dm.lands.50kb.3L$thetaC <- dm.lands.50kb.3L$theta- mean(dm.lands.50kb.3L$theta)
dm.lands.50kb.3L$tmrcaC <- dm.lands.50kb.3L$tmrca - mean(dm.lands.50kb.3L$tmrca)
dm.lands.50kb.3L$rhoC <- dm.lands.50kb.3L$rho - mean(dm.lands.50kb.3L$rho)

g.div.dm.50kb.3L <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                        data = dm.lands.50kb.3L, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin),

summary(g.div.dm.50kb.3L)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.50kb.3L
##         AIC       BIC   logLik
##   -5025.491 -4994.674 2520.746
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~bin
##  Parameter estimate(s):
##      Phi1
## 0.2696667
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##      power
## 0.06685875
##
## Coefficients:
##                   Value   Std.Error  t-value p-value
## (Intercept)   0.0089556 0.000012514 715.6471  0.0000
## thetaC        0.9656448 0.004479703 215.5600  0.0000
## rhoC          0.0028522 0.001435660   1.9867  0.0477
## tmrcaC        0.0118244 0.000153683  76.9400  0.0000
## thetaC:tmrcaC 1.0917168 0.030139577  36.2220  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.004
## rhoC           0.004  0.150
## tmrcaC        -0.106 -0.427 -0.517
## thetaC:tmrcaC -0.323  0.045 -0.029  0.309
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.70241395 -0.68659890  0.01012415  0.61472090  3.79680208
##
```

```
## Residual standard error: 0.0002214495
## Degrees of freedom: 348 total; 343 residual
```

```r
# Chr 3R

# recombination landscapes
rho.dm.50kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.rho.50kb.bedgraph", header = T)

# diversity
diversity.dm.50kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.diversity.50kb.bedgraph", header =
diversity.dm.50kb.3R$avg <- apply(diversity.dm.50kb.3R[4:ncol(diversity.dm.50kb.3R)], 1, mean)

# mutation landscapes
theta.dm.50kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.theta.50kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.50kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.TMRCA.50kb.bedgraph", header = T)
tmrca.dm.50kb.3R$sample_mean <- apply(tmrca.dm.50kb.3R[4:ncol(tmrca.dm.50kb.3R)], 1, mean)

# missing data
missing.prop.50kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.missing.prop.50kb.bedgraph", header
intersect.50kb.3R <- apply(missing.prop.50kb.3R[4:ncol(missing.prop.50kb.3R)], 1, function(x) any(x > 0

dm.lands.50kb.3R <- as.data.frame(cbind(diversity.dm.50kb.3R$chromStart,
                                        diversity.dm.50kb.3R$chromEnd,
                                        diversity.dm.50kb.3R$avg,
                                        theta.dm.50kb.3R$sample_mean,
                                        rho.dm.50kb.3R$sample_mean,
                                        tmrca.dm.50kb.3R$sample_mean))
names(dm.lands.50kb.3R) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.50kb.3R$bin <- 1:nrow(dm.lands.50kb.3R)

# filters based on missing data
dm.lands.50kb.3R <- dm.lands.50kb.3R[which(intersect.50kb.3R == F),]

dm.lands.50kb.3R$chr <- "3R"

dm.lands.50kb.3R$thetaC <- dm.lands.50kb.3R$theta- mean(dm.lands.50kb.3R$theta)
dm.lands.50kb.3R$tmrcaC <- dm.lands.50kb.3R$tmrca - mean(dm.lands.50kb.3R$tmrca)
dm.lands.50kb.3R$rhoC <- dm.lands.50kb.3R$rho - mean(dm.lands.50kb.3R$rho)

g.div.dm.50kb.3R <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                        data = dm.lands.50kb.3R, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin),

summary(g.div.dm.50kb.3R)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.50kb.3R
##         AIC       BIC   logLik
##   -6576.594 -6543.846 3296.297
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~bin
##  Parameter estimate(s):
```

```
##      Phi1
## 0.2356751
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##     power
## 0.1042665
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0075334 0.00000895 841.7772  0.0000
## thetaC         0.9644922 0.00331480 290.9655  0.0000
## rhoC          -0.0017265 0.00118208  -1.4606  0.1449
## tmrcaC         0.0106635 0.00013264  80.3973  0.0000
## thetaC:tmrcaC  1.0663698 0.03719362  28.6708  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.003
## rhoC          -0.022  0.107
## tmrcaC        -0.203 -0.346 -0.501
## thetaC:tmrcaC -0.384 -0.089  0.039  0.507
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.64020083 -0.59320311  0.03347561  0.68857658  4.67175243
##
## Residual standard error: 0.0002041515
## Degrees of freedom: 443 total; 438 residual
```

```
# all together now
dm.lands.50kb <- rbind.data.frame(dm.lands.50kb.2L, dm.lands.50kb.2R, dm.lands.50kb.3L, dm.lands.50kb.3

write.table(dm.lands.50kb, "dm_data/dm_chr_maps/dm.maps.50kb.tsv", quote = F, sep = "\t", row.names = F


# Plots
scale.3d <- function(x) sprintf("%.3f", x) # digits shown in y axis

setTimeLimit(cpu = Inf, elapsed = Inf, transient = F)

molten.diversity <- melt(dm.lands.50kb[c(3,7,8)], id.vars = c("bin", "chr"))
diversity.map <- ggplot(data = molten.diversity, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
diversity.map <- diversity.map + geom_line(data = molten.diversity, colour = "#F8766D")
diversity.map <- diversity.map + geom_smooth(method = "loess", se = F, colour = "#F8766D")
diversity.map <- diversity.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breal
diversity.map <- diversity.map + labs(title = NULL, x = NULL, y = expression(pi))
diversity.map <- diversity.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_te
diversity.map
```

```
molten.rho <- melt(dm.lands.50kb[c(5,7,8)], id.vars = c("bin", "chr"))
rho.map <- ggplot(data = molten.rho, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
rho.map <- rho.map + geom_line(data = molten.rho, colour = "#7CAE00")
rho.map <- rho.map + geom_smooth(method = "loess", se = F, colour = "#7CAE00")
rho.map <- rho.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks = pretty_
rho.map <- rho.map + labs(title = NULL, x = NULL, y = expression(rho))
rho.map <- rho.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size = 2
rho.map
```

```
molten.theta <- melt(dm.lands.50kb[c(4,7,8)], id.vars = c("bin", "chr"))
theta.map <- ggplot(data = molten.theta, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
theta.map <- theta.map + geom_line(data = molten.theta, colour = "#00BFC4")
theta.map <- theta.map + geom_smooth(method = "loess", se = F, colour = "#00BFC4")
theta.map <- theta.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks = pre
theta.map <- theta.map + labs(title = NULL, x = NULL, y = expression(theta))
theta.map <- theta.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size
theta.map
```

```
molten.tmrca <- melt(dm.lands.50kb[c(6,7,8)], id.vars = c("bin", "chr"))
tmrca.map <- ggplot(data = molten.tmrca, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
tmrca.map <- tmrca.map + geom_line(data = molten.tmrca, colour = "#88419d")
tmrca.map <- tmrca.map + geom_smooth(method = "loess", se = F, colour = "#88419d")
tmrca.map <- tmrca.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks = pre
tmrca.map <- tmrca.map + labs(title = NULL, x = "Position (kb)", y = expression(tau))
tmrca.map <- tmrca.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size
tmrca.map
```

```
# genome-wide correlations between maps
cor.test(~rho + diversity, data = dm.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and diversity
## S = 369190884, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.2197879
```

```
cor.test(~rho + tmrca, data = dm.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 245063950, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.4821057
```

```
cor.test(~theta + tmrca, data = dm.lands.50kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
```

```
##
## data:  theta and tmrca
## S = 254806288, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.4615172
```

```r
# Linear models
# centering
dm.lands.50kb$thetaC <- dm.lands.50kb$theta - mean(dm.lands.50kb$theta)
dm.lands.50kb$tmrcaC <- dm.lands.50kb$tmrca - mean(dm.lands.50kb$tmrca)
dm.lands.50kb$rhoC <- dm.lands.50kb$rho - mean(dm.lands.50kb$rho)

dm.lands.50kb$bin <- 1:nrow(dm.lands.50kb)

m.div.dm.50kb <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC), data = dm.lands.50kb)

plot(resid(m.div.dm.50kb)~fitted(m.div.dm.50kb))
```

```r
hist(resid(m.div.dm.50kb), nclass = 30)
```

```
dwtest(m.div.dm.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.dm.50kb
## DW = 1.5747, p-value = 3.694e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.dm.50kb, nsim = 10000)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.dm.50kb
## HMC = 0.50534, p-value = 0.6096
```

```
summary(m.div.dm.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
##     data = dm.lands.50kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -7.369e-04 -1.007e-04  1.500e-07  9.777e-05  8.124e-04
##
## Coefficients:
##                Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   8.603e-03  4.838e-06 1778.154   <2e-16 ***
```

```
## thetaC        9.759e-01  1.822e-03  535.514   <2e-16 ***
## rhoC          1.209e-03  6.874e-04    1.759   0.0788 .
## tmrcaC        1.139e-02  7.780e-05  146.379   <2e-16 ***
## thetaC:tmrcaC 1.065e+00  1.654e-02   64.386   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001675 on 1411 degrees of freedom
## Multiple R-squared:  0.9973, Adjusted R-squared:  0.9973
## F-statistic: 1.286e+05 on 4 and 1411 DF,  p-value: < 2.2e-16
```

```
# type 2 ANOVA
anova.diversity <- Anova(m.div.dm.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
anova.diversity
```

```
## Anova Table (Type II tests)
##
## Response: diversity
##                  Sum Sq   Df    F value   Pr(>F)  VarExp
## thetaC        0.0081479    1 2.9046e+05 0.000000 0.92684
## rhoC          0.0000001    1 3.0932e+00 0.078838 0.00001
## tmrcaC        0.0004872    1 1.7367e+04 0.000000 0.05542
## thetaC:tmrcaC 0.0001163    1 4.1456e+03 0.000000 0.01323
## Residuals     0.0000396 1411                     0.00450
```

```
# GLS
g.div.dm.50kb.1 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                       data = dm.lands.50kb, weights = varPower(0, ~tmrcaC|chr), cor = corAR1(0, ~bin|ch

g.div.dm.50kb.2 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                       data = dm.lands.50kb, weights = varPower(0, ~theta|chr), cor = corAR1(0, ~bin|chr

g.div.dm.50kb.3 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                       data = dm.lands.50kb, weights = varPower(0, ~theta|chr), method = "ML")

g.div.dm.50kb.4 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                       data = dm.lands.50kb, weights = varPower(0, ~tmrcaC|chr), method = "ML")

AIC(g.div.dm.50kb.1, g.div.dm.50kb.2, g.div.dm.50kb.3, g.div.dm.50kb.4)
```

```
##                 df       AIC
## g.div.dm.50kb.1 11 -20716.73
## g.div.dm.50kb.2 11 -20681.77
## g.div.dm.50kb.3 10 -20620.30
## g.div.dm.50kb.4 10 -20655.40
```

```
summary(g.div.dm.50kb.1)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.50kb
##         AIC       BIC   logLik
##   -20716.73 -20658.92 10369.36
##
```

```
## Correlation Structure: AR(1)
##  Formula: ~bin | chr
##  Parameter estimate(s):
##       Phi
## 0.2144958
## Variance function:
##  Structure: Power of variance covariate, different strata
##  Formula: ~tmrcaC | chr
##  Parameter estimates:
##         2L         2R         3L         3R
## 0.05445085 0.09431089 0.06301189 0.11582370
##
## Coefficients:
##                   Value   Std.Error   t-value p-value
## (Intercept)   0.0085894 0.000005516 1557.0928  0.0000
## thetaC        0.9739854 0.002055447  473.8558  0.0000
## rhoC          0.0004325 0.000704757    0.6136  0.5396
## tmrcaC        0.0115855 0.000084597  136.9497  0.0000
## thetaC:tmrcaC 1.0997196 0.018815526   58.4475  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.044
## rhoC          -0.005  0.137
## tmrcaC        -0.143 -0.375 -0.480
## thetaC:tmrcaC -0.289 -0.046  0.003  0.409
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.04794537 -0.58184890  0.04817139  0.65222011  4.30624880
##
## Residual standard error: 0.0002182617
## Degrees of freedom: 1416 total; 1411 residual
```

```
vif(g.div.dm.50kb.1)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.191539      1.394432      1.938114      1.306769
```

```
# Linear model without TMRCA --> rho becomes significant
g.div.dm.50kb.5 <- gls(diversity ~ (thetaC + rhoC),
                       data = dm.lands.50kb, weights = varPower(0, ~thetaC|chr), cor = corAR1(0, ~bin|c

summary(g.div.dm.50kb.5)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC)
##   Data: dm.lands.50kb
##         AIC       BIC   logLik
##   -16799.87 -16752.56 8408.933
##
## Correlation Structure: AR(1)
##  Formula: ~bin | chr
##  Parameter estimate(s):
##       Phi
```

```
## 0.2247151
## Variance function:
##  Structure: Power of variance covariate, different strata
##  Formula: ~thetaC | chr
##  Parameter estimates:
##         2L         2R         3L         3R
## 0.09783160 0.09779667 0.05144519 0.09035490
##
## Coefficients:
##                 Value   Std.Error  t-value p-value
## (Intercept) 0.0087332 0.000021244 411.0954       0
## thetaC      1.0944710 0.008181042 133.7814       0
## rhoC        0.0499971 0.002361411  21.1725       0
##
##  Correlation:
##        (Intr) thetaC
## thetaC -0.012
## rhoC   -0.003 -0.148
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -12.55241945  -0.53031140   0.01367413   0.61083754   3.01722550
##
## Residual standard error: 0.001130577
## Degrees of freedom: 1416 total; 1413 residual
```

3.2 200 kb windows

```
# Chr 2L

# recombination landscapes
rho.dm.200kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.rho.200kb.bedgraph", header = T)

# diversity
diversity.dm.200kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.diversity.200kb.bedgraph", header
diversity.dm.200kb.2L$avg <- apply(diversity.dm.200kb.2L[4:ncol(diversity.dm.200kb.2L)], 1, mean)

# mutation landscapes
theta.dm.200kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.theta.200kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.200kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.TMRCA.200kb.bedgraph", header = T)
tmrca.dm.200kb.2L$sample_mean <- apply(tmrca.dm.200kb.2L[4:ncol(tmrca.dm.200kb.2L)], 1, mean)

# missing data
missing.prop.200kb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.missing.prop.200kb.bedgraph", hea
intersect.200kb.2L <- apply(missing.prop.200kb.2L[4:ncol(missing.prop.200kb.2L)], 1, function(x) any(x

dm.lands.200kb.2L <- as.data.frame(cbind(diversity.dm.200kb.2L$chromStart,
                                         diversity.dm.200kb.2L$chromEnd,
                                         diversity.dm.200kb.2L$avg,
                                         theta.dm.200kb.2L$sample_mean,
                                         rho.dm.200kb.2L$sample_mean,
                                         tmrca.dm.200kb.2L$sample_mean))
```

```r
names(dm.lands.200kb.2L) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.200kb.2L$bin <- 1:nrow(dm.lands.200kb.2L)

# filters based on missing data
dm.lands.200kb.2L <- dm.lands.200kb.2L[which(intersect.200kb.2L == F),]

dm.lands.200kb.2L$chr <- "2L"

dm.lands.200kb.2L$thetaC <- dm.lands.200kb.2L$theta - mean(dm.lands.200kb.2L$theta)
dm.lands.200kb.2L$tmrcaC <- dm.lands.200kb.2L$tmrca - mean(dm.lands.200kb.2L$tmrca)
dm.lands.200kb.2L$rhoC <- dm.lands.200kb.2L$rho - mean(dm.lands.200kb.2L$rho)

g.div.dm.200kb.2L <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                         data = dm.lands.200kb.2L, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin)

summary(g.div.dm.200kb.2L)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.200kb.2L
##         AIC       BIC  logLik
##   -1244.082 -1224.731 630.041
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~bin
##  Parameter estimate(s):
##      Phi1
## 0.1002962
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##      power
## 0.05640964
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0096851 0.00001590 609.2540  0.0000
## thetaC        0.9912075 0.00852140 116.3197  0.0000
## rhoC          0.0035710 0.00328708   1.0864  0.2807
## tmrcaC        0.0122532 0.00061930  19.7855  0.0000
## thetaC:tmrcaC 1.0357572 0.10202136  10.1524  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.034
## rhoC           0.100  0.195
## tmrcaC        -0.244 -0.449 -0.572
## thetaC:tmrcaC -0.337 -0.112 -0.276  0.696
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.8896987 -0.6840157 -0.1976799  0.4999174  3.1380216
##
```

```
## Residual standard error: 0.0001535204
## Degrees of freedom: 83 total; 78 residual
```

```r
# Chr 2R

# recombination landscapes
rho.dm.200kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.rho.200kb.bedgraph", header = T)

# diversity
diversity.dm.200kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.diversity.200kb.bedgraph", header
diversity.dm.200kb.2R$avg <- apply(diversity.dm.200kb.2R[4:ncol(diversity.dm.200kb.2R)], 1, mean)

# mutation landscapes
theta.dm.200kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.theta.200kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.200kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.TMRCA.200kb.bedgraph", header = T)
tmrca.dm.200kb.2R$sample_mean <- apply(tmrca.dm.200kb.2R[4:ncol(tmrca.dm.200kb.2R)], 1, mean)

# missing data
missing.prop.200kb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.missing.prop.200kb.bedgraph", hea
intersect.200kb.2R <- apply(missing.prop.200kb.2R[4:ncol(missing.prop.200kb.2R)], 1, function(x) any(x 

dm.lands.200kb.2R <- as.data.frame(cbind(diversity.dm.200kb.2R$chromStart,
                                         diversity.dm.200kb.2R$chromEnd,
                                         diversity.dm.200kb.2R$avg,
                                         theta.dm.200kb.2R$sample_mean,
                                         rho.dm.200kb.2R$sample_mean,
                                         tmrca.dm.200kb.2R$sample_mean))
names(dm.lands.200kb.2R) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.200kb.2R$bin <- 1:nrow(dm.lands.200kb.2R)

# filters based on missing data
dm.lands.200kb.2R <- dm.lands.200kb.2R[which(intersect.200kb.2R == F),]

dm.lands.200kb.2R$chr <- "2R"

dm.lands.200kb.2R$thetaC <- dm.lands.200kb.2R$theta- mean(dm.lands.200kb.2R$theta)
dm.lands.200kb.2R$tmrcaC <- dm.lands.200kb.2R$tmrca - mean(dm.lands.200kb.2R$tmrca)
dm.lands.200kb.2R$rhoC <- dm.lands.200kb.2R$rho - mean(dm.lands.200kb.2R$rho)

g.div.dm.200kb.2R <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                         data = dm.lands.200kb.2R, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin)

summary(g.div.dm.200kb.2R)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.200kb.2R
##         AIC       BIC   logLik
##   -1165.932 -1147.499 590.9659
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~bin
##  Parameter estimate(s):
```

```
##      Phi1
## 0.2372428
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##      power
## 0.06275131
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0085392 0.00001264 675.8307  0.0000
## thetaC        0.9699152 0.00505288 191.9530  0.0000
## rhoC          0.0027465 0.00292834   0.9379  0.3516
## tmrcaC        0.0110299 0.00043604  25.2959  0.0000
## thetaC:tmrcaC 0.9373743 0.11041693   8.4894  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.030
## rhoC          -0.063  0.097
## tmrcaC        -0.153 -0.376 -0.219
## thetaC:tmrcaC -0.290 -0.069  0.120  0.497
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.13836840 -0.68198223  0.08032991  0.80898896  2.57448459
##
## Residual standard error: 0.0001098196
## Degrees of freedom: 74 total; 69 residual
```

```
# recombination landscapes
rho.dm.200kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.rho.200kb.bedgraph", header = T)

# diversity
diversity.dm.200kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.diversity.200kb.bedgraph", header
diversity.dm.200kb.3L$avg <- apply(diversity.dm.200kb.3L[4:ncol(diversity.dm.200kb.3L)], 1, mean)

# mutation landscapes
theta.dm.200kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.theta.200kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.200kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.TMRCA.200kb.bedgraph", header = T)
tmrca.dm.200kb.3L$sample_mean <- apply(tmrca.dm.200kb.3L[4:ncol(tmrca.dm.200kb.3L)], 1, mean)

# missing data
missing.prop.200kb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.missing.prop.200kb.bedgraph", hea
intersect.200kb.3L <- apply(missing.prop.200kb.3L[4:ncol(missing.prop.200kb.3L)], 1, function(x) any(x

dm.lands.200kb.3L <- as.data.frame(cbind(diversity.dm.200kb.3L$chromStart,
                                         diversity.dm.200kb.3L$chromEnd,
                                         diversity.dm.200kb.3L$avg,
                                         theta.dm.200kb.3L$sample_mean,
                                         rho.dm.200kb.3L$sample_mean,
```

```r
                                        tmrca.dm.200kb.3L$sample_mean))
names(dm.lands.200kb.3L) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.200kb.3L$bin <- 1:nrow(dm.lands.200kb.3L)

# filters based on missing data
dm.lands.200kb.3L <- dm.lands.200kb.3L[which(intersect.200kb.3L == F),]

dm.lands.200kb.3L$chr <- "3L"

dm.lands.200kb.3L$thetaC <- dm.lands.200kb.3L$theta- mean(dm.lands.200kb.3L$theta)
dm.lands.200kb.3L$tmrcaC <- dm.lands.200kb.3L$tmrca - mean(dm.lands.200kb.3L$tmrca)
dm.lands.200kb.3L$rhoC <- dm.lands.200kb.3L$rho - mean(dm.lands.200kb.3L$rho)

g.div.dm.200kb.3L <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                         data = dm.lands.200kb.3L, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin)

summary(g.div.dm.200kb.3L)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.200kb.3L
##         AIC       BIC   logLik
##   -1307.956 -1288.321 661.9781
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~bin
##  Parameter estimate(s):
##      Phi1
## 0.2105682
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##       power
## -0.01470136
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0089474 0.00001816 492.7438  0.0000
## thetaC        0.9811147 0.00736731 133.1713  0.0000
## rhoC          0.0078270 0.00261567   2.9923  0.0037
## tmrcaC        0.0106115 0.00042145  25.1784  0.0000
## thetaC:tmrcaC 0.7871946 0.07703046  10.2193  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.133
## rhoC           0.079  0.300
## tmrcaC        -0.394 -0.661 -0.463
## thetaC:tmrcaC -0.587 -0.238 -0.153  0.680
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.89079240 -0.54320978  0.05844326  0.51224633  2.31374278
```

```
##
## Residual standard error: 0.0001059737
## Degrees of freedom: 86 total; 81 residual
```

```r
# Chr 3R

# recombination landscapes
rho.dm.200kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.rho.200kb.bedgraph", header = T)

# diversity
diversity.dm.200kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.diversity.200kb.bedgraph", header

diversity.dm.200kb.3R$avg <- apply(diversity.dm.200kb.3R[4:ncol(diversity.dm.200kb.3R)], 1, mean)

# mutation landscapes
theta.dm.200kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.theta.200kb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.200kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.TMRCA.200kb.bedgraph", header = T)
tmrca.dm.200kb.3R$sample_mean <- apply(tmrca.dm.200kb.3R[4:ncol(tmrca.dm.200kb.3R)], 1, mean)

# missing data
missing.prop.200kb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.missing.prop.200kb.bedgraph", hea
intersect.200kb.3R <- apply(missing.prop.200kb.3R[4:ncol(missing.prop.200kb.3R)], 1, function(x) any(x 

dm.lands.200kb.3R <- as.data.frame(cbind(diversity.dm.200kb.3R$chromStart,
                                         diversity.dm.200kb.3R$chromEnd,
                                         diversity.dm.200kb.3R$avg,
                                         theta.dm.200kb.3R$sample_mean,
                                         rho.dm.200kb.3R$sample_mean,
                                         tmrca.dm.200kb.3R$sample_mean))
names(dm.lands.200kb.3R) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.200kb.3R$bin <- 1:nrow(dm.lands.200kb.3R)

# filters based on missing data
dm.lands.200kb.3R <- dm.lands.200kb.3R[which(intersect.200kb.3R == F),]

dm.lands.200kb.3R$chr <- "3R"

dm.lands.200kb.3R$thetaC <- dm.lands.200kb.3R$theta- mean(dm.lands.200kb.3R$theta)
dm.lands.200kb.3R$tmrcaC <- dm.lands.200kb.3R$tmrca - mean(dm.lands.200kb.3R$tmrca)
dm.lands.200kb.3R$rhoC <- dm.lands.200kb.3R$rho - mean(dm.lands.200kb.3R$rho)

g.div.dm.200kb.3R <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                         data = dm.lands.200kb.3R, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin)

summary(g.div.dm.200kb.3R)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.200kb.3R
##         AIC       BIC  logLik
##   -1883.368 -1861.135 949.684
##
## Correlation Structure: ARMA(1,0)
```

```
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi1
## 0.09974948
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##     power
## 0.1960113
##
## Coefficients:
##                     Value  Std.Error  t-value p-value
## (Intercept)    0.0075661 0.00001040 727.7696  0.0000
## thetaC         0.9683546 0.00387775 249.7209  0.0000
## rhoC          -0.0041012 0.00175339  -2.3390  0.0211
## tmrcaC         0.0103602 0.00028386  36.4976  0.0000
## thetaC:tmrcaC  0.7922675 0.07387641  10.7242  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.283
## rhoC           0.214  0.306
## tmrcaC        -0.574 -0.559 -0.546
## thetaC:tmrcaC -0.625 -0.374 -0.215  0.781
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.78869608 -0.72931864 -0.06801681  0.64535064  2.75174311
##
## Residual standard error: 0.0001631813
## Degrees of freedom: 119 total; 114 residual
```

```
# all together now
dm.lands.200kb <- rbind.data.frame(dm.lands.200kb.2L, dm.lands.200kb.2R, dm.lands.200kb.3L, dm.lands.20

write.table(dm.lands.200kb, "dm_data/dm_chr_maps/dm.maps.200kb.tsv", quote = F, sep = "\t", row.names =


# Plots
scale.3d <- function(x) sprintf("%.3f", x) # digits shown in y axis

setTimeLimit(cpu = Inf, elapsed = Inf, transient = F)

molten.diversity <- melt(dm.lands.200kb[c(3,7,8)], id.vars = c("bin", "chr"))
diversity.map <- ggplot(data = molten.diversity, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
diversity.map <- diversity.map + geom_line(data = molten.diversity, colour = "#F8766D")
diversity.map <- diversity.map + geom_smooth(method = "loess", se = F, colour = "#F8766D")
diversity.map <- diversity.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(brea
diversity.map <- diversity.map + labs(title = NULL, x = NULL, y = expression(pi))
diversity.map <- diversity.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_te
diversity.map
```

```
molten.rho <- melt(dm.lands.200kb[c(5,7,8)], id.vars = c("bin", "chr"))
rho.map <- ggplot(data = molten.rho, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
rho.map <- rho.map + geom_line(data = molten.rho, colour = "#7CAE00")
rho.map <- rho.map + geom_smooth(method = "loess", se = F, colour = "#7CAE00")
rho.map <- rho.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks = pretty_
rho.map <- rho.map + labs(title = NULL, x = NULL, y = expression(rho))
rho.map <- rho.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size = 2
rho.map
```

```
molten.theta <- melt(dm.lands.200kb[c(4,7,8)], id.vars = c("bin", "chr"))
theta.map <- ggplot(data = molten.theta, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
theta.map <- theta.map + geom_line(data = molten.theta, colour = "#00BFC4")
theta.map <- theta.map + geom_smooth(method = "loess", se = F, colour = "#00BFC4")
theta.map <- theta.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks = pre
theta.map <- theta.map + labs(title = NULL, x = NULL, y = expression(theta))
theta.map <- theta.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size
theta.map
```

```
molten.tmrca <- melt(dm.lands.200kb[c(6,7,8)], id.vars = c("bin", "chr"))
tmrca.map <- ggplot(data = molten.tmrca, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
tmrca.map <- tmrca.map + geom_line(data = molten.tmrca, colour = "#88419d")
tmrca.map <- tmrca.map + geom_smooth(method = "loess", se = F, colour = "#88419d")
tmrca.map <- tmrca.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks = pre
tmrca.map <- tmrca.map + labs(title = NULL, x = "Position (kb)", y = expression(tau))
tmrca.map <- tmrca.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size
tmrca.map
```

```r
# genome-wide correlations
cor.test(~rho + diversity, data = dm.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and diversity
## S = 6652998, p-value = 0.002515
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.1585153
```

```r
cor.test(~rho + tmrca, data = dm.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 4384696, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.4454147
```

```r
cor.test(~theta + tmrca, data = dm.lands.200kb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
```

```
##
## data:  theta and tmrca
## S = 3844174, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.513781
```

```
# Linear models
# centering
dm.lands.200kb$thetaC <- dm.lands.200kb$theta - mean(dm.lands.200kb$theta)
dm.lands.200kb$tmrcaC <- dm.lands.200kb$tmrca - mean(dm.lands.200kb$tmrca)
dm.lands.200kb$rhoC <- dm.lands.200kb$rho - mean(dm.lands.200kb$rho)

dm.lands.200kb$bin <- 1:nrow(dm.lands.200kb)

m.div.dm.200kb <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC), data = dm.lands.200kb)

plot(resid(m.div.dm.200kb)~fitted(m.div.dm.200kb))
```

```
hist(resid(m.div.dm.200kb), nclass = 30)
```

```
dwtest(m.div.dm.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.dm.200kb
## DW = 1.4463, p-value = 3.114e-08
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.dm.200kb, nsim = 10000)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.dm.200kb
## HMC = 0.56948, p-value = 0.9698
```

```
summary(m.div.dm.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
##     data = dm.lands.200kb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.481e-04 -6.976e-05 -4.240e-06  6.351e-05  4.418e-04
##
## Coefficients:
##                Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   8.571e-03  7.361e-06 1164.402   <2e-16 ***
```

```
## thetaC        9.802e-01  2.925e-03  335.083   <2e-16 ***
## rhoC          1.652e-03  1.319e-03    1.252    0.211
## tmrcaC        1.104e-02  2.092e-04   52.753   <2e-16 ***
## thetaC:tmrcaC 8.772e-01  4.258e-02   20.604   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001143 on 357 degrees of freedom
## Multiple R-squared:  0.9985, Adjusted R-squared:  0.9985
## F-statistic: 6.074e+04 on 4 and 357 DF,  p-value: < 2.2e-16
```

```
# type 2 ANOVA
anova.diversity <- Anova(m.div.dm.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
anova.diversity
```

```
## Anova Table (Type II tests)
##
## Response: diversity
##                    Sum Sq  Df    F value  Pr(>F)  VarExp
## thetaC        0.00165091   1 126265.729 0.00000 0.96688
## rhoC          0.00000002   1      1.567 0.21147 0.00001
## tmrcaC        0.00004631   1   3541.894 0.00000 0.02712
## thetaC:tmrcaC 0.00000555   1    424.530 0.00000 0.00325
## Residuals     0.00000467 357                    0.00273
```

```
# GLS
g.div.dm.200kb.1 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                        data = dm.lands.200kb, weights = varPower(0, ~tmrcaC|chr), cor = corAR1(0, ~bin|c

g.div.dm.200kb.2 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                        data = dm.lands.200kb, weights = varPower(0, ~thetaC|chr), cor = corAR1(0, ~bin|c

g.div.dm.200kb.3 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                        data = dm.lands.200kb, weights = varPower(0, ~thetaC|chr), method = "ML")

g.div.dm.200kb.4 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                        data = dm.lands.200kb, weights = varPower(0, ~tmrcaC|chr), method = "ML")

AIC(g.div.dm.200kb.1, g.div.dm.200kb.2, g.div.dm.200kb.3, g.div.dm.200kb.4)
```

```
##                  df       AIC
## g.div.dm.200kb.1 11 -5583.726
## g.div.dm.200kb.2 11 -5586.917
## g.div.dm.200kb.3 10 -5561.165
## g.div.dm.200kb.4 10 -5559.348
```

```
summary(g.div.dm.200kb.1)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.200kb
##         AIC       BIC   logLik
##   -5583.726 -5540.918 2802.863
##
```

```
## Correlation Structure: AR(1)
##  Formula: ~bin | chr
##  Parameter estimate(s):
##       Phi
## 0.2843776
## Variance function:
##  Structure: Power of variance covariate, different strata
##  Formula: ~tmrcaC | chr
##  Parameter estimates:
##           2L           2R           3L           3R
## -0.041143168  0.074096129  0.001953365  0.075769644
##
## Coefficients:
##                   Value  Std.Error   t-value p-value
## (Intercept)   0.0085528 0.00000826 1035.3856  0.0000
## thetaC        0.9746365 0.00308267  316.1668  0.0000
## rhoC          0.0005062 0.00134208    0.3772  0.7062
## tmrcaC        0.0111855 0.00019829   56.4105  0.0000
## thetaC:tmrcaC 0.9085940 0.04279033   21.2336  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.133
## rhoC          -0.012  0.189
## tmrcaC        -0.315 -0.455 -0.439
## thetaC:tmrcaC -0.440 -0.196 -0.068  0.727
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.9288491 -0.5661772  0.1198220  0.6423762  3.2163952
##
## Residual standard error: 0.0001234783
## Degrees of freedom: 362 total; 357 residual
```

```
vif(g.div.dm.200kb.1)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.344037      1.505614      3.925969      2.712069
```

```
# Linear model without TMRCA --> rho becomes significant
g.div.dm.200kb.5 <- gls(diversity ~ (thetaC + rhoC),
                        data = dm.lands.200kb, weights = varPower(0, ~thetaC), cor = corAR1(0, ~bin), me

summary(g.div.dm.200kb.5)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC)
##   Data: dm.lands.200kb
##         AIC       BIC   logLik
##   -4723.291 -4699.941 2367.645
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi
```

```
## 0.4168254
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~thetaC
##  Parameter estimates:
##   power
## 0.13171
##
## Coefficients:
##                 Value   Std.Error   t-value p-value
## (Intercept) 0.0086889 0.000030456 285.29066       0
## thetaC      1.0809727 0.011736952  92.09995       0
## rhoC        0.0472191 0.003846123  12.27706       0
##
##  Correlation:
##        (Intr) thetaC
## thetaC -0.016
## rhoC    0.048 -0.145
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -5.445432702 -0.523327124 -0.009243674  0.572229138  2.972069433
##
## Residual standard error: 0.0009029211
## Degrees of freedom: 362 total; 359 residual
```

3.3 1 Mb windows

```
# Chr 2L

# recombination landscapes
rho.dm.1Mb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.rho.1Mb.bedgraph", header = T)

# diversity
diversity.dm.1Mb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.diversity.1Mb.bedgraph", header = T
diversity.dm.1Mb.2L$avg <- apply(diversity.dm.1Mb.2L[4:ncol(diversity.dm.1Mb.2L)], 1, mean)

# mutation landscapes
theta.dm.1Mb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.theta.1Mb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.1Mb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.TMRCA.1Mb.bedgraph", header = T)
tmrca.dm.1Mb.2L$sample_mean <- apply(tmrca.dm.1Mb.2L[4:ncol(tmrca.dm.1Mb.2L)], 1, mean)

# missing data
missing.prop.1Mb.2L <- read.table("dm_data/dm_chr_maps/2L/dm_30x5x5.missing.prop.1Mb.bedgraph", header =
intersect.1Mb.2L <- apply(missing.prop.1Mb.2L[4:ncol(missing.prop.1Mb.2L)], 1, function(x) any(x > 0.1)

dm.lands.1Mb.2L <- as.data.frame(cbind(diversity.dm.1Mb.2L$chromStart,
                                       diversity.dm.1Mb.2L$chromEnd,
                                       diversity.dm.1Mb.2L$avg,
                                       theta.dm.1Mb.2L$sample_mean,
                                       rho.dm.1Mb.2L$sample_mean,
                                       tmrca.dm.1Mb.2L$sample_mean))
```

```r
names(dm.lands.1Mb.2L) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.1Mb.2L$bin <- 1:nrow(dm.lands.1Mb.2L)

# filters based on missing data
dm.lands.1Mb.2L <- dm.lands.1Mb.2L[which(intersect.1Mb.2L == F),]

dm.lands.1Mb.2L$chr <- "2L"

dm.lands.1Mb.2L$thetaC <- dm.lands.1Mb.2L$theta - mean(dm.lands.1Mb.2L$theta)
dm.lands.1Mb.2L$tmrcaC <- dm.lands.1Mb.2L$tmrca - mean(dm.lands.1Mb.2L$tmrca)
dm.lands.1Mb.2L$rhoC <- dm.lands.1Mb.2L$rho - mean(dm.lands.1Mb.2L$rho)

g.div.dm.1Mb.2L <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                       data = dm.lands.1Mb.2L, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), m

summary(g.div.dm.1Mb.2L)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.1Mb.2L
##         AIC       BIC   logLik
##   -299.3268 -291.7713 157.6634
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~bin
##  Parameter estimate(s):
##      Phi1
## 0.3144048
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##       power
## -0.04167166
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0101124 0.00002513 402.3734  0.0000
## thetaC        0.9865547 0.02033816  48.5076  0.0000
## rhoC          0.0010558 0.00560117   0.1885  0.8532
## tmrcaC        0.0129654 0.00122549  10.5798  0.0000
## thetaC:tmrcaC 1.8344985 0.31522351   5.8197  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.006
## rhoC           0.168  0.107
## tmrcaC        -0.247 -0.591 -0.594
## thetaC:tmrcaC -0.447  0.063 -0.394  0.584
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.4305539 -0.5991356 -0.1638086  0.5088896  2.4443876
##
```

```
## Residual standard error: 5.223258e-05
## Degrees of freedom: 19 total; 14 residual
```

```r
# Chr 2R

# recombination landscapes
rho.dm.1Mb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.rho.1Mb.bedgraph", header = T)

# diversity
diversity.dm.1Mb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.diversity.1Mb.bedgraph", header = T
diversity.dm.1Mb.2R$avg <- apply(diversity.dm.1Mb.2R[4:ncol(diversity.dm.1Mb.2R)], 1, mean)

# mutation landscapes
theta.dm.1Mb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.theta.1Mb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.1Mb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.TMRCA.1Mb.bedgraph", header = T)
tmrca.dm.1Mb.2R$sample_mean <- apply(tmrca.dm.1Mb.2R[4:ncol(tmrca.dm.1Mb.2R)], 1, mean)

# missing data
missing.prop.1Mb.2R <- read.table("dm_data/dm_chr_maps/2R/dm_30x5x5.missing.prop.1Mb.bedgraph", header =
intersect.1Mb.2R <- apply(missing.prop.1Mb.2R[4:ncol(missing.prop.1Mb.2R)], 1, function(x) any(x > 0.1))

dm.lands.1Mb.2R <- as.data.frame(cbind(diversity.dm.1Mb.2R$chromStart,
                                       diversity.dm.1Mb.2R$chromEnd,
                                       diversity.dm.1Mb.2R$avg,
                                       theta.dm.1Mb.2R$sample_mean,
                                       rho.dm.1Mb.2R$sample_mean,
                                       tmrca.dm.1Mb.2R$sample_mean))
names(dm.lands.1Mb.2R) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.1Mb.2R$bin <- 1:nrow(dm.lands.1Mb.2R)

# filters based on missing data
dm.lands.1Mb.2R <- dm.lands.1Mb.2R[which(intersect.1Mb.2R == F),]

dm.lands.1Mb.2R$chr <- "2R"

dm.lands.1Mb.2R$thetaC <- dm.lands.1Mb.2R$theta- mean(dm.lands.1Mb.2R$theta)
dm.lands.1Mb.2R$tmrcaC <- dm.lands.1Mb.2R$tmrca - mean(dm.lands.1Mb.2R$tmrca)
dm.lands.1Mb.2R$rhoC <- dm.lands.1Mb.2R$rho - mean(dm.lands.1Mb.2R$rho)

g.div.dm.1Mb.2R <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                       data = dm.lands.1Mb.2R, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), n

summary(g.div.dm.1Mb.2R)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.1Mb.2R
##         AIC       BIC   logLik
##   -260.3174 -253.6517 138.1587
##
## Correlation Structure: AR(1)
##  Formula: ~bin
##  Parameter estimate(s):
```

```
##        Phi
## -0.0218456
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##     power
## -0.286626
##
## Coefficients:
##                   Value Std.Error  t-value p-value
## (Intercept)   0.0091552 0.0000283 323.2402  0.0000
## thetaC        0.9718147 0.0212995  45.6262  0.0000
## rhoC          0.0083025 0.0084990   0.9769  0.3479
## tmrcaC        0.0103641 0.0019099   5.4266  0.0002
## thetaC:tmrcaC 0.5202501 0.5098563   1.0204  0.3277
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.233
## rhoC           0.302 -0.140
## tmrcaC        -0.353 -0.638 -0.213
## thetaC:tmrcaC -0.803  0.231 -0.355  0.536
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.4712709 -0.4619228 -0.3324611  0.4352917  2.9199214
##
## Residual standard error: 1.478574e-05
## Degrees of freedom: 17 total; 12 residual
```

```
# recombination landscapes
rho.dm.1Mb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.rho.1Mb.bedgraph", header = T)

# diversity
diversity.dm.1Mb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.diversity.1Mb.bedgraph", header = T
diversity.dm.1Mb.3L$avg <- apply(diversity.dm.1Mb.3L[4:ncol(diversity.dm.1Mb.3L)], 1, mean)

# mutation landscapes
theta.dm.1Mb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.theta.1Mb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.1Mb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.TMRCA.1Mb.bedgraph", header = T)
tmrca.dm.1Mb.3L$sample_mean <- apply(tmrca.dm.1Mb.3L[4:ncol(tmrca.dm.1Mb.3L)], 1, mean)

# missing data
missing.prop.1Mb.3L <- read.table("dm_data/dm_chr_maps/3L/dm_30x5x5.missing.prop.1Mb.bedgraph", header =
intersect.1Mb.3L <- apply(missing.prop.1Mb.3L[4:ncol(missing.prop.1Mb.3L)], 1, function(x) any(x > 0.1)

dm.lands.1Mb.3L <- as.data.frame(cbind(diversity.dm.1Mb.3L$chromStart,
                                       diversity.dm.1Mb.3L$chromEnd,
                                       diversity.dm.1Mb.3L$avg,
                                       theta.dm.1Mb.3L$sample_mean,
                                       rho.dm.1Mb.3L$sample_mean,
```

```
                                        tmrca.dm.1Mb.3L$sample_mean))
names(dm.lands.1Mb.3L) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.1Mb.3L$bin <- 1:nrow(dm.lands.1Mb.3L)

# filters based on missing data
dm.lands.1Mb.3L <- dm.lands.1Mb.3L[which(intersect.1Mb.3L == F),]

dm.lands.1Mb.3L$chr <- "3L"

dm.lands.1Mb.3L$thetaC <- dm.lands.1Mb.3L$theta- mean(dm.lands.1Mb.3L$theta)
dm.lands.1Mb.3L$tmrcaC <- dm.lands.1Mb.3L$tmrca - mean(dm.lands.1Mb.3L$tmrca)
dm.lands.1Mb.3L$rhoC <- dm.lands.1Mb.3L$rho - mean(dm.lands.1Mb.3L$rho)


g.div.dm.1Mb.3L <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                        data = dm.lands.1Mb.3L, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), n

summary(g.div.dm.1Mb.3L)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.1Mb.3L
##         AIC       BIC   logLik
##   -348.8084 -340.0801 182.4042
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~bin
##  Parameter estimate(s):
##     Phi1
## 0.249026
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##     power
## 0.3377385
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0091813 0.00002405 381.7940  0.0000
## thetaC        1.0032652 0.01244385  80.6233  0.0000
## rhoC          0.0070903 0.00396199   1.7896  0.0913
## tmrcaC        0.0084717 0.00094627   8.9528  0.0000
## thetaC:tmrcaC 0.3950013 0.16662491   2.3706  0.0298
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.400
## rhoC           0.531  0.320
## tmrcaC        -0.622 -0.738 -0.571
## thetaC:tmrcaC -0.618 -0.348 -0.438  0.689
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
```

```
## -2.71883315 -0.45696303  0.01830762  0.69607904  1.69584501
##
## Residual standard error: 0.0002107777
## Degrees of freedom: 22 total; 17 residual
```

```
# Chr 3R

# recombination landscapes
rho.dm.1Mb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.rho.1Mb.bedgraph", header = T)

# diversity
diversity.dm.1Mb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.diversity.1Mb.bedgraph", header = T
diversity.dm.1Mb.3R$avg <- apply(diversity.dm.1Mb.3R[4:ncol(diversity.dm.1Mb.3R)], 1, mean)

# mutation landscapes
theta.dm.1Mb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.theta.1Mb.bedgraph", header = T)

# TMRCA landscapes
tmrca.dm.1Mb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.TMRCA.1Mb.bedgraph", header = T)
tmrca.dm.1Mb.3R$sample_mean <- apply(tmrca.dm.1Mb.3R[4:ncol(tmrca.dm.1Mb.3R)], 1, mean)

# missing data
missing.prop.1Mb.3R <- read.table("dm_data/dm_chr_maps/3R/dm_30x5x5.missing.prop.1Mb.bedgraph", header =
intersect.1Mb.3R <- apply(missing.prop.1Mb[4:ncol(missing.prop.1Mb.3R)], 1, function(x) any(x > 0.1))

dm.lands.1Mb.3R <- as.data.frame(cbind(diversity.dm.1Mb.3R$chromStart,
                                       diversity.dm.1Mb.3R$chromEnd,
                                       diversity.dm.1Mb.3R$avg,
                                       theta.dm.1Mb.3R$sample_mean,
                                       rho.dm.1Mb.3R$sample_mean,
                                       tmrca.dm.1Mb.3R$sample_mean))
names(dm.lands.1Mb.3R) <- c("start", "end", "diversity", "theta", "rho", "tmrca")
dm.lands.1Mb.3R$bin <- 1:nrow(dm.lands.1Mb.3R)

# filters based on missing data
dm.lands.1Mb.3R <- dm.lands.1Mb.3R[which(intersect.1Mb.3R == F),]

dm.lands.1Mb.3R$chr <- "3R"

dm.lands.1Mb.3R$thetaC <- dm.lands.1Mb.3R$theta- mean(dm.lands.1Mb.3R$theta)
dm.lands.1Mb.3R$tmrcaC <- dm.lands.1Mb.3R$tmrca - mean(dm.lands.1Mb.3R$tmrca)
dm.lands.1Mb.3R$rhoC <- dm.lands.1Mb.3R$rho - mean(dm.lands.1Mb.3R$rho)

g.div.dm.1Mb.3R <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                       data = dm.lands.1Mb.3R, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), m

summary(g.div.dm.1Mb.3R)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.1Mb.3R
##         AIC       BIC   logLik
##   -324.5017 -316.9462 170.2509
##
## Correlation Structure: ARMA(1,0)
```

```
##  Formula: ~bin
##  Parameter estimate(s):
##       Phi1
## -0.4823527
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~tmrcaC
##  Parameter estimates:
##     power
## 0.8128992
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0067246 0.00002252 298.66684  0.0000
## thetaC        0.9664707 0.00658600 146.74630  0.0000
## rhoC          0.0031144 0.00241054   1.29200  0.2173
## tmrcaC        0.0094832 0.00049711  19.07661  0.0000
## thetaC:tmrcaC 0.7457601 0.17910163   4.16389  0.0010
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.911
## rhoC          -0.280  0.066
## tmrcaC        -0.923 -0.944 -0.023
## thetaC:tmrcaC -0.940 -0.881  0.320  0.856
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.7308894 -0.4913490  0.1058365  0.6374128  1.9989782
##
## Residual standard error: 0.0004245867
## Degrees of freedom: 19 total; 14 residual
```

```
# all together now
dm.lands.1Mb <- rbind.data.frame(dm.lands.1Mb.2L, dm.lands.1Mb.2R, dm.lands.1Mb.3L, dm.lands.1Mb.3R)

write.table(dm.lands.1Mb, "dm_data/dm_chr_maps/dm.maps.1Mb.tsv", quote = F, sep = "\t", row.names = F, 


# Plots
scale.3d <- function(x) sprintf("%.3f", x) # digits shown in y axis

setTimeLimit(cpu = Inf, elapsed = Inf, transient = F)

molten.diversity <- melt(dm.lands.1Mb[c(3,7,8)], id.vars = c("bin", "chr"))
diversity.map <- ggplot(data = molten.diversity, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
diversity.map <- diversity.map + geom_line(data = molten.diversity, colour = "#F8766D")
diversity.map <- diversity.map + geom_smooth(method = "loess", se = F, colour = "#F8766D")
diversity.map <- diversity.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breal
diversity.map <- diversity.map + labs(title = NULL, x = NULL, y = expression(pi))
diversity.map <- diversity.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_te
diversity.map
```

```
molten.rho <- melt(dm.lands.1Mb[c(5,7,8)], id.vars = c("bin", "chr"))
rho.map <- ggplot(data = molten.rho, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
rho.map <- rho.map + geom_line(data = molten.rho, colour = "#7CAE00")
rho.map <- rho.map + geom_smooth(method = "loess", se = F, colour = "#7CAE00")
rho.map <- rho.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks = pretty_
rho.map <- rho.map + labs(title = NULL, x = NULL, y = expression(rho))
rho.map <- rho.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size = 2
rho.map
```

```
molten.theta <- melt(dm.lands.1Mb[c(4,7,8)], id.vars = c("bin", "chr"))
theta.map <- ggplot(data = molten.theta, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
theta.map <- theta.map + geom_line(data = molten.theta, colour = "#00BFC4")
theta.map <- theta.map + geom_smooth(method = "loess", se = F, colour = "#00BFC4")
theta.map <- theta.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks = pre
theta.map <- theta.map + labs(title = NULL, x = NULL, y = expression(theta))
theta.map <- theta.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size
theta.map
```

```
molten.tmrca <- melt(dm.lands.1Mb[c(6,7,8)], id.vars = c("bin", "chr"))
tmrca.map <- ggplot(data = molten.tmrca, aes(x = bin * 50, y = value)) + facet_wrap(~chr)
tmrca.map <- tmrca.map + geom_line(data = molten.tmrca, colour = "#88419d")
tmrca.map <- tmrca.map + geom_smooth(method = "loess", se = F, colour = "#88419d")
tmrca.map <- tmrca.map + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuous(breaks = pre
tmrca.map <- tmrca.map + labs(title = NULL, x = "Position (kb)", y = expression(tau))
tmrca.map <- tmrca.map + theme(axis.title.x = element_text(size = 20), axis.title.y = element_text(size
tmrca.map
```

```
# genome-wide correlations
cor.test(~rho + diversity, data = dm.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and diversity
## S = 60404, p-value = 0.07236
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.2060045
```

```
cor.test(~rho + tmrca, data = dm.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 42930, p-value = 8.798e-05
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.4356959
```

```
cor.test(~theta + tmrca, data = dm.lands.1Mb, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
```

```
##
## data:  theta and tmrca
## S = 22964, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.698144
```

```r
# Linear models
# centering
dm.lands.1Mb$thetaC <- dm.lands.1Mb$theta - mean(dm.lands.1Mb$theta)
dm.lands.1Mb$tmrcaC <- dm.lands.1Mb$tmrca - mean(dm.lands.1Mb$tmrca)
dm.lands.1Mb$rhoC <- dm.lands.1Mb$rho - mean(dm.lands.1Mb$rho)

dm.lands.1Mb$bin <- 1:nrow(dm.lands.1Mb)

m.div.dm.1Mb <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC), data = dm.lands.1Mb)

plot(resid(m.div.dm.1Mb)~fitted(m.div.dm.1Mb))
```

```r
hist(resid(m.div.dm.1Mb), nclass = 30)
```

```
dwtest(m.div.dm.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.dm.1Mb
## DW = 1.6563, p-value = 0.03648
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.dm.1Mb, nsim = 10000)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.dm.1Mb
## HMC = 0.72492, p-value = 0.999
```

```
summary(m.div.dm.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
##     data = dm.lands.1Mb)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.268e-04 -5.063e-05 -1.004e-05  4.519e-05  2.656e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   8.808e-03  1.551e-05 568.036  < 2e-16 ***
```

```
## thetaC        9.919e-01  6.725e-03 147.497  < 2e-16 ***
## rhoC          6.510e-03  3.034e-03   2.145   0.0353 *
## tmrcaC        9.525e-03  6.035e-04  15.784  < 2e-16 ***
## thetaC:tmrcaC 4.881e-01  1.131e-01   4.315 4.99e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.283e-05 on 72 degrees of freedom
## Multiple R-squared:  0.999,  Adjusted R-squared:  0.9989
## F-statistic: 1.74e+04 on 4 and 72 DF,  p-value: < 2.2e-16
```

```
# type 2 ANOVA
anova.diversity <- Anova(m.div.dm.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
anova.diversity
```

```
## Anova Table (Type II tests)
##
## Response: diversity
##                    Sum Sq Df    F value   Pr(>F)  VarExp
## thetaC        2.4658e-04  1 28614.4891 0.000000 0.97371
## rhoC          4.0000e-08  1     4.6028 0.035293 0.00016
## tmrcaC        5.8360e-06  1   677.2706 0.000000 0.02305
## thetaC:tmrcaC 1.6000e-07  1    18.6226 0.000050 0.00063
## Residuals     6.2000e-07 72                     0.00245
```

```
# GLS
g.div.dm.1Mb.1 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                      data = dm.lands.1Mb, weights = varPower(0, ~tmrcaC|chr), cor = corAR1(0, ~bin|chr

g.div.dm.1Mb.2 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                      data = dm.lands.1Mb, weights = varPower(0, ~theta|chr), cor = corAR1(0, ~bin|chr

g.div.dm.1Mb.3 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                      data = dm.lands.1Mb, weights = varPower(0, ~theta|chr), method = "ML")

g.div.dm.1Mb.4 <- gls(diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC),
                      data = dm.lands.1Mb, weights = varPower(0, ~tmrcaC|chr), method = "ML")

AIC(g.div.dm.1Mb.1, g.div.dm.1Mb.2, g.div.dm.1Mb.3, g.div.dm.1Mb.4)
```

```
##                df       AIC
## g.div.dm.1Mb.1 11 -1228.719
## g.div.dm.1Mb.2 11 -1202.198
## g.div.dm.1Mb.3 10 -1202.415
## g.div.dm.1Mb.4 10 -1222.559
```

```
summary(g.div.dm.1Mb.1)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC + tmrcaC + thetaC:tmrcaC)
##   Data: dm.lands.1Mb
##         AIC       BIC   logLik
##   -1228.719 -1202.937 625.3593
##
```

```
## Correlation Structure: AR(1)
##  Formula: ~bin | chr
##  Parameter estimate(s):
##       Phi
## 0.4040378
## Variance function:
##  Structure: Power of variance covariate, different strata
##  Formula: ~tmrcaC | chr
##  Parameter estimates:
##        2L        2R        3L        3R
## 0.3814217 0.3225539 0.4579989 0.5841943
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0087611 0.00000890 984.6352  0.0000
## thetaC        0.9830140 0.00395599 248.4877  0.0000
## rhoC          0.0012460 0.00207145   0.6015  0.5494
## tmrcaC        0.0103831 0.00051388  20.2055  0.0000
## thetaC:tmrcaC 0.6061715 0.12215915   4.9621  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.225
## rhoC           0.466  0.603
## tmrcaC        -0.364 -0.486 -0.568
## thetaC:tmrcaC -0.286 -0.210 -0.299  0.636
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.6371469 -0.3675723  0.2903415  0.9878832  2.4854661
##
## Residual standard error: 0.0004263997
## Degrees of freedom: 77 total; 72 residual
```

```
vif(g.div.dm.1Mb.1)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.674634      1.863553      2.423500      1.719226
```

```
# Linear model without TMRCA --> rho becomes significant
g.div.dm.1Mb.5 <- gls(diversity ~ (thetaC + rhoC),
                      data = dm.lands.1Mb, weights = varPower(0, ~thetaC|chr), cor = corAR1(0, ~bin|chr

summary(g.div.dm.1Mb.5)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ (thetaC + rhoC)
##   Data: dm.lands.1Mb
##         AIC   BIC   logLik
##   -1055.094 -1034 536.5469
##
## Correlation Structure: AR(1)
##  Formula: ~bin | chr
##  Parameter estimate(s):
##       Phi
```

```
## 0.5243617
## Variance function:
##  Structure: Power of variance covariate, different strata
##  Formula: ~thetaC | chr
##  Parameter estimates:
##        2L        2R        3L        3R
## 0.2518860 0.2065555 0.2564038 0.1788172
##
## Coefficients:
##                 Value   Std.Error   t-value p-value
## (Intercept) 0.0089103 0.000042565 209.33432       0
## thetaC      1.1056720 0.018207498  60.72619       0
## rhoC        0.0395498 0.006816697   5.80191       0
##
##  Correlation:
##        (Intr) thetaC
## thetaC -0.115
## rhoC    0.055 -0.292
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.46396737 -0.69180396 -0.06437384  0.58736967  1.68761843
##
## Residual standard error: 0.001151706
## Degrees of freedom: 77 total; 74 residual
```

3.4 Evolutionary (Protein) Rates

This is how we generated the table from raw data:

```r
# loads data
dm.raw <- read.table("dm_data/dm_misc/dpgp2_Dsimulans_bpp.all.csv", header = TRUE, sep = "\t") %>%
          select(GeneID, MeanNumberSynPos, PiN, PiS, MeanNumberSynPosDiv, dS, dN) %>% na.omit()

# summarises
dm.tbl <- ddply(dm.raw, .variables = "GeneID", .fun = plyr::summarize,
                MeanNumberNonSynPos = sum(3 - MeanNumberSynPos),
                MeanNumberSynPos = sum(MeanNumberSynPos),
                PiN = sum(PiN),
                PiS = sum(PiS),
                MeanNumberNonSynPosDiv = sum(3 - MeanNumberSynPosDiv),
                MeanNumberSynPosDiv = sum(MeanNumberSynPosDiv),
                dN = sum(dN),
                dS = sum(dS))

dm.tbl$PiN <- (dm.tbl$PiN / dm.tbl$MeanNumberNonSynPos)
dm.tbl$PiS <- (dm.tbl$PiS / dm.tbl$MeanNumberSynPos)
dm.tbl$dN <- (dm.tbl$dN / dm.tbl$MeanNumberNonSynPosDiv)
dm.tbl$dS <- (dm.tbl$dS / dm.tbl$MeanNumberSynPosDiv)

# computes ratios
dm.tbl$PiN_PiS <- with(dm.tbl, PiN / PiS)
dm.tbl$dN_dS <- with(dm.tbl, dN / dS)

# we ignore genes with ratios above 1
```

```
dm.tbl <- dm.tbl[dm.tbl$PiN_PiS < 1,]
dm.tbl <- dm.tbl[dm.tbl$dN_dS < 1,]
dm.tbl <- na.omit(dm.tbl)

dm.genes.coord <- read.table("dm_data/dm_misc/Dmel_trans_nooverlaps-final.txt", header = F) %>% select(V
names(dm.genes.coord) <- c("chr", "start", "end", "GeneID")

dm.evol <- merge(dm.genes.coord, dm.tbl, by = "GeneID")
dm.evol <- arrange(dm.evol, chr, start)

# reorder maps to use GR
dm.maps.50kb <- dm.lands.50kb[,c(8,1:6)]

# mapping genes to 50kb windows in ismc maps
dm.lands.gr <- makeGRangesFromDataFrame(dm.maps.50kb)
values(dm.lands.gr) <- select(dm.maps.50kb, diversity, theta, rho, tmrca)
evolrate.gr <- makeGRangesFromDataFrame(dm.evol)
values(evolrate.gr) <- select(dm.evol, PiN, PiS, dN, dS)

hits_1 <- findOverlaps(query = evolrate.gr, subject = dm.lands.gr, type = "within")

evolrate.gr.df <- as.data.frame(evolrate.gr[queryHits(hits_1)], row.names = NULL)
dm.lands.gr.df <- as.data.frame(dm.lands.gr[subjectHits(hits_1)], row.names = NULL)

dm.lands.evolrate <- cbind.data.frame(dm.lands.gr.df[,c(1:3,6:9)], evolrate.gr.df[,c(2:4,6:9)])

names(dm.lands.evolrate)[1] <- "chr"
names(dm.lands.evolrate)[2] <- "start.window"
names(dm.lands.evolrate)[3] <- "end.window"
names(dm.lands.evolrate)[8] <- "start.gene"
names(dm.lands.evolrate)[9] <- "end.gene"
names(dm.lands.evolrate)[10] <- "length.gene"

dm.lands.evolrate <- arrange(dm.lands.evolrate, chr, start.window, start.gene)

dm.lands.evolrate.bins <- ddply(dm.lands.evolrate, .variables = c("chr", "start.window"), .fun = plyr::s
                                diversity = mean(diversity),
                                theta = mean(theta),
                                rho = mean(rho),
                                tmrca = mean(tmrca),
                                exonic.sites = sum(length.gene),
                                PiN.window = sum(PiN*length.gene) / sum(length.gene),
                                PiS.window = sum(PiS*length.gene) / sum(length.gene),
                                dN.window = sum(dN*length.gene) / sum(length.gene),
                                dS.window = sum(dS*length.gene) / sum(length.gene))


dm.lands.evolrate.bins$bin <- 1:nrow(dm.lands.evolrate.bins)

write.table(dm.lands.evolrate.bins, "dm_data/dm_chr_maps/dm_maps_50kb_protein_rates.tsv",
            sep = "\t", quote = F, col.names = T, row.names = F)
```

We can read the table generated and move on with the analyses:

```r
dm.lands.evolrate.bins <- read.table("dm_data/dm_chr_maps/dm_maps_50kb_protein_rates.tsv", header = T, s

# linear model in windows with more than 20kb exonic sites
dm.lands.genes <- filter(dm.lands.evolrate.bins, exonic.sites > 2e+4)

# centering
dm.lands.genes$thetaC <- dm.lands.genes$theta - mean(dm.lands.genes$theta)
dm.lands.genes$tmrcaC <- dm.lands.genes$tmrca - mean(dm.lands.genes$tmrca)
dm.lands.genes$rhoC <- dm.lands.genes$rho - mean(dm.lands.genes$rho)

m.dm.cds.1 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC*tmrcaC), data = dm.lands.genes)
m.dm.cds.2 <- lm(diversity ~ (thetaC + rhoC + tmrcaC + thetaC*tmrcaC + rhoC*tmrcaC), data = dm.lands.gen

AIC(m.dm.cds.1, m.dm.cds.2)
```

```
##            df       AIC
## m.dm.cds.1  6 -4778.177
## m.dm.cds.2  7 -4782.425
```

```r
summary(m.dm.cds.2)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC * tmrcaC +
##     rhoC * tmrcaC), data = dm.lands.genes)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.723e-04 -9.596e-05  3.810e-06  8.994e-05  6.661e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   7.890e-03  9.716e-06 812.013   <2e-16 ***
## thetaC        9.511e-01  3.665e-03 259.551   <2e-16 ***
## rhoC          1.053e-03  1.530e-03   0.688   0.4919
## tmrcaC        1.120e-02  1.964e-04  57.010   <2e-16 ***
## thetaC:tmrcaC 1.079e+00  4.299e-02  25.098   <2e-16 ***
## rhoC:tmrcaC   2.110e-02  8.481e-03   2.488   0.0133 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000149 on 318 degrees of freedom
## Multiple R-squared:  0.9977, Adjusted R-squared:  0.9977
## F-statistic: 2.785e+04 on 5 and 318 DF,  p-value: < 2.2e-16
```

```r
# picking best OLS for ANOVA
plot(resid(m.dm.cds.2)~fitted(m.dm.cds.2))
```

```
hist(resid(m.dm.cds.2), nclass = 30)
```

```
dwtest(m.dm.cds.2)
```

```
##
##  Durbin-Watson test
##
## data:  m.dm.cds.2
## DW = 2.0035, p-value = 0.4719
```

```
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.dm.cds.2, nsim = 10000)
```

```
##
##  Harrison-McCabe test
##
## data:  m.dm.cds.2
## HMC = 0.41117, p-value = 0.0128
```

```r
summary(m.dm.cds.2)
```

```
##
## Call:
## lm(formula = diversity ~ (thetaC + rhoC + tmrcaC + thetaC * tmrcaC +
##     rhoC * tmrcaC), data = dm.lands.genes)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.723e-04 -9.596e-05  3.810e-06  8.994e-05  6.661e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   7.890e-03  9.716e-06 812.013   <2e-16 ***
## thetaC        9.511e-01  3.665e-03 259.551   <2e-16 ***
## rhoC          1.053e-03  1.530e-03   0.688   0.4919
## tmrcaC        1.120e-02  1.964e-04  57.010   <2e-16 ***
## thetaC:tmrcaC 1.079e+00  4.299e-02  25.098   <2e-16 ***
## rhoC:tmrcaC   2.110e-02  8.481e-03   2.488   0.0133 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000149 on 318 degrees of freedom
## Multiple R-squared:  0.9977, Adjusted R-squared:  0.9977
## F-statistic: 2.785e+04 on 5 and 318 DF,  p-value: < 2.2e-16
```

```r
# type 2 anova
anova.diversity.cds <- Anova(m.dm.cds.2)
apiss <- anova.diversity.cds$"Sum Sq"
anova.diversity.cds$VarExp <- apiss / sum(apiss)

anova.diversity.cds
```

```
## Anova Table (Type II tests)
##
## Response: diversity
##                  Sum Sq  Df    F value   Pr(>F)  VarExp
## thetaC        1.544e-03   1 69536.4716 0.000000 0.93236
## rhoC          5.000e-08   1     2.1199 0.146386 0.00003
## tmrcaC        9.078e-05   1  4088.1586 0.000000 0.05482
## thetaC:tmrcaC 1.399e-05   1   629.9096 0.000000 0.00845
## rhoC:tmrcaC   1.400e-07   1     6.1917 0.013347 0.00008
## Residuals     7.060e-06 318                     0.00426
```

```r
# simple correlations
cor.test(dm.lands.evolrate.bins$theta, dm.lands.evolrate.bins$exonic.sites, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  dm.lands.evolrate.bins$theta and dm.lands.evolrate.bins$exonic.sites
## S = 339155743, p-value = 6.388e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.1289611
```

```
cor.test(dm.lands.evolrate.bins$tmrca, dm.lands.evolrate.bins$exonic.sites, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  dm.lands.evolrate.bins$tmrca and dm.lands.evolrate.bins$exonic.sites
## S = 360421366, p-value = 2.034e-12
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## -0.1997488
```

```
pcor.test(dm.lands.evolrate.bins$tmrca, dm.lands.evolrate.bins$exonic.sites, dm.lands.evolrate.bins$the
```

```
##     estimate      p.value statistic    n gp   Method
## 1 -0.1582194 2.912866e-08 -5.583083 1217  1 spearman
```

```
pcor.test(dm.lands.evolrate.bins$theta, dm.lands.evolrate.bins$exonic.sites, dm.lands.evolrate.bins$tmr
```

```
##     estimate   p.value statistic    n gp   Method
## 1 -0.0374759 0.1915716 -1.306672 1217  1 spearman
```

```
cor.test(dm.lands.evolrate.bins$PiN, dm.lands.evolrate.bins$theta, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  dm.lands.evolrate.bins$PiN and dm.lands.evolrate.bins$theta
## S = 192322126, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.3598097
```

```
cor.test(dm.lands.evolrate.bins$PiS, dm.lands.evolrate.bins$theta, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  dm.lands.evolrate.bins$PiS and dm.lands.evolrate.bins$theta
## S = 78488708, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.7387315
```

```
pcor.test(dm.lands.evolrate.bins$PiN, dm.lands.evolrate.bins$theta, dm.lands.evolrate.bins$tmrca, metho
```

```
##    estimate      p.value statistic    n gp   Method
## 1 0.2711637 6.149318e-22  9.815788 1217  1 spearman
```

```
pcor.test(dm.lands.evolrate.bins$PiS, dm.lands.evolrate.bins$theta, dm.lands.evolrate.bins$tmrca, metho
```

```
##    estimate       p.value statistic    n gp   Method
## 1 0.6767774 1.129784e-163  32.03071 1217  1 spearman
```

```
# bootstrapping PiN and PiS correlations
boot.tbl <- dm.lands.evolrate.bins %>% dplyr::select(PiN.window, theta, PiS.window, tmrca)
boot.PiN.theta <- boot(boot.tbl,
  statistic = function(boot.tbl, i) {
    pcor.test(boot.tbl[i, "PiN.window"], boot.tbl[i, "theta"], boot.tbl[i, "tmrca"], method='spearman')$
  },
  R = 10000
)
boot.ci(boot.PiN.theta, type = "basic")
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 10000 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = boot.PiN.theta, type = "basic")
##
## Intervals :
## Level      Basic
## 95%   ( 0.2191,  0.3262 )
## Calculations and Intervals on Original Scale
```

```
boot.PiS.theta <- boot(boot.tbl,
  statistic = function(boot.tbl, i) {
    pcor.test(boot.tbl[i, "PiS.window"], boot.tbl[i, "theta"], boot.tbl[i, "tmrca"], method='spearman')$
  },
  R = 10000
)
boot.ci(boot.PiS.theta, type = "basic")
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 10000 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = boot.PiS.theta, type = "basic")
##
## Intervals :
## Level      Basic
## 95%   ( 0.6413,  0.7153 )
## Calculations and Intervals on Original Scale
```

4 R^2 plot with real, simulated (true) and simulated (inferred) data.

```
# true landscapes
r2.sim.avg <- rbind.data.frame(r2.sim.50kb$average, r2.sim.200kb$average, r2.sim.1Mb$average, make.row.n
colnames(r2.sim.avg) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
r2.sim.avg$bin.size <- c(50, 200, 1000)
```

```
# inferred landscapes
r2.inf.avg <- rbind.data.frame(r2.inf.50kb$average, r2.inf.200kb$average, r2.inf.1Mb$average, make.row.
colnames(r2.inf.avg) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
r2.inf.avg$bin.size <- c(50, 200, 1000)

r2.tab.comb <- rbind.data.frame(r2.dm.tab, r2.sim.avg, r2.inf.avg)
r2.tab.comb$type <- c(rep("iSMC-inferred (Drosophila)", 3), rep("True landscapes (simulations)", 3), rep

molten.r2 <- melt(r2.tab.comb, id.vars = c("bin.size", "type"))
r2.coal.plot <- ggplot(data = molten.r2, aes(x = bin.size, y = value, shape = variable))
r2.coal.plot <- r2.coal.plot + geom_line(data = molten.r2)
r2.coal.plot <- r2.coal.plot + geom_point(aes(shape = variable), size = 4) + facet_wrap(~type)
r2.coal.plot <- r2.coal.plot + scale_x_continuous(breaks = c(50, 200, 1000), trans="log10")
r2.coal.plot <- r2.coal.plot + scale_y_continuous(breaks = pretty_breaks())
r2.coal.plot <- r2.coal.plot + scale_shape_manual(values = c(0, 1, 2, 3, 4))
r2.coal.plot <- r2.coal.plot + labs(title = NULL, x = NULL, y = "Variance Explained (%)")
r2.coal.plot <- r2.coal.plot + theme_bw()
r2.coal.plot <- r2.coal.plot + theme(axis.title = element_text(size = 20), axis.text = element_text(size
                                     legend.position = "none")
r2.coal.plot
```

5 Simulations with BGS and flat Mutation Landscape in SLiM

5.1 50 kb scale

```
r2.bgs.50kb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.bgs.50kb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.bgs.50kb) <- reps

sim.theta.50kb <- read.table("dm_bgs_sims/MutationMap_50kb.csv", header = T)
sim.rho.50kb <- read.table("dm_bgs_sims/RecombinationMap_50kb.csv", header = T)
```

5.1.1 Replicate 1

```
rep_1.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.divers
rep_1.pi.50kb$avg <- apply(rep_1.pi.50kb[4:ncol(rep_1.pi.50kb)], 1, mean)
rep_1.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.rho.5
rep_1.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.the
rep_1.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.TMR
rep_1.tmrca.50kb$avg <- apply(rep_1.tmrca.50kb[4:ncol(rep_1.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep1 <- as.data.frame(cbind(rep_1.pi.50kb$avg, rep_1.theta.50kb$sample_mean, rep_1.tmrca
names(inf.lands.50kb.rep1) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep1$thetaC <- inf.lands.50kb.rep1$theta - mean(inf.lands.50kb.rep1$theta)
inf.lands.50kb.rep1$tmrcaC <- inf.lands.50kb.rep1$tmrca - mean(inf.lands.50kb.rep1$tmrca)
inf.lands.50kb.rep1$rhoC <- inf.lands.50kb.rep1$rho - mean(inf.lands.50kb.rep1$rho)

inf.lands.50kb.rep1$bin <- 1:nrow(inf.lands.50kb.rep1)

# for merging:
inf.lands.50kb.rep1 <- inf.lands.50kb.rep1
inf.lands.50kb.rep1$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep1)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep1)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7318.145
## m.div.50kb.2  7 -7451.068
## m.div.50kb.3  8 -7486.698
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.4321, p-value = 1.478e-10
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.52172, p-value = 0.765
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep1)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.947e-04 -6.541e-05 -1.228e-05  5.931e-05  4.683e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.277e-03  4.998e-06 655.616  < 2e-16 ***
## thetaC         1.006e+00  6.702e-03 150.043  < 2e-16 ***
## rhoC          -9.708e-03  2.472e-03  -3.927 9.89e-05 ***
## tmrcaC         4.427e-03  4.633e-05  95.553  < 2e-16 ***
## thetaC:tmrcaC  1.440e+00  4.349e-02  33.119  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001015 on 466 degrees of freedom
## Multiple R-squared:  0.9913, Adjusted R-squared:  0.9912
## F-statistic: 1.33e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```r
r2.bgs.50kb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 1] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep1, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep1, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep1, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep1, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7399.286
## g.div.50kb.2  8 -7355.715
## g.div.50kb.3  7 -7316.900
## g.div.50kb.4  7 -7364.037
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep1
##       AIC       BIC  logLik
##   -7316.9 -7287.816 3665.45
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##      power
## -0.1014917
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0032769 0.00000503 651.9396   0e+00
## thetaC         1.0054351 0.00719805 139.6817   0e+00
## rhoC          -0.0095979 0.00250781  -3.8272   1e-04
## tmrcaC         0.0044287 0.00004609  96.0789   0e+00
## thetaC:tmrcaC  1.4438099 0.04510765  32.0081   0e+00
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.151
## rhoC           0.148 -0.373
## tmrcaC        -0.213 -0.043 -0.406
## thetaC:tmrcaC -0.365  0.276 -0.390  0.587
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.8674339 -0.6482883 -0.1174017  0.5737792  4.7306842
##
## Residual standard error: 5.692921e-05
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.367234      1.469587      1.859346      1.775887
```

5.1.2 Replicate 2

```
rep_2.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.divers
rep_2.pi.50kb$avg <- apply(rep_2.pi.50kb[4:ncol(rep_2.pi.50kb)], 1, mean)
rep_2.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.rho.5
rep_2.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.the
rep_2.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.TMR
rep_2.tmrca.50kb$avg <- apply(rep_2.tmrca.50kb[4:ncol(rep_2.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep2 <- as.data.frame(cbind(rep_2.pi.50kb$avg, rep_2.theta.50kb$sample_mean, rep_2.tmrca
names(inf.lands.50kb.rep2) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep2$thetaC <- inf.lands.50kb.rep2$theta - mean(inf.lands.50kb.rep2$theta)
inf.lands.50kb.rep2$tmrcaC <- inf.lands.50kb.rep2$tmrca - mean(inf.lands.50kb.rep2$tmrca)
inf.lands.50kb.rep2$rhoC <- inf.lands.50kb.rep2$rho - mean(inf.lands.50kb.rep2$rho)

inf.lands.50kb.rep2$bin <- 1:nrow(inf.lands.50kb.rep2)

# for merging:
inf.lands.50kb.rep2 <- inf.lands.50kb.rep2
inf.lands.50kb.rep2$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep2)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep2)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7276.144
## m.div.50kb.2  7 -7410.948
## m.div.50kb.3  8 -7431.395
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.2946, p-value = 3.222e-15
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.54072, p-value = 0.915
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.170e-04 -6.312e-05 -6.800e-06  7.111e-05  4.170e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.259e-03  5.004e-06 651.204   <2e-16 ***
## thetaC         1.011e+00  5.917e-03 170.941   <2e-16 ***
## rhoC          -4.534e-03  2.399e-03  -1.891   0.0593 .
## tmrcaC         4.443e-03  4.698e-05  94.580   <2e-16 ***
## thetaC:tmrcaC  1.398e+00  3.709e-02  37.684   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001062 on 466 degrees of freedom
## Multiple R-squared:  0.9914, Adjusted R-squared:  0.9913
## F-statistic: 1.34e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```r
r2.bgs.50kb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 2] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep2, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep2, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep2, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep2, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7357.344
## g.div.50kb.2  8 -7336.580
## g.div.50kb.3  7 -7288.080
## g.div.50kb.4  7 -7305.879
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep2
##        AIC       BIC  logLik
##   -7288.08 -7258.996 3651.04
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##      power
## -0.4081818
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0032575 0.00000508 641.8553  0.0000
## thetaC         1.0124073 0.00775179 130.6030  0.0000
## rhoC          -0.0040721 0.00246568  -1.6515  0.0993
## tmrcaC         0.0044536 0.00004513  98.6740  0.0000
## thetaC:tmrcaC  1.4273988 0.04122365  34.6257  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.304
## rhoC           0.127 -0.379
## tmrcaC        -0.134  0.038 -0.454
## thetaC:tmrcaC -0.236  0.195 -0.380  0.565
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.52489862 -0.63206070 -0.05459749  0.67564509  4.32924101
##
## Residual standard error: 1.020588e-05
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.229717      1.521322      1.708997      1.550185
```

5.1.3 Replicate 3

```
rep_3.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.divers
rep_3.pi.50kb$avg <- apply(rep_3.pi.50kb[4:ncol(rep_3.pi.50kb)], 1, mean)
rep_3.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.rho.5
rep_3.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.the
rep_3.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.TMR
rep_3.tmrca.50kb$avg <- apply(rep_3.tmrca.50kb[4:ncol(rep_3.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep3 <- as.data.frame(cbind(rep_3.pi.50kb$avg, rep_3.theta.50kb$sample_mean, rep_3.tmrca
names(inf.lands.50kb.rep3) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep3$thetaC <- inf.lands.50kb.rep3$theta - mean(inf.lands.50kb.rep3$theta)
inf.lands.50kb.rep3$tmrcaC <- inf.lands.50kb.rep3$tmrca - mean(inf.lands.50kb.rep3$tmrca)
inf.lands.50kb.rep3$rhoC <- inf.lands.50kb.rep3$rho - mean(inf.lands.50kb.rep3$rho)

inf.lands.50kb.rep3$bin <- 1:nrow(inf.lands.50kb.rep3)

# for merging:
inf.lands.50kb.rep3 <- inf.lands.50kb.rep3
inf.lands.50kb.rep3$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep3)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep3)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7236.648
## m.div.50kb.2  7 -7340.444
## m.div.50kb.3  8 -7339.716
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.1119, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.52223, p-value = 0.788
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.309e-04 -6.593e-05 -3.010e-06  6.406e-05  5.898e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.240e-03  5.846e-06 554.234   <2e-16 ***
## thetaC         9.990e-01  1.080e-02  92.469   <2e-16 ***
## rhoC          -2.159e-02  3.353e-03  -6.439    3e-10 ***
## tmrcaC         4.853e-03  5.411e-05  89.678   <2e-16 ***
## thetaC:tmrcaC  1.564e+00  4.725e-02  33.099   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001107 on 466 degrees of freedom
## Multiple R-squared:  0.9897, Adjusted R-squared:  0.9896
## F-statistic: 1.12e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```
r2.bgs.50kb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 3] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep3, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep3, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep3, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep3, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7359.900
## g.div.50kb.2  8 -7340.098
## g.div.50kb.3  7 -7288.913
## g.div.50kb.4  7 -7274.406
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep3
##         AIC       BIC   logLik
##   -7288.913 -7259.829 3651.457
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.091799
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0032438 0.00000609 532.3525   0e+00
## thetaC         0.9768237 0.01338390  72.9850   0e+00
## rhoC          -0.0119449 0.00322386  -3.7051   2e-04
## tmrcaC         0.0048080 0.00005006  96.0371   0e+00
## thetaC:tmrcaC  1.5193043 0.05851355  25.9650   0e+00
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.529
## rhoC           0.263 -0.501
## tmrcaC        -0.247 -0.034 -0.434
## thetaC:tmrcaC -0.531  0.442 -0.444  0.441
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.59963451 -0.68458381 -0.01785194  0.62376484  6.50658904
##
## Residual standard error: 2.251601e-07
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.816128      1.836383      1.703670      1.676572
```

5.1.4 Replicate 4

```
rep_4.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.divers
rep_4.pi.50kb$avg <- apply(rep_4.pi.50kb[4:ncol(rep_4.pi.50kb)], 1, mean)
rep_4.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.rho.5
rep_4.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.the
rep_4.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.TMR
rep_4.tmrca.50kb$avg <- apply(rep_4.tmrca.50kb[4:ncol(rep_4.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep4 <- as.data.frame(cbind(rep_4.pi.50kb$avg, rep_4.theta.50kb$sample_mean, rep_4.tmrca
names(inf.lands.50kb.rep4) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep4$thetaC <- inf.lands.50kb.rep4$theta - mean(inf.lands.50kb.rep4$theta)
inf.lands.50kb.rep4$tmrcaC <- inf.lands.50kb.rep4$tmrca - mean(inf.lands.50kb.rep4$tmrca)
inf.lands.50kb.rep4$rhoC <- inf.lands.50kb.rep4$rho - mean(inf.lands.50kb.rep4$rho)

inf.lands.50kb.rep4$bin <- 1:nrow(inf.lands.50kb.rep4)

# for merging:
inf.lands.50kb.rep4 <- inf.lands.50kb.rep4
inf.lands.50kb.rep4$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep4)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep4)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7148.596
## m.div.50kb.2  7 -7314.670
## m.div.50kb.3  8 -7320.485
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 0.8328, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.30453, p-value < 2.2e-16
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -6.159e-04 -6.584e-05 -3.550e-06  6.891e-05  3.507e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.272e-03  6.258e-06  522.93  < 2e-16 ***
## thetaC         1.072e+00  9.427e-03  113.76  < 2e-16 ***
## rhoC          -2.087e-02  2.742e-03   -7.61 1.53e-13 ***
## tmrcaC         4.847e-03  6.387e-05   75.89  < 2e-16 ***
## thetaC:tmrcaC  1.500e+00  6.429e-02   23.33  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001216 on 466 degrees of freedom
## Multiple R-squared:  0.9875, Adjusted R-squared:  0.9874
## F-statistic:  9231 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```r
r2.bgs.50kb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 4] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep4, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep4, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep4, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep4, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7364.127
## g.div.50kb.2  8 -7380.679
## g.div.50kb.3  7 -7328.156
## g.div.50kb.4  7 -7185.307
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep4
##         AIC       BIC   logLik
##   -7328.156 -7299.072 3671.078
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.716365
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0032745 0.00000641 510.5089   0e+00
## thetaC         1.0530201 0.01402599  75.0764   0e+00
## rhoC          -0.0078260 0.00228445  -3.4258   7e-04
## tmrcaC         0.0047769 0.00004948  96.5484   0e+00
## thetaC:tmrcaC  1.5114666 0.07777455  19.4339   0e+00
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.717
## rhoC           0.208 -0.339
## tmrcaC         0.007 -0.209 -0.292
## thetaC:tmrcaC -0.468  0.403 -0.269 -0.042
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.09763551 -0.64794574 -0.02894227  0.67255210  3.68496533
##
## Residual standard error: 6.320069e-09
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.428139      1.368975      1.239488      1.223467
```

5.1.5 Replicate 5

```
rep_5.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.divers
rep_5.pi.50kb$avg <- apply(rep_5.pi.50kb[4:ncol(rep_5.pi.50kb)], 1, mean)
rep_5.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.rho.5
rep_5.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.the
rep_5.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.TMR
rep_5.tmrca.50kb$avg <- apply(rep_5.tmrca.50kb[4:ncol(rep_5.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep5 <- as.data.frame(cbind(rep_5.pi.50kb$avg, rep_5.theta.50kb$sample_mean, rep_5.tmrca
names(inf.lands.50kb.rep5) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep5$thetaC <- inf.lands.50kb.rep5$theta - mean(inf.lands.50kb.rep5$theta)
inf.lands.50kb.rep5$tmrcaC <- inf.lands.50kb.rep5$tmrca - mean(inf.lands.50kb.rep5$tmrca)
inf.lands.50kb.rep5$rhoC <- inf.lands.50kb.rep5$rho - mean(inf.lands.50kb.rep5$rho)

inf.lands.50kb.rep5$bin <- 1:nrow(inf.lands.50kb.rep5)

# for merging:
inf.lands.50kb.rep5 <- inf.lands.50kb.rep5
inf.lands.50kb.rep5$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep5)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep5)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7172.172
## m.div.50kb.2  7 -7341.719
## m.div.50kb.3  8 -7354.145
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 0.85321, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.2333, p-value < 2.2e-16
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep5)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.961e-04 -6.375e-05 -3.090e-06  6.777e-05  3.736e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.270e-03  5.753e-06 568.437  < 2e-16 ***
## thetaC         1.020e+00  7.027e-03 145.100  < 2e-16 ***
## rhoC          -1.412e-02  2.587e-03  -5.459  7.8e-08 ***
## tmrcaC         4.663e-03  6.224e-05  74.918  < 2e-16 ***
## thetaC:tmrcaC  1.404e+00  5.510e-02  25.477  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001186 on 466 degrees of freedom
## Multiple R-squared:  0.988,  Adjusted R-squared:  0.9879
## F-statistic:  9613 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```r
r2.bgs.50kb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 5] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep5, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep5, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep5, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep5, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7372.655
## g.div.50kb.2  8 -7396.802
## g.div.50kb.3  7 -7353.762
## g.div.50kb.4  7 -7191.430
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep5
##         AIC       BIC   logLik
##   -7353.762 -7324.678 3683.881
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.250143
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0032684 0.00000574 568.9219  0.0000
## thetaC         1.0163725 0.01103937  92.0680  0.0000
## rhoC          -0.0081960 0.00256257  -3.1984  0.0015
## tmrcaC         0.0046435 0.00004864  95.4763  0.0000
## thetaC:tmrcaC  1.5350345 0.06594486  23.2775  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.668
## rhoC           0.244 -0.335
## tmrcaC        -0.061 -0.092 -0.399
## thetaC:tmrcaC -0.382  0.327 -0.295  0.084
```

```
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -2.682153307 -0.700850203 -0.007002239  0.581689051  3.153007219
##
## Residual standard error: 8.22055e-08
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.290406      1.468923      1.276986      1.171610
```

5.1.6 Replicate 6

```
rep_6.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.divers
rep_6.pi.50kb$avg <- apply(rep_6.pi.50kb[4:ncol(rep_6.pi.50kb)], 1, mean)
rep_6.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.rho.5
rep_6.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.the
rep_6.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.TMR
rep_6.tmrca.50kb$avg <- apply(rep_6.tmrca.50kb[4:ncol(rep_6.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep6 <- as.data.frame(cbind(rep_6.pi.50kb$avg, rep_6.theta.50kb$sample_mean, rep_6.tmrca
names(inf.lands.50kb.rep6) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep6$thetaC <- inf.lands.50kb.rep6$theta - mean(inf.lands.50kb.rep6$theta)
inf.lands.50kb.rep6$tmrcaC <- inf.lands.50kb.rep6$tmrca - mean(inf.lands.50kb.rep6$tmrca)
inf.lands.50kb.rep6$rhoC <- inf.lands.50kb.rep6$rho - mean(inf.lands.50kb.rep6$rho)

inf.lands.50kb.rep6$bin <- 1:nrow(inf.lands.50kb.rep6)

# for merging:
inf.lands.50kb.rep6 <- inf.lands.50kb.rep6
inf.lands.50kb.rep6$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep6)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep6)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7335.144
## m.div.50kb.2  7 -7498.540
## m.div.50kb.3  8 -7514.166
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.3358, p-value = 9.341e-14
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.47386, p-value = 0.218
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.943e-04 -5.947e-05 -8.070e-06  6.832e-05  3.485e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.220e-03  5.176e-06 622.042   <2e-16 ***
## thetaC         9.534e-01  9.236e-03 103.226   <2e-16 ***
## rhoC          -5.393e-03  2.447e-03  -2.204    0.028 *
## tmrcaC         4.525e-03  3.735e-05 121.133   <2e-16 ***
## thetaC:tmrcaC  1.504e+00  4.103e-02  36.646   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.972e-05 on 466 degrees of freedom
## Multiple R-squared:  0.9921, Adjusted R-squared:  0.9921
## F-statistic: 1.468e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```r
r2.bgs.50kb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 6] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep6, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep6, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep6, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep6, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7401.361
## g.div.50kb.2  8 -7395.754
## g.div.50kb.3  7 -7343.374
## g.div.50kb.4  7 -7344.669
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep6
##         AIC      BIC   logLik
##   -7343.374 -7314.29 3678.687
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##      power
## -0.4875493
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0032193 0.00000530 607.6494  0.0000
## thetaC         0.9475585 0.01054102  89.8925  0.0000
## rhoC          -0.0045657 0.00232628  -1.9627  0.0503
## tmrcaC         0.0045371 0.00003659 123.9958  0.0000
## thetaC:tmrcaC  1.5262923 0.04601779  33.1674  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.324
## rhoC           0.154 -0.468
## tmrcaC        -0.233 -0.058 -0.392
## thetaC:tmrcaC -0.495  0.358 -0.295  0.481
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.08862261 -0.60657489 -0.08314079  0.71335416  3.70206913
##
## Residual standard error: 6.337495e-06
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.745874      1.684710      1.813824      1.645992
```

5.1.7 Replicate 7

```
rep_7.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.divers
rep_7.pi.50kb$avg <- apply(rep_7.pi.50kb[4:ncol(rep_7.pi.50kb)], 1, mean)
rep_7.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.rho.5
rep_7.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.the
rep_7.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.TMR
rep_7.tmrca.50kb$avg <- apply(rep_7.tmrca.50kb[4:ncol(rep_7.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep7 <- as.data.frame(cbind(rep_7.pi.50kb$avg, rep_7.theta.50kb$sample_mean, rep_7.tmrca
names(inf.lands.50kb.rep7) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep7$thetaC <- inf.lands.50kb.rep7$theta - mean(inf.lands.50kb.rep7$theta)
inf.lands.50kb.rep7$tmrcaC <- inf.lands.50kb.rep7$tmrca - mean(inf.lands.50kb.rep7$tmrca)
inf.lands.50kb.rep7$rhoC <- inf.lands.50kb.rep7$rho - mean(inf.lands.50kb.rep7$rho)

inf.lands.50kb.rep7$bin <- 1:nrow(inf.lands.50kb.rep7)

# for merging:
inf.lands.50kb.rep7 <- inf.lands.50kb.rep7
inf.lands.50kb.rep7$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep7)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep7)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7218.123
## m.div.50kb.2  7 -7414.252
## m.div.50kb.3  8 -7443.965
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.4443, p-value = 3.053e-10
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.50794, p-value = 0.565
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep7)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.662e-04 -8.213e-05 -4.060e-06  7.155e-05  3.156e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.226e-03  5.376e-06 600.011   <2e-16 ***
## thetaC        9.766e-01  8.060e-03 121.160   <2e-16 ***
## rhoC          1.410e-04  2.565e-03   0.055    0.956
## tmrcaC        4.100e-03  3.387e-05 121.046   <2e-16 ***
## thetaC:tmrcaC 1.142e+00  3.736e-02  30.582   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001129 on 466 degrees of freedom
## Multiple R-squared:  0.9901, Adjusted R-squared:   0.99
## F-statistic: 1.168e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```r
r2.bgs.50kb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 7] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep7, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep7, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep7, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep7, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7272.529
## g.div.50kb.2  8 -7269.935
## g.div.50kb.3  7 -7216.131
## g.div.50kb.4  7 -7223.209
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep7
##         AIC       BIC   logLik
##   -7216.131 -7187.047 3615.065
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##      power
## 0.01202936
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0032257 0.00000537 600.2221  0.0000
## thetaC        0.9766507 0.00800060 122.0722  0.0000
## rhoC          0.0000939 0.00255835   0.0367  0.9707
## tmrcaC        0.0040999 0.00003387 121.0465  0.0000
## thetaC:tmrcaC 1.1430656 0.03716320  30.7580  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.073
## rhoC           0.059 -0.347
## tmrcaC        -0.073 -0.001 -0.471
## thetaC:tmrcaC -0.250  0.312 -0.236  0.293
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.27722127 -0.73114058 -0.03610908  0.63566694  2.80728591
##
## Residual standard error: 0.0001202057
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.306438      1.520048      1.447676      1.224616
```

5.1.8 Replicate 8

```
rep_8.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.divers
rep_8.pi.50kb$avg <- apply(rep_8.pi.50kb[4:ncol(rep_8.pi.50kb)], 1, mean)
rep_8.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.rho.5
rep_8.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.the
rep_8.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.TMR
rep_8.tmrca.50kb$avg <- apply(rep_8.tmrca.50kb[4:ncol(rep_8.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep8 <- as.data.frame(cbind(rep_8.pi.50kb$avg, rep_8.theta.50kb$sample_mean, rep_8.tmrca
names(inf.lands.50kb.rep8) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep8$thetaC <- inf.lands.50kb.rep8$theta - mean(inf.lands.50kb.rep8$theta)
inf.lands.50kb.rep8$tmrcaC <- inf.lands.50kb.rep8$tmrca - mean(inf.lands.50kb.rep8$tmrca)
inf.lands.50kb.rep8$rhoC <- inf.lands.50kb.rep8$rho - mean(inf.lands.50kb.rep8$rho)

inf.lands.50kb.rep8$bin <- 1:nrow(inf.lands.50kb.rep8)

# for merging:
inf.lands.50kb.rep8 <- inf.lands.50kb.rep8
inf.lands.50kb.rep8$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep8)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep8)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7306.572
## m.div.50kb.2  7 -7395.464
## m.div.50kb.3  8 -7417.026
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.1894, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.43038, p-value = 0.015
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep8)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.769e-04 -6.997e-05  7.200e-07  6.476e-05  4.430e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.222e-03  5.257e-06 612.845   <2e-16 ***
## thetaC         1.024e+00  7.003e-03 146.273   <2e-16 ***
## rhoC          -2.032e-02  2.326e-03  -8.735   <2e-16 ***
## tmrcaC         4.540e-03  4.741e-05  95.750   <2e-16 ***
## thetaC:tmrcaC  1.539e+00  5.006e-02  30.742   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001028 on 466 degrees of freedom
## Multiple R-squared:  0.9913, Adjusted R-squared:  0.9912
## F-statistic: 1.329e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```r
r2.bgs.50kb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 8] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep8, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep8, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep8, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep8, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7416.341
## g.div.50kb.2  8 -7387.986
## g.div.50kb.3  7 -7327.716
## g.div.50kb.4  7 -7355.330
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep8
##         AIC       BIC   logLik
##   -7327.716 -7298.632 3670.858
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -0.734904
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0032208 0.00000559 575.7028       0
## thetaC         1.0152924 0.01030530  98.5214       0
## rhoC          -0.0180657 0.00260902  -6.9243       0
## tmrcaC         0.0045590 0.00004496 101.4012       0
## thetaC:tmrcaC  1.5987953 0.05970086  26.7801       0
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.465
## rhoC           0.232 -0.300
## tmrcaC        -0.208 -0.109 -0.454
## thetaC:tmrcaC -0.500  0.342 -0.328  0.416
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.80727802 -0.69707408 -0.02648353  0.58385802  5.31468972
##
## Residual standard error: 1.572529e-06
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.424254      1.490470      1.682699      1.481909
```

5.1.9 Replicate 9

```
rep_9.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.divers
rep_9.pi.50kb$avg <- apply(rep_9.pi.50kb[4:ncol(rep_9.pi.50kb)], 1, mean)
rep_9.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.rho.5
rep_9.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.the
rep_9.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.TMR
rep_9.tmrca.50kb$avg <- apply(rep_9.tmrca.50kb[4:ncol(rep_9.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep9 <- as.data.frame(cbind(rep_9.pi.50kb$avg, rep_9.theta.50kb$sample_mean, rep_9.tmrca
names(inf.lands.50kb.rep9) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep9$thetaC <- inf.lands.50kb.rep9$theta - mean(inf.lands.50kb.rep9$theta)
inf.lands.50kb.rep9$tmrcaC <- inf.lands.50kb.rep9$tmrca - mean(inf.lands.50kb.rep9$tmrca)
inf.lands.50kb.rep9$rhoC <- inf.lands.50kb.rep9$rho - mean(inf.lands.50kb.rep9$rho)

inf.lands.50kb.rep9$bin <- 1:nrow(inf.lands.50kb.rep9)

# for merging:
inf.lands.50kb.rep9 <- inf.lands.50kb.rep9
inf.lands.50kb.rep9$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep9)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep9)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7369.759
## m.div.50kb.2  7 -7424.115
## m.div.50kb.3  8 -7428.899
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.2528, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.46539, p-value = 0.142
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep9)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.264e-04 -6.316e-05 -1.027e-05  6.801e-05  4.899e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.248e-03  5.227e-06 621.439   <2e-16 ***
## thetaC         1.007e+00  1.036e-02  97.243   <2e-16 ***
## rhoC          -1.100e-02  3.402e-03  -3.235   0.0013 **
## tmrcaC         4.379e-03  4.958e-05  88.332   <2e-16 ***
## thetaC:tmrcaC  1.555e+00  5.186e-02  29.995   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.612e-05 on 466 degrees of freedom
## Multiple R-squared:  0.9917, Adjusted R-squared:  0.9917
## F-statistic: 1.396e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```
r2.bgs.50kb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 9] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep9, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep9, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep9, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep9, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7451.009
## g.div.50kb.2  8 -7439.797
## g.div.50kb.3  7 -7376.047
## g.div.50kb.4  7 -7383.267
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep9
##         AIC       BIC   logLik
##   -7376.047 -7346.963 3695.024
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##      power
## -0.4270749
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0032474 0.00000528 614.7741  0.0000
## thetaC         1.0062408 0.01106024  90.9782  0.0000
## rhoC          -0.0102097 0.00323678  -3.1543  0.0017
## tmrcaC         0.0043865 0.00004812  91.1520  0.0000
## thetaC:tmrcaC  1.5757705 0.05403968  29.1595  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.243
## rhoC           0.170 -0.643
## tmrcaC        -0.350  0.002 -0.396
## thetaC:tmrcaC -0.544  0.245 -0.301  0.648
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.2113836 -0.6525291 -0.1192687  0.7100956  5.1287732
##
## Residual standard error: 8.453983e-06
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.305163      2.465079      2.509466      2.033860
```

5.1.10 Replicate 10

```
rep_10.pi.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5.di
rep_10.pi.50kb$avg <- apply(rep_10.pi.50kb[4:ncol(rep_10.pi.50kb)], 1, mean)
rep_10.rho.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5.r
rep_10.theta.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5
rep_10.tmrca.50kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5
rep_10.tmrca.50kb$avg <- apply(rep_10.tmrca.50kb[4:ncol(rep_10.tmrca.50kb)], 1, mean)

inf.lands.50kb.rep10 <- as.data.frame(cbind(rep_10.pi.50kb$avg, rep_10.theta.50kb$sample_mean, rep_10.t
names(inf.lands.50kb.rep10) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.50kb.rep10$thetaC <- inf.lands.50kb.rep10$theta - mean(inf.lands.50kb.rep10$theta)
inf.lands.50kb.rep10$tmrcaC <- inf.lands.50kb.rep10$tmrca - mean(inf.lands.50kb.rep10$tmrca)
inf.lands.50kb.rep10$rhoC <- inf.lands.50kb.rep10$rho - mean(inf.lands.50kb.rep10$rho)

inf.lands.50kb.rep10$bin <- 1:nrow(inf.lands.50kb.rep10)

# for merging:
inf.lands.50kb.rep10 <- inf.lands.50kb.rep10
inf.lands.50kb.rep10$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep10)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep10)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -7460.959
## m.div.50kb.2  7 -7475.549
## m.div.50kb.3  8 -7483.607
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.5178, p-value = 3.811e-08
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.53736, p-value = 0.884
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep10)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.478e-04 -6.053e-05 -2.924e-06  5.485e-05  2.698e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.198e-03  5.109e-06 625.860   <2e-16 ***
## thetaC         9.866e-01  8.927e-03 110.518   <2e-16 ***
## rhoC          -7.779e-03  2.350e-03  -3.311    0.001 **
## tmrcaC         4.495e-03  4.372e-05 102.823   <2e-16 ***
## thetaC:tmrcaC  1.665e+00  4.678e-02  35.582   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.725e-05 on 466 degrees of freedom
## Multiple R-squared:  0.9935, Adjusted R-squared:  0.9934
## F-statistic: 1.774e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.bgs.50kb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 10] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep10, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep10, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep10, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep10, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -7484.624
## g.div.50kb.2  8 -7485.078
## g.div.50kb.3  7 -7459.832
## g.div.50kb.4  7 -7459.998
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep10
##         AIC       BIC   logLik
##   -7459.832 -7430.748 3736.916
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##      power
## -0.1592648
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)    0.0031967 0.00000514 621.9328  0.0000
## thetaC         0.9849415 0.00928349 106.0960  0.0000
## rhoC          -0.0075689 0.00236213  -3.2043  0.0014
## tmrcaC         0.0045075 0.00004366 103.2465  0.0000
## thetaC:tmrcaC  1.6785292 0.04770442  35.1860  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.269
## rhoC           0.302 -0.536
## tmrcaC        -0.434 -0.104 -0.319
## thetaC:tmrcaC -0.623  0.370 -0.480  0.698
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.91439880 -0.70455045 -0.02126716  0.64062857  2.97141938
##
## Residual standard error: 3.517815e-05
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.216705      1.751910      3.009599      3.171116
```

5.2 200 kb scale

```
r2.bgs.200kb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.bgs.200kb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.bgs.200kb) <- reps

sim.theta.200kb <- read.table("dm_bgs_sims/MutationMap_200kb.csv", header = T)
sim.rho.200kb <- read.table("dm_bgs_sims/RecombinationMap_200kb.csv", header = T)
```

5.2.1 Replicate 1

```
rep_1.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.diver
rep_1.pi.200kb$avg <- apply(rep_1.pi.200kb[4:ncol(rep_1.pi.200kb)], 1, mean)
rep_1.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.rho
rep_1.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.th
rep_1.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.TM
rep_1.tmrca.200kb$avg <- apply(rep_1.tmrca.200kb[4:ncol(rep_1.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep1 <- as.data.frame(cbind(rep_1.pi.200kb$avg, rep_1.theta.200kb$sample_mean, rep_1.tmr
names(inf.lands.200kb.rep1) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep1$thetaC <- inf.lands.200kb.rep1$theta - mean(inf.lands.200kb.rep1$theta)
inf.lands.200kb.rep1$tmrcaC <- inf.lands.200kb.rep1$tmrca - mean(inf.lands.200kb.rep1$tmrca)
inf.lands.200kb.rep1$rhoC <- inf.lands.200kb.rep1$rho - mean(inf.lands.200kb.rep1$rho)

inf.lands.200kb.rep1$bin <- 1:nrow(inf.lands.200kb.rep1)

# for merging:
inf.lands.200kb.rep1 <- inf.lands.200kb.rep1
inf.lands.200kb.rep1$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep1)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep1)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1914.977
## m.div.200kb.2  7 -1944.689
```

```
## m.div.200kb.3  8 -1949.080
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.2584, p-value = 1.066e-05
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.44852, p-value = 0.213
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep1)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.283e-04 -4.365e-05 -6.902e-06  4.766e-05  1.538e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.281e-03  7.378e-06 444.748   <2e-16 ***
## thetaC         9.974e-01  9.968e-03 100.055   <2e-16 ***
## rhoC          -4.104e-03  3.993e-03  -1.028    0.306
## tmrcaC         4.264e-03  9.788e-05  43.569   <2e-16 ***
## thetaC:tmrcaC  1.296e+00  9.192e-02  14.098   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.032e-05 on 113 degrees of freedom
## Multiple R-squared:  0.9953, Adjusted R-squared:  0.9951
## F-statistic:  5984 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.bgs.200kb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 1] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep1, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep1, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep1, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep1, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1929.159
## g.div.200kb.2  8 -1933.494
## g.div.200kb.3  7 -1926.365
## g.div.200kb.4  7 -1917.258
```

```
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep1
##         AIC       BIC   logLik
##   -1926.365 -1906.971 970.1827
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.080861
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0032885 0.00000834 394.2498  0.0000
## thetaC        0.9708036 0.01721523  56.3921  0.0000
## rhoC          0.0033041 0.00402510   0.8209  0.4134
## tmrcaC        0.0042215 0.00008753  48.2274  0.0000
## thetaC:tmrcaC 1.2036805 0.11732841  10.2591  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.639
## rhoC           0.349 -0.467
## tmrcaC        -0.180 -0.069 -0.485
## thetaC:tmrcaC -0.584  0.488 -0.484  0.339
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.99907520 -0.66401443 -0.03059178  0.50174113  2.54737837
##
## Residual standard error: 1.462783e-07
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.831093      1.973108      1.688173      1.625769
```

5.2.2 Replicate 2

```
rep_2.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.dive
rep_2.pi.200kb$avg <- apply(rep_2.pi.200kb[4:ncol(rep_2.pi.200kb)], 1, mean)
rep_2.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.rho
rep_2.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.t
rep_2.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.T
rep_2.tmrca.200kb$avg <- apply(rep_2.tmrca.200kb[4:ncol(rep_2.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep2 <- as.data.frame(cbind(rep_2.pi.200kb$avg, rep_2.theta.200kb$sample_mean, rep_2.tm
names(inf.lands.200kb.rep2) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep2$thetaC <- inf.lands.200kb.rep2$theta - mean(inf.lands.200kb.rep2$theta)
inf.lands.200kb.rep2$tmrcaC <- inf.lands.200kb.rep2$tmrca - mean(inf.lands.200kb.rep2$tmrca)
inf.lands.200kb.rep2$rhoC <- inf.lands.200kb.rep2$rho - mean(inf.lands.200kb.rep2$rho)

inf.lands.200kb.rep2$bin <- 1:nrow(inf.lands.200kb.rep2)

# for merging:
inf.lands.200kb.rep2 <- inf.lands.200kb.rep2
inf.lands.200kb.rep2$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep2)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep2)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1887.453
## m.div.200kb.2  7 -1919.176
## m.div.200kb.3  8 -1920.353
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.063, p-value = 5.447e-08
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.49635, p-value = 0.472
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.533e-04 -3.652e-05 -6.176e-06  4.179e-05  1.788e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.262e-03  7.579e-06 430.439   <2e-16 ***
## thetaC        9.986e-01  8.965e-03 111.396   <2e-16 ***
## rhoC          8.743e-04  4.151e-03   0.211    0.834
## tmrcaC        4.301e-03  1.082e-04  39.768   <2e-16 ***
## thetaC:tmrcaC 1.155e+00  8.152e-02  14.172   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.902e-05 on 113 degrees of freedom
## Multiple R-squared:  0.9946, Adjusted R-squared:  0.9944
## F-statistic:  5214 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 2] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep2, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep2, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep2, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep2, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1917.295
## g.div.200kb.2  8 -1916.677
## g.div.200kb.3  7 -1903.456
## g.div.200kb.4  7 -1896.019
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep2
##         AIC       BIC   logLik
##   -1903.456 -1884.061 958.7279
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.085434
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0032643 0.00000853 382.8006  0.0000
## thetaC        0.9899751 0.01622286  61.0235  0.0000
## rhoC          0.0069842 0.00405325   1.7231  0.0876
## tmrcaC        0.0042272 0.00009516  44.4238  0.0000
## thetaC:tmrcaC 1.1253916 0.11680302   9.6350  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.654
## rhoC           0.247 -0.385
## tmrcaC        -0.083 -0.021 -0.561
## thetaC:tmrcaC -0.395  0.329 -0.409  0.237
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.51891730 -0.57824590 -0.03723372  0.70670288  2.40975217
##
## Residual standard error: 1.488727e-07
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.363138      1.985520      1.622713      1.260591
```

5.2.3 Replicate 3

```
rep_3.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.diver
rep_3.pi.200kb$avg <- apply(rep_3.pi.200kb[4:ncol(rep_3.pi.200kb)], 1, mean)
rep_3.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.rho
rep_3.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.th
rep_3.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.TM
rep_3.tmrca.200kb$avg <- apply(rep_3.tmrca.200kb[4:ncol(rep_3.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep3 <- as.data.frame(cbind(rep_3.pi.200kb$avg, rep_3.theta.200kb$sample_mean, rep_3.tm
names(inf.lands.200kb.rep3) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep3$thetaC <- inf.lands.200kb.rep3$theta - mean(inf.lands.200kb.rep3$theta)
inf.lands.200kb.rep3$tmrcaC <- inf.lands.200kb.rep3$tmrca - mean(inf.lands.200kb.rep3$tmrca)
inf.lands.200kb.rep3$rhoC <- inf.lands.200kb.rep3$rho - mean(inf.lands.200kb.rep3$rho)

inf.lands.200kb.rep3$bin <- 1:nrow(inf.lands.200kb.rep3)

# for merging:
inf.lands.200kb.rep3 <- inf.lands.200kb.rep3
inf.lands.200kb.rep3$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep3)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep3)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1876.293
## m.div.200kb.2  7 -1898.010
## m.div.200kb.3  8 -1902.143
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.1111, p-value = 1.783e-07
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.46233, p-value = 0.281
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.252e-04 -4.059e-05 -6.880e-06  2.642e-05  2.636e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.242e-03  9.601e-06 337.658  < 2e-16 ***
## thetaC         9.685e-01  1.825e-02  53.057  < 2e-16 ***
## rhoC          -2.030e-02  5.679e-03  -3.575 0.000517 ***
## tmrcaC         4.991e-03  1.229e-04  40.605  < 2e-16 ***
## thetaC:tmrcaC  1.520e+00  9.968e-02  15.248  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.285e-05 on 113 degrees of freedom
## Multiple R-squared:  0.9935, Adjusted R-squared:  0.9932
## F-statistic:  4302 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 3] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep3, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep3, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep3, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep3, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1922.374
## g.div.200kb.2  8 -1930.481
## g.div.200kb.3  7 -1924.104
## g.div.200kb.4  7 -1894.752
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep3
##         AIC       BIC   logLik
##   -1924.104 -1904.709 969.0518
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -2.126211
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0032622 0.00001028 317.2138  0.0000
## thetaC        0.9224921 0.02375101  38.8401  0.0000
## rhoC          0.0027288 0.00490482   0.5564  0.5791
## tmrcaC        0.0046998 0.00010351  45.4062  0.0000
## thetaC:tmrcaC 1.1347165 0.14316037   7.9262  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.743
## rhoC           0.398 -0.416
## tmrcaC        -0.206 -0.122 -0.565
## thetaC:tmrcaC -0.646  0.503 -0.522  0.260
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.50460278 -0.67087750 -0.06109084  0.52042110  4.19078904
##
## Residual standard error: 4.250353e-10
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.848355      2.367957      1.945836      1.628937
```

5.2.4 Replicate 4

```
rep_4.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.diver
rep_4.pi.200kb$avg <- apply(rep_4.pi.200kb[4:ncol(rep_4.pi.200kb)], 1, mean)
rep_4.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.rho
rep_4.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.th
rep_4.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.TM
rep_4.tmrca.200kb$avg <- apply(rep_4.tmrca.200kb[4:ncol(rep_4.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep4 <- as.data.frame(cbind(rep_4.pi.200kb$avg, rep_4.theta.200kb$sample_mean, rep_4.tmr
names(inf.lands.200kb.rep4) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep4$thetaC <- inf.lands.200kb.rep4$theta - mean(inf.lands.200kb.rep4$theta)
inf.lands.200kb.rep4$tmrcaC <- inf.lands.200kb.rep4$tmrca - mean(inf.lands.200kb.rep4$tmrca)
inf.lands.200kb.rep4$rhoC <- inf.lands.200kb.rep4$rho - mean(inf.lands.200kb.rep4$rho)

inf.lands.200kb.rep4$bin <- 1:nrow(inf.lands.200kb.rep4)

# for merging:
inf.lands.200kb.rep4 <- inf.lands.200kb.rep4
inf.lands.200kb.rep4$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep4)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep4)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1827.605
## m.div.200kb.2  7 -1873.540
## m.div.200kb.3  8 -1871.569
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.3604, p-value = 0.0001091
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.2301, p-value < 2.2e-16
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -5.342e-04 -4.098e-05 -7.900e-06  4.752e-05  3.074e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.276e-03  1.143e-05 286.660  < 2e-16 ***
## thetaC         1.045e+00  1.761e-02  59.334  < 2e-16 ***
## rhoC          -2.233e-02  5.132e-03  -4.350 2.99e-05 ***
## tmrcaC         5.026e-03  1.709e-04  29.414  < 2e-16 ***
## thetaC:tmrcaC  1.339e+00  1.551e-01   8.634 4.35e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001018 on 113 degrees of freedom
## Multiple R-squared:  0.9904, Adjusted R-squared:  0.9901
## F-statistic:  2915 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 4] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep4, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep4, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep4, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep4, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1855.917
## g.div.200kb.2  8 -1925.934
## g.div.200kb.3  7 -1927.927
## g.div.200kb.4  7 -1847.741
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep4
##         AIC       BIC   logLik
##   -1927.927 -1908.532 970.9635
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -2.669007
##
## Coefficients:
##                     Value  Std.Error   t-value p-value
## (Intercept)    0.0032818 0.00001095 299.72122  0.0000
## thetaC         1.0207123 0.02542774  40.14169  0.0000
## rhoC          -0.0009231 0.00320329  -0.28818  0.7737
## tmrcaC         0.0047770 0.00010882  43.89864  0.0000
## thetaC:tmrcaC  1.2979389 0.18863144   6.88082  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.848
## rhoC           0.274 -0.302
## tmrcaC         0.155 -0.357 -0.311
## thetaC:tmrcaC -0.544  0.455 -0.271 -0.311
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.02838223 -0.65770633 -0.02151832  0.58911537  2.66454244
##
## Residual standard error: 1.912967e-11
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.554579      1.528687      1.603813      1.410626
```

5.2.5 Replicate 5

```
rep_5.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.diver
rep_5.pi.200kb$avg <- apply(rep_5.pi.200kb[4:ncol(rep_5.pi.200kb)], 1, mean)
rep_5.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.rho
rep_5.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.th
rep_5.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.TM
rep_5.tmrca.200kb$avg <- apply(rep_5.tmrca.200kb[4:ncol(rep_5.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep5 <- as.data.frame(cbind(rep_5.pi.200kb$avg, rep_5.theta.200kb$sample_mean, rep_5.tmr
names(inf.lands.200kb.rep5) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep5$thetaC <- inf.lands.200kb.rep5$theta - mean(inf.lands.200kb.rep5$theta)
inf.lands.200kb.rep5$tmrcaC <- inf.lands.200kb.rep5$tmrca - mean(inf.lands.200kb.rep5$tmrca)
inf.lands.200kb.rep5$rhoC <- inf.lands.200kb.rep5$rho - mean(inf.lands.200kb.rep5$rho)

inf.lands.200kb.rep5$bin <- 1:nrow(inf.lands.200kb.rep5)

# for merging:
inf.lands.200kb.rep5 <- inf.lands.200kb.rep5
inf.lands.200kb.rep5$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep5)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep5)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1832.288
## m.div.200kb.2  7 -1908.006
## m.div.200kb.3  8 -1906.692
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 0.92809, p-value = 6.318e-10
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.11675, p-value < 2.2e-16
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep5)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.791e-04 -3.534e-05 -6.800e-06  5.343e-05  3.238e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.283e-03  1.059e-05 310.058  < 2e-16 ***
## thetaC         9.887e-01  1.432e-02  69.023  < 2e-16 ***
## rhoC          -1.473e-02  5.029e-03  -2.928  0.00413 **
## tmrcaC         4.787e-03  1.789e-04  26.750  < 2e-16 ***
## thetaC:tmrcaC  8.767e-01  1.750e-01   5.009 2.04e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.983e-05 on 113 degrees of freedom
## Multiple R-squared:  0.9905, Adjusted R-squared:  0.9902
## F-statistic:  2954 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 5] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep5, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep5, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep5, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep5, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1904.235
## g.div.200kb.2  8 -1966.271
## g.div.200kb.3  7 -1957.545
## g.div.200kb.4  7 -1877.048
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep5
##         AIC      BIC   logLik
##   -1957.545 -1938.15 985.7726
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -2.391887
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0032774 0.00001096 299.03645  0.0000
## thetaC        0.9895543 0.02287693  43.25556  0.0000
## rhoC          0.0066653 0.00365417   1.82403  0.0708
## tmrcaC        0.0042963 0.00012657  33.94444  0.0000
## thetaC:tmrcaC 1.5492412 0.25844201   5.99454  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.904
## rhoC           0.342 -0.341
## tmrcaC         0.348 -0.416 -0.313
## thetaC:tmrcaC -0.627  0.574 -0.266 -0.618
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.51067986 -0.72940766 -0.08537948  0.66721967  2.52110876
##
## Residual standard error: 7.541972e-11
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.770022      2.101209      3.007245      2.679306
```

5.2.6 Replicate 6

```
rep_6.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.diver
rep_6.pi.200kb$avg <- apply(rep_6.pi.200kb[4:ncol(rep_6.pi.200kb)], 1, mean)
rep_6.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.rho
rep_6.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.th
rep_6.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.TM
rep_6.tmrca.200kb$avg <- apply(rep_6.tmrca.200kb[4:ncol(rep_6.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep6 <- as.data.frame(cbind(rep_6.pi.200kb$avg, rep_6.theta.200kb$sample_mean, rep_6.tm
names(inf.lands.200kb.rep6) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep6$thetaC <- inf.lands.200kb.rep6$theta - mean(inf.lands.200kb.rep6$theta)
inf.lands.200kb.rep6$tmrcaC <- inf.lands.200kb.rep6$tmrca - mean(inf.lands.200kb.rep6$tmrca)
inf.lands.200kb.rep6$rhoC <- inf.lands.200kb.rep6$rho - mean(inf.lands.200kb.rep6$rho)

inf.lands.200kb.rep6$bin <- 1:nrow(inf.lands.200kb.rep6)

# for merging:
inf.lands.200kb.rep6 <- inf.lands.200kb.rep6
inf.lands.200kb.rep6$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep6)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep6)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1911.034
## m.div.200kb.2  7 -1936.062
## m.div.200kb.3  8 -1935.098
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.5065, p-value = 0.001877
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.3467, p-value = 0.006
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.776e-04 -3.617e-05 -5.260e-06  3.802e-05  1.743e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.232e-03  8.240e-06 392.222   <2e-16 ***
## thetaC        9.288e-01  1.516e-02  61.272   <2e-16 ***
## rhoC          4.115e-04  3.911e-03   0.105    0.916
## tmrcaC        4.368e-03  6.441e-05  67.810   <2e-16 ***
## thetaC:tmrcaC 1.264e+00  8.898e-02  14.208   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.151e-05 on 113 degrees of freedom
## Multiple R-squared:  0.9955, Adjusted R-squared:  0.9953
## F-statistic:  6262 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 6] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep6, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep6, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep6, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep6, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1920.492
## g.div.200kb.2  8 -1940.555
## g.div.200kb.3  7 -1936.838
## g.div.200kb.4  7 -1914.435
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep6
##         AIC       BIC   logLik
##   -1936.838 -1917.444 975.4192
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.630375
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0032306 0.00000894 361.3092  0.0000
## thetaC        0.9172863 0.02050378  44.7374  0.0000
## rhoC          0.0022991 0.00298926   0.7691  0.4434
## tmrcaC        0.0043875 0.00005702  76.9426  0.0000
## thetaC:tmrcaC 1.3282349 0.11467784  11.5823  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.665
## rhoC           0.227 -0.361
## tmrcaC        -0.112 -0.204 -0.404
## thetaC:tmrcaC -0.650  0.440 -0.318  0.249
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -3.6868115 -0.5651622 -0.1157522  0.6510378  3.6401674
##
## Residual standard error: 6.533835e-09
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.776494      1.583826      1.627773      1.457581
```

5.2.7 Replicate 7

```
rep_7.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.diver
rep_7.pi.200kb$avg <- apply(rep_7.pi.200kb[4:ncol(rep_7.pi.200kb)], 1, mean)
rep_7.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.rho
rep_7.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.th
rep_7.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.TM
rep_7.tmrca.200kb$avg <- apply(rep_7.tmrca.200kb[4:ncol(rep_7.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep7 <- as.data.frame(cbind(rep_7.pi.200kb$avg, rep_7.theta.200kb$sample_mean, rep_7.tmr
names(inf.lands.200kb.rep7) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep7$thetaC <- inf.lands.200kb.rep7$theta - mean(inf.lands.200kb.rep7$theta)
inf.lands.200kb.rep7$tmrcaC <- inf.lands.200kb.rep7$tmrca - mean(inf.lands.200kb.rep7$tmrca)
inf.lands.200kb.rep7$rhoC <- inf.lands.200kb.rep7$rho - mean(inf.lands.200kb.rep7$rho)

inf.lands.200kb.rep7$bin <- 1:nrow(inf.lands.200kb.rep7)

# for merging:
inf.lands.200kb.rep7 <- inf.lands.200kb.rep7
inf.lands.200kb.rep7$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep7)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep7)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1911.034
## m.div.200kb.2  7 -1936.062
## m.div.200kb.3  8 -1935.098
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.5065, p-value = 0.001877
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.3467, p-value = 0.009
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep7)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.776e-04 -3.617e-05 -5.260e-06  3.802e-05  1.743e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.232e-03  8.240e-06 392.222   <2e-16 ***
## thetaC        9.288e-01  1.516e-02  61.272   <2e-16 ***
## rhoC          4.115e-04  3.911e-03   0.105    0.916
## tmrcaC        4.368e-03  6.441e-05  67.810   <2e-16 ***
## thetaC:tmrcaC 1.264e+00  8.898e-02  14.208   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.151e-05 on 113 degrees of freedom
## Multiple R-squared:  0.9955, Adjusted R-squared:  0.9953
## F-statistic:  6262 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 7] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep7, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep7, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep7, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep7, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1920.492
## g.div.200kb.2  8 -1940.555
## g.div.200kb.3  7 -1936.838
## g.div.200kb.4  7 -1914.435
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep7
##         AIC       BIC   logLik
##   -1936.838 -1917.444 975.4192
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.630375
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0032306 0.00000894 361.3092  0.0000
## thetaC        0.9172863 0.02050378  44.7374  0.0000
## rhoC          0.0022991 0.00298926   0.7691  0.4434
## tmrcaC        0.0043875 0.00005702  76.9426  0.0000
## thetaC:tmrcaC 1.3282349 0.11467784  11.5823  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.665
## rhoC           0.227 -0.361
## tmrcaC        -0.112 -0.204 -0.404
## thetaC:tmrcaC -0.650  0.440 -0.318  0.249
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -3.6868115 -0.5651622 -0.1157522  0.6510378  3.6401674
##
## Residual standard error: 6.533835e-09
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.776494      1.583826      1.627773      1.457581
```

5.2.8 Replicate 8

```
rep_8.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.diver
rep_8.pi.200kb$avg <- apply(rep_8.pi.200kb[4:ncol(rep_8.pi.200kb)], 1, mean)
rep_8.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.rho
rep_8.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.th
rep_8.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.TM
rep_8.tmrca.200kb$avg <- apply(rep_8.tmrca.200kb[4:ncol(rep_8.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep8 <- as.data.frame(cbind(rep_8.pi.200kb$avg, rep_8.theta.200kb$sample_mean, rep_8.tmr
names(inf.lands.200kb.rep8) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep8$thetaC <- inf.lands.200kb.rep8$theta - mean(inf.lands.200kb.rep8$theta)
inf.lands.200kb.rep8$tmrcaC <- inf.lands.200kb.rep8$tmrca - mean(inf.lands.200kb.rep8$tmrca)
inf.lands.200kb.rep8$rhoC <- inf.lands.200kb.rep8$rho - mean(inf.lands.200kb.rep8$rho)

inf.lands.200kb.rep8$bin <- 1:nrow(inf.lands.200kb.rep8)

# for merging:
inf.lands.200kb.rep8 <- inf.lands.200kb.rep8
inf.lands.200kb.rep8$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep8)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep8)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df      AIC
## m.div.200kb    6 -1905.50
## m.div.200kb.2  7 -1923.90
## m.div.200kb.3  8 -1925.21
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.2303, p-value = 5.166e-06
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.37056, p-value = 0.027
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep8)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.390e-04 -3.825e-05 -5.591e-06  4.443e-05  1.988e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.221e-03  8.322e-06 387.016  < 2e-16 ***
## thetaC         1.017e+00  1.094e-02  92.962  < 2e-16 ***
## rhoC          -2.092e-02  3.786e-03  -5.526 2.13e-07 ***
## tmrcaC         4.560e-03  1.005e-04  45.393  < 2e-16 ***
## thetaC:tmrcaC  1.522e+00  1.131e-01  13.456  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.32e-05 on 113 degrees of freedom
## Multiple R-squared:  0.9951, Adjusted R-squared:  0.9949
## F-statistic:  5732 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 8] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep8, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep8, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep8, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep8, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1932.600
## g.div.200kb.2  8 -1937.089
## g.div.200kb.3  7 -1937.377
## g.div.200kb.4  7 -1918.893
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep8
##         AIC       BIC   logLik
##   -1937.377 -1917.982 975.6885
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.878383
##
## Coefficients:
##                     Value  Std.Error  t-value p-value
## (Intercept)    0.0032344 0.00000985 328.4171  0.0000
## thetaC         0.9627689 0.02185825  44.0460  0.0000
## rhoC          -0.0082138 0.00386613  -2.1245  0.0358
## tmrcaC         0.0045252 0.00008179  55.3261  0.0000
## thetaC:tmrcaC  1.5036433 0.13840840  10.8638  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.784
## rhoC           0.358 -0.351
## tmrcaC        -0.030 -0.234 -0.483
## thetaC:tmrcaC -0.638  0.482 -0.338  0.067
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.9946307 -0.6775280 -0.1141325  0.6237563  3.1278548
##
## Residual standard error: 1.530777e-09
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.811248      1.875762      1.743492      1.375969
```

5.2.9 Replicate 9

```
rep_9.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.diver
rep_9.pi.200kb$avg <- apply(rep_9.pi.200kb[4:ncol(rep_9.pi.200kb)], 1, mean)
rep_9.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.rho
rep_9.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.th
rep_9.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.TM
rep_9.tmrca.200kb$avg <- apply(rep_9.tmrca.200kb[4:ncol(rep_9.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep9 <- as.data.frame(cbind(rep_9.pi.200kb$avg, rep_9.theta.200kb$sample_mean, rep_9.tm
names(inf.lands.200kb.rep9) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep9$thetaC <- inf.lands.200kb.rep9$theta - mean(inf.lands.200kb.rep9$theta)
inf.lands.200kb.rep9$tmrcaC <- inf.lands.200kb.rep9$tmrca - mean(inf.lands.200kb.rep9$tmrca)
inf.lands.200kb.rep9$rhoC <- inf.lands.200kb.rep9$rho - mean(inf.lands.200kb.rep9$rho)

inf.lands.200kb.rep9$bin <- 1:nrow(inf.lands.200kb.rep9)

# for merging:
inf.lands.200kb.rep9 <- inf.lands.200kb.rep9
inf.lands.200kb.rep9$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep9)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep9)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1906.537
## m.div.200kb.2  7 -1906.701
## m.div.200kb.3  8 -1910.859
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.2787, p-value = 2.2e-05
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.35665, p-value = 0.017
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep9)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.076e-04 -4.682e-05 -6.010e-06  3.996e-05  3.232e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.266e-03  8.920e-06 366.151   <2e-16 ***
## thetaC        9.871e-01  1.741e-02  56.693   <2e-16 ***
## rhoC          9.046e-04  5.911e-03   0.153    0.879
## tmrcaC        4.050e-03  1.058e-04  38.264   <2e-16 ***
## thetaC:tmrcaC 1.218e+00  1.156e-01  10.530   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.288e-05 on 113 degrees of freedom
## Multiple R-squared:  0.9946, Adjusted R-squared:  0.9944
## F-statistic:  5230 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 9] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep9, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep9, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep9, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep9, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1931.011
## g.div.200kb.2  8 -1922.265
## g.div.200kb.3  7 -1916.833
## g.div.200kb.4  7 -1924.388
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep9
##         AIC       BIC   logLik
##   -1916.833 -1897.438 965.4164
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.082583
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0032663 0.00000920 354.9684  0.0000
## thetaC        0.9842793 0.02012000  48.9204  0.0000
## rhoC          0.0029596 0.00501867   0.5897  0.5566
## tmrcaC        0.0040304 0.00009727  41.4361  0.0000
## thetaC:tmrcaC 1.2099218 0.12627567   9.5816  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.462
## rhoC           0.248 -0.556
## tmrcaC        -0.319 -0.110 -0.435
## thetaC:tmrcaC -0.676  0.345 -0.338  0.492
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.9751365 -0.6158304 -0.1083347  0.6213443  4.6377522
##
## Residual standard error: 1.478625e-07
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.436334      2.397958      2.387627      1.772638
```

5.2.10 Replicate 10

```
rep_10.pi.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5.di
rep_10.pi.200kb$avg <- apply(rep_10.pi.200kb[4:ncol(rep_10.pi.200kb)], 1, mean)
rep_10.rho.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5.r
rep_10.theta.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5
rep_10.tmrca.200kb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5
rep_10.tmrca.200kb$avg <- apply(rep_10.tmrca.200kb[4:ncol(rep_10.tmrca.200kb)], 1, mean)

inf.lands.200kb.rep10 <- as.data.frame(cbind(rep_10.pi.200kb$avg, rep_10.theta.200kb$sample_mean, rep_10
names(inf.lands.200kb.rep10) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.200kb.rep10$thetaC <- inf.lands.200kb.rep10$theta - mean(inf.lands.200kb.rep10$theta)
inf.lands.200kb.rep10$tmrcaC <- inf.lands.200kb.rep10$tmrca - mean(inf.lands.200kb.rep10$tmrca)
inf.lands.200kb.rep10$rhoC <- inf.lands.200kb.rep10$rho - mean(inf.lands.200kb.rep10$rho)

inf.lands.200kb.rep10$bin <- 1:nrow(inf.lands.200kb.rep10)

# for merging:
inf.lands.200kb.rep10 <- inf.lands.200kb.rep10
inf.lands.200kb.rep10$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep10)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep10)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1953.891
## m.div.200kb.2  7 -1957.296
## m.div.200kb.3  8 -1956.404
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.4907, p-value = 0.001364
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.4609, p-value = 0.254
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep10)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.353e-04 -4.230e-05 -2.486e-06  3.708e-05  1.866e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.208e-03  8.005e-06  400.71   <2e-16 ***
## thetaC         9.757e-01  1.418e-02   68.82   <2e-16 ***
## rhoC          -4.932e-03  3.681e-03   -1.34    0.183
## tmrcaC         4.336e-03  8.327e-05   52.08   <2e-16 ***
## thetaC:tmrcaC  1.504e+00  8.914e-02   16.87   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.963e-05 on 113 degrees of freedom
## Multiple R-squared:  0.9966, Adjusted R-squared:  0.9965
## F-statistic:  8373 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 10] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep10, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), m

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep10, weights = varPower(0, ~theta), cor = corAR1(0, ~bin),

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep10, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep10, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1956.857
## g.div.200kb.2  8 -1961.174
## g.div.200kb.3  7 -1962.059
## g.div.200kb.4  7 -1954.326
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep10
##         AIC       BIC   logLik
##   -1962.059 -1942.664 988.0295
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.050657
##
## Coefficients:
##                     Value  Std.Error  t-value p-value
## (Intercept)    0.0032026 0.00000819 390.9569  0.0000
## thetaC         0.9673050 0.01663839  58.1369  0.0000
## rhoC          -0.0036182 0.00342358  -1.0568  0.2928
## tmrcaC         0.0044132 0.00008237  53.5765  0.0000
## thetaC:tmrcaC  1.5967041 0.09757204  16.3644  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.442
## rhoC           0.392 -0.493
## tmrcaC        -0.447 -0.247 -0.319
## thetaC:tmrcaC -0.732  0.327 -0.496  0.612
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.44406069 -0.66547647 -0.04305055  0.56882100  2.81110079
##
## Residual standard error: 1.454558e-07
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.391981      1.820450      2.860658      2.619865
```

5.3 1 Mb scale

```
r2.bgs.1Mb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.bgs.1Mb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.bgs.1Mb) <- reps

sim.theta.1Mb <- read.table("dm_bgs_sims/MutationMap_1000kb.csv", header = T)
sim.rho.1Mb <- read.table("dm_bgs_sims/RecombinationMap_1000kb.csv", header = T)
```

5.3.1 Replicate 1

```
rep_1.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.diversi
rep_1.pi.1Mb$avg <- apply(rep_1.pi.1Mb[4:ncol(rep_1.pi.1Mb)], 1, mean)
rep_1.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.rho.1M
rep_1.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.thet
rep_1.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.TMRC
rep_1.tmrca.1Mb$avg <- apply(rep_1.tmrca.1Mb[4:ncol(rep_1.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep1 <- as.data.frame(cbind(rep_1.pi.1Mb$avg, rep_1.theta.1Mb$sample_mean, rep_1.tmrca.1M
names(inf.lands.1Mb.rep1) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.1Mb.rep1$thetaC <- inf.lands.1Mb.rep1$theta - mean(inf.lands.1Mb.rep1$theta)
inf.lands.1Mb.rep1$tmrcaC <- inf.lands.1Mb.rep1$tmrca - mean(inf.lands.1Mb.rep1$tmrca)
inf.lands.1Mb.rep1$rhoC <- inf.lands.1Mb.rep1$rho - mean(inf.lands.1Mb.rep1$rho)

inf.lands.1Mb.rep1$bin <- 1:nrow(inf.lands.1Mb.rep1)

# for merging:
inf.lands.1Mb.rep1 <- inf.lands.1Mb.rep1
inf.lands.1Mb.rep1$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep1)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep1)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -406.0092
## m.div.1Mb.2  7 -430.3498
```

```
## m.div.1Mb.3  8 -429.6365
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.5555, p-value = 0.08007
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.45959, p-value = 0.376
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep1)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -8.953e-05 -3.030e-05  2.250e-06  2.058e-05  7.938e-05
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.254e-03  1.291e-05 251.954  < 2e-16 ***
## thetaC         9.715e-01  1.576e-02  61.645  < 2e-16 ***
## rhoC          -9.396e-03  8.296e-03  -1.133 0.271502
## tmrcaC         4.173e-03  2.523e-04  16.540 9.75e-13 ***
## thetaC:tmrcaC  9.890e-01  2.369e-01   4.175 0.000513 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.492e-05 on 19 degrees of freedom
## Multiple R-squared:  0.9982, Adjusted R-squared:  0.9978
## F-statistic:  2604 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```r
r2.bgs.1Mb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 1] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep1, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep1, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep1, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep1, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -404.0002
## g.div.1Mb.2  8 -409.1515
## g.div.1Mb.3  7 -410.3386
## g.div.1Mb.4  7 -404.4055
```

```r
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep1
##         AIC       BIC   logLik
##   -410.3386 -402.0922 212.1693
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -2.023167
##
## Coefficients:
##                   Value  Std.Error   t-value p-value
## (Intercept)   0.0032718 0.00001510 216.74597  0.0000
## thetaC        0.9215871 0.03309801  27.84419  0.0000
## rhoC          0.0120765 0.00797878   1.51358  0.1466
## tmrcaC        0.0039579 0.00019161  20.65590  0.0000
## thetaC:tmrcaC 0.6717915 0.25595339   2.62466  0.0167
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.784
## rhoC           0.548 -0.526
## tmrcaC        -0.243 -0.106 -0.551
## thetaC:tmrcaC -0.726  0.554 -0.634  0.360
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.4378441 -0.7324371  0.1090189  0.7573097  1.4770720
##
## Residual standard error: 3.880261e-10
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.427407      2.973115      2.239180      2.020861
```

5.3.2 Replicate 2

```
rep_2.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.diversi
rep_2.pi.1Mb$avg <- apply(rep_2.pi.1Mb[4:ncol(rep_2.pi.1Mb)], 1, mean)
rep_2.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.rho.1M
rep_2.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.thet
rep_2.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.TMRC
rep_2.tmrca.1Mb$avg <- apply(rep_2.tmrca.1Mb[4:ncol(rep_2.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep2 <- as.data.frame(cbind(rep_2.pi.1Mb$avg, rep_2.theta.1Mb$sample_mean, rep_2.tmrca.1Mb
names(inf.lands.1Mb.rep2) <- c("diversity", "theta", "tmrca", "rho")

# centering
inf.lands.1Mb.rep2$thetaC <- inf.lands.1Mb.rep2$theta - mean(inf.lands.1Mb.rep2$theta)
inf.lands.1Mb.rep2$tmrcaC <- inf.lands.1Mb.rep2$tmrca - mean(inf.lands.1Mb.rep2$tmrca)
inf.lands.1Mb.rep2$rhoC <- inf.lands.1Mb.rep2$rho - mean(inf.lands.1Mb.rep2$rho)

inf.lands.1Mb.rep2$bin <- 1:nrow(inf.lands.1Mb.rep2)

# for merging:
inf.lands.1Mb.rep2 <- inf.lands.1Mb.rep2
inf.lands.1Mb.rep2$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep2)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1Mb
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep2)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -399.3496
## m.div.1Mb.2  7 -400.4083
## m.div.1Mb.3  8 -404.4321
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.9912, p-value = 0.4076
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.30667, p-value = 0.097
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.454e-04 -2.155e-05  3.937e-06  1.863e-05  6.974e-05
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.212e-03  1.309e-05 245.351  < 2e-16 ***
## thetaC         1.004e+00  1.562e-02  64.253  < 2e-16 ***
## rhoC          -8.039e-03  8.127e-03  -0.989 0.335038
## tmrcaC         4.377e-03  2.656e-04  16.477 1.04e-12 ***
## thetaC:tmrcaC  1.422e+00  3.310e-01   4.297 0.000389 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.16e-05 on 19 degrees of freedom
## Multiple R-squared:  0.9979, Adjusted R-squared:  0.9975
## F-statistic:  2268 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```r
r2.bgs.1Mb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 2] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep2, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep2, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep2, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep2, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -400.8891
## g.div.1Mb.2  8 -398.2412
## g.div.1Mb.3  7 -400.2042
## g.div.1Mb.4  7 -402.8436
```

```r
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep2
##         AIC       BIC   logLik
##   -400.2042 -391.9578 207.1021
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.212998
##
## Coefficients:
##                    Value Std.Error   t-value p-value
## (Intercept)    0.0032150 0.0000175 184.14687  0.0000
## thetaC         1.0011436 0.0318203  31.46238  0.0000
## rhoC          -0.0029708 0.0095617  -0.31070  0.7594
## tmrcaC         0.0042837 0.0002888  14.83268  0.0000
## thetaC:tmrcaC  1.2596398 0.4834910   2.60530  0.0174
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.811
## rhoC           0.368 -0.444
## tmrcaC         0.227 -0.293 -0.505
## thetaC:tmrcaC -0.732  0.654 -0.334 -0.348
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.3171033 -0.6352535  0.0043190  0.4947650  1.9662715
##
## Residual standard error: 4.267979e-08
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.484863      3.135203      2.875812      2.075453
```

5.3.3 Replicate 3

```
rep_3.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.diversi
rep_3.pi.1Mb$avg <- apply(rep_3.pi.1Mb[4:ncol(rep_3.pi.1Mb)], 1, mean)
rep_3.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.rho.1M
rep_3.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.thet
rep_3.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.TMRC
rep_3.tmrca.1Mb$avg <- apply(rep_3.tmrca.1Mb[4:ncol(rep_3.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep3 <- as.data.frame(cbind(rep_3.pi.1Mb$avg, rep_3.theta.1Mb$sample_mean, rep_3.tmrca.1M
names(inf.lands.1Mb.rep3) <- c("diversity", "theta", "tmrca", "rho")

cor.test(~theta+tmrca, data = inf.lands.1Mb.rep3, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 440, p-value = 3.288e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.8086957
```

```
cor.test(~theta+rho, data = inf.lands.1Mb.rep3, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 770, p-value = 0.0005364
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.6652174
```

```
cor.test(~rho+tmrca, data = inf.lands.1Mb.rep3, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 754, p-value = 0.0004483
```

```
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.6721739
```

```
# centering
inf.lands.1Mb.rep3$thetaC <- inf.lands.1Mb.rep3$theta - mean(inf.lands.1Mb.rep3$theta)
inf.lands.1Mb.rep3$tmrcaC <- inf.lands.1Mb.rep3$tmrca - mean(inf.lands.1Mb.rep3$tmrca)
inf.lands.1Mb.rep3$rhoC <- inf.lands.1Mb.rep3$rho - mean(inf.lands.1Mb.rep3$rho)

inf.lands.1Mb.rep3$bin <- 1:nrow(inf.lands.1Mb.rep3)

# for merging:
inf.lands.1Mb.rep3 <- inf.lands.1Mb.rep3
inf.lands.1Mb.rep3$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep3)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep3)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -409.5885
## m.div.1Mb.2  7 -421.5471
## m.div.1Mb.3  8 -425.6301
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
```

```
## data:  m.div.1Mb
## DW = 1.5916, p-value = 0.08878
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.49696, p-value = 0.464
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -6.175e-05 -2.983e-05 -9.534e-06  2.798e-05  8.075e-05
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.226e-03  1.424e-05 226.584  < 2e-16 ***
## thetaC         8.712e-01  3.284e-02  26.531  < 2e-16 ***
## rhoC          -1.881e-02  8.705e-03  -2.161   0.0437 *
## tmrcaC         5.277e-03  1.800e-04  29.312  < 2e-16 ***
```

```
## thetaC:tmrcaC  1.378e+00  2.275e-01   6.057 7.96e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.169e-05 on 19 degrees of freedom
## Multiple R-squared:  0.9984, Adjusted R-squared:  0.998
## F-statistic:  2900 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 3] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep3, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep3, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), met

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep3, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep3, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -412.9887
## g.div.1Mb.2  8 -421.6118
## g.div.1Mb.3  7 -422.9079
## g.div.1Mb.4  7 -407.8511
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep3
##         AIC       BIC   logLik
##   -422.9079 -414.6615 218.4539
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -3.724078
##
## Coefficients:
##                     Value Std.Error   t-value p-value
```

```
## (Intercept)    0.0032933 0.0000184 178.77408  0.0000
## thetaC         0.7870738 0.0431431  18.24332  0.0000
## rhoC           0.0273497 0.0075170   3.63836  0.0017
## tmrcaC         0.0046109 0.0001840  25.06263  0.0000
## thetaC:tmrcaC -0.2390013 0.3660344  -0.65295  0.5216
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.860
## rhoC           0.566 -0.397
## tmrcaC         0.089 -0.454 -0.450
## thetaC:tmrcaC -0.831  0.651 -0.650 -0.037
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.9969434 -0.7226474 -0.1659388  0.8513819  1.7336059
##
## Residual standard error: 2.191358e-14
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      3.317025      3.688734      3.131145      2.561962
```

5.3.4 Replicate 4

```
rep_4.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.diversi
rep_4.pi.1Mb$avg <- apply(rep_4.pi.1Mb[4:ncol(rep_4.pi.1Mb)], 1, mean)
rep_4.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.rho.1M
rep_4.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.thet
rep_4.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.TMRC
rep_4.tmrca.1Mb$avg <- apply(rep_4.tmrca.1Mb[4:ncol(rep_4.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep4 <- as.data.frame(cbind(rep_4.pi.1Mb$avg, rep_4.theta.1Mb$sample_mean, rep_4.tmrca.1Mb
names(inf.lands.1Mb.rep4) <- c("diversity", "theta", "tmrca", "rho")

cor.test(~theta+tmrca, data = inf.lands.1Mb.rep4, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 500, p-value = 9.977e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.7826087
```

```
cor.test(~theta+rho, data = inf.lands.1Mb.rep4, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
```

```
## S = 940, p-value = 0.002813
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.5913043
```

```
cor.test(~rho+tmrca, data = inf.lands.1Mb.rep4, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 902, p-value = 0.00201
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.6078261
```

```
# centering
inf.lands.1Mb.rep4$thetaC <- inf.lands.1Mb.rep4$theta - mean(inf.lands.1Mb.rep4$theta)
inf.lands.1Mb.rep4$tmrcaC <- inf.lands.1Mb.rep4$tmrca - mean(inf.lands.1Mb.rep4$tmrca)
inf.lands.1Mb.rep4$rhoC <- inf.lands.1Mb.rep4$rho - mean(inf.lands.1Mb.rep4$rho)

inf.lands.1Mb.rep4$bin <- 1:nrow(inf.lands.1Mb.rep4)

# for merging:
inf.lands.1Mb.rep4 <- inf.lands.1Mb.rep4
inf.lands.1Mb.rep4$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep4)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep4)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -397.3660
## m.div.1Mb.2  7 -427.2283
## m.div.1Mb.3  8 -446.7941
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.3655, p-value = 0.02222
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.48067, p-value = 0.477
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -9.179e-05 -2.866e-05  2.328e-06  3.847e-05  1.106e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.238e-03  1.922e-05 168.487  < 2e-16 ***
## thetaC         1.010e+00  3.264e-02  30.927  < 2e-16 ***
## rhoC          -4.312e-02  8.617e-03  -5.004 7.88e-05 ***
## tmrcaC         5.734e-03  3.744e-04  15.312 3.83e-12 ***
## thetaC:tmrcaC  1.420e+00  4.169e-01   3.406  0.00296 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.378e-05 on 19 degrees of freedom
## Multiple R-squared:  0.9974, Adjusted R-squared:  0.9968
## F-statistic:  1813 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```r
r2.bgs.1Mb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 4] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep4, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep4, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep4, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep4, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -404.8346
## g.div.1Mb.2  8 -424.5156
## g.div.1Mb.3  7 -422.1707
## g.div.1Mb.4  7 -400.0480
```

```r
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep4
##         AIC       BIC   logLik
##   -422.1707 -413.9243 218.0853
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -4.737286
##
## Coefficients:
##                   Value Std.Error   t-value p-value
## (Intercept)   0.0032689 0.0000159 205.75108  0.0000
## thetaC        0.9760242 0.0361547  26.99576  0.0000
## rhoC          0.0039813 0.0046685   0.85280  0.4044
## tmrcaC        0.0047403 0.0002452  19.33405  0.0000
## thetaC:tmrcaC 0.8318990 0.4600537   1.80826  0.0864
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.883
## rhoC           0.383 -0.230
## tmrcaC         0.246 -0.494 -0.334
## thetaC:tmrcaC -0.601  0.469 -0.337 -0.505
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -2.6711787 -0.6405707  0.1119790  0.6454920  1.4954806
##
## Residual standard error: 7.05933e-17
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.724453      2.187442      2.913950      2.187653
```

5.3.5 Replicate 5

```
rep_5.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.diversi
rep_5.pi.1Mb$avg <- apply(rep_5.pi.1Mb[4:ncol(rep_5.pi.1Mb)], 1, mean)
rep_5.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.rho.1M
rep_5.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.thet
rep_5.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.TMRC
rep_5.tmrca.1Mb$avg <- apply(rep_5.tmrca.1Mb[4:ncol(rep_5.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep5 <- as.data.frame(cbind(rep_5.pi.1Mb$avg, rep_5.theta.1Mb$sample_mean, rep_5.tmrca.1Mb
names(inf.lands.1Mb.rep5) <- c("diversity", "theta", "tmrca", "rho")

cor.test(~theta+tmrca, data = inf.lands.1Mb.rep5, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 268, p-value = 2.67e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.8834783
```

```
cor.test(~theta+rho, data = inf.lands.1Mb.rep5, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 1460, p-value = 0.08
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.3652174
```

```
cor.test(~rho+tmrca, data = inf.lands.1Mb.rep5, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 1062, p-value = 0.007459
```

```
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.5382609
```

```
# centering
inf.lands.1Mb.rep5$thetaC <- inf.lands.1Mb.rep5$theta - mean(inf.lands.1Mb.rep5$theta)
inf.lands.1Mb.rep5$tmrcaC <- inf.lands.1Mb.rep5$tmrca - mean(inf.lands.1Mb.rep5$tmrca)
inf.lands.1Mb.rep5$rhoC <- inf.lands.1Mb.rep5$rho - mean(inf.lands.1Mb.rep5$rho)

inf.lands.1Mb.rep5$bin <- 1:nrow(inf.lands.1Mb.rep5)

# for merging:
inf.lands.1Mb.rep5 <- inf.lands.1Mb.rep5
inf.lands.1Mb.rep5$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep5)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep5)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -382.7189
## m.div.1Mb.2  7 -416.9266
## m.div.1Mb.3  8 -419.8725
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
```

```
## data:  m.div.1Mb
## DW = 1.7255, p-value = 0.1405
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.138, p-value = 0.002
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep5)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.491e-04 -3.724e-05 -6.155e-06  3.251e-05  1.498e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.254e-03  2.031e-05 160.184  < 2e-16 ***
## thetaC         9.496e-01  3.075e-02  30.884  < 2e-16 ***
## rhoC          -3.654e-02  1.253e-02  -2.916  0.00886 **
## tmrcaC         5.750e-03  5.084e-04  11.310    7e-10 ***
```

```
## thetaC:tmrcaC  3.073e-01  4.749e-01   0.647  0.52524
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.297e-05 on 19 degrees of freedom
## Multiple R-squared:  0.9954, Adjusted R-squared:  0.9945
## F-statistic:  1034 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 5] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep5, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep5, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), met

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep5, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep5, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -405.5094
## g.div.1Mb.2  8 -416.4060
## g.div.1Mb.3  7 -417.3755
## g.div.1Mb.4  7 -404.1949
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep5
##         AIC       BIC   logLik
##   -417.3755 -409.1291 215.6878
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -3.557677
##
## Coefficients:
##                    Value Std.Error   t-value p-value
```

```
## (Intercept)   0.0032539 0.0000171 190.36062  0.0000
## thetaC        0.9547500 0.0341476  27.95952  0.0000
## rhoC          0.0135203 0.0067801   1.99412  0.0607
## tmrcaC        0.0041562 0.0003606  11.52515  0.0000
## thetaC:tmrcaC 1.5129197 0.6394793   2.36586  0.0288
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.934
## rhoC           0.246 -0.150
## tmrcaC         0.417 -0.495 -0.425
## thetaC:tmrcaC -0.608  0.546 -0.103 -0.770
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.8935863 -0.7573840  0.2241411  0.7328966  1.5503079
##
## Residual standard error: 5.423247e-14
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.686193      3.183125      7.798430      5.560645
```

5.3.6 Replicate 6

```
rep_6.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.diversi
rep_6.pi.1Mb$avg <- apply(rep_6.pi.1Mb[4:ncol(rep_6.pi.1Mb)], 1, mean)
rep_6.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.rho.1M
rep_6.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.thet
rep_6.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.TMRC
rep_6.tmrca.1Mb$avg <- apply(rep_6.tmrca.1Mb[4:ncol(rep_6.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep6 <- as.data.frame(cbind(rep_6.pi.1Mb$avg, rep_6.theta.1Mb$sample_mean, rep_6.tmrca.1Mb
names(inf.lands.1Mb.rep6) <- c("diversity", "theta", "tmrca", "rho")

cor.test(~theta+tmrca, data = inf.lands.1Mb.rep6, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 442, p-value = 3.392e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.8078261
```

```
cor.test(~theta+rho, data = inf.lands.1Mb.rep6, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
```

```
## S = 550, p-value = 2.535e-05
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.7608696
```

```
cor.test(~rho+tmrca, data = inf.lands.1Mb.rep6, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 392, p-value = 2.091e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.8295652
```

```
# centering
inf.lands.1Mb.rep6$thetaC <- inf.lands.1Mb.rep6$theta - mean(inf.lands.1Mb.rep6$theta)
inf.lands.1Mb.rep6$tmrcaC <- inf.lands.1Mb.rep6$tmrca - mean(inf.lands.1Mb.rep6$tmrca)
inf.lands.1Mb.rep6$rhoC <- inf.lands.1Mb.rep6$rho - mean(inf.lands.1Mb.rep6$rho)

inf.lands.1Mb.rep6$bin <- 1:nrow(inf.lands.1Mb.rep6)

# for merging:
inf.lands.1Mb.rep6 <- inf.lands.1Mb.rep6
inf.lands.1Mb.rep6$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep6)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep6)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -415.3703
## m.div.1Mb.2  7 -417.0328
## m.div.1Mb.3  8 -415.8507
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.6421, p-value = 0.08247
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.44033, p-value = 0.317
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.385e-05 -2.409e-05 -8.850e-06  2.257e-05  7.852e-05
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.221e-03  1.452e-05 221.916  < 2e-16 ***
## thetaC         8.720e-01  2.701e-02  32.287  < 2e-16 ***
## rhoC          -2.162e-03  6.894e-03  -0.314  0.75725
## tmrcaC         4.418e-03  7.076e-05  62.438  < 2e-16 ***
## thetaC:tmrcaC  8.861e-01  2.794e-01   3.172  0.00502 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.696e-05 on 19 degrees of freedom
## Multiple R-squared:  0.9989, Adjusted R-squared:  0.9987
## F-statistic:  4266 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.bgs.1Mb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 6] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep6, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep6, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep6, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep6, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -419.5980
## g.div.1Mb.2  8 -414.2332
## g.div.1Mb.3  7 -416.0917
## g.div.1Mb.4  7 -420.2497
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep6
##         AIC       BIC   logLik
##   -416.0917 -407.8454 215.0459
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.520312
##
## Coefficients:
##                    Value Std.Error   t-value p-value
## (Intercept)   0.0032272 0.0000172 187.11448  0.0000
## thetaC        0.8493976 0.0323282  26.27419  0.0000
## rhoC          0.0010118 0.0057291   0.17661  0.8617
## tmrcaC        0.0044253 0.0000679  65.21370  0.0000
## thetaC:tmrcaC 0.7973603 0.3536114   2.25491  0.0361
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.786
## rhoC           0.436 -0.580
## tmrcaC        -0.057 -0.130 -0.388
## thetaC:tmrcaC -0.886  0.706 -0.483  0.109
```

```
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.4816278 -0.6660409 -0.3104259  0.5962878  2.2696281
##
## Residual standard error: 6.080401e-09
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      3.222770      2.256455      1.632443      2.180967
```

5.3.7 Replicate 7

```
rep_7.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.diversi
rep_7.pi.1Mb$avg <- apply(rep_7.pi.1Mb[4:ncol(rep_7.pi.1Mb)], 1, mean)
rep_7.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.rho.1M
rep_7.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.thet
rep_7.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.TMRC
rep_7.tmrca.1Mb$avg <- apply(rep_7.tmrca.1Mb[4:ncol(rep_7.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep7 <- as.data.frame(cbind(rep_7.pi.1Mb$avg, rep_7.theta.1Mb$sample_mean, rep_7.tmrca.1Mb
names(inf.lands.1Mb.rep7) <- c("diversity", "theta", "tmrca", "rho")

cor.test(~theta+tmrca, data = inf.lands.1Mb.rep7, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 442, p-value = 3.392e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.8078261
```

```
cor.test(~theta+rho, data = inf.lands.1Mb.rep7, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 550, p-value = 2.535e-05
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.7608696
```

```
cor.test(~rho+tmrca, data = inf.lands.1Mb.rep7, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 392, p-value = 2.091e-06
```

```
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.8295652
```

```
# centering
inf.lands.1Mb.rep7$thetaC <- inf.lands.1Mb.rep7$theta - mean(inf.lands.1Mb.rep7$theta)
inf.lands.1Mb.rep7$tmrcaC <- inf.lands.1Mb.rep7$tmrca - mean(inf.lands.1Mb.rep7$tmrca)
inf.lands.1Mb.rep7$rhoC <- inf.lands.1Mb.rep7$rho - mean(inf.lands.1Mb.rep7$rho)

inf.lands.1Mb.rep7$bin <- 1:nrow(inf.lands.1Mb.rep7)

# for merging:
inf.lands.1Mb.rep7 <- inf.lands.1Mb.rep7
inf.lands.1Mb.rep7$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep7)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep7)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -415.3703
## m.div.1Mb.2  7 -417.0328
## m.div.1Mb.3  8 -415.8507
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
```

```
##
## data:  m.div.1Mb
## DW = 1.6421, p-value = 0.08247
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.44033, p-value = 0.331
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep7)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.385e-05 -2.409e-05 -8.850e-06  2.257e-05  7.852e-05
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.221e-03  1.452e-05 221.916  < 2e-16 ***
## thetaC         8.720e-01  2.701e-02  32.287  < 2e-16 ***
## rhoC          -2.162e-03  6.894e-03  -0.314  0.75725
```

```
## tmrcaC         4.418e-03  7.076e-05  62.438  < 2e-16 ***
## thetaC:tmrcaC  8.861e-01  2.794e-01   3.172  0.00502 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.696e-05 on 19 degrees of freedom
## Multiple R-squared:  0.9989, Adjusted R-squared:  0.9987
## F-statistic:  4266 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 7] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep7, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep7, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep7, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep7, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -419.5980
## g.div.1Mb.2  8 -414.2332
## g.div.1Mb.3  7 -416.0917
## g.div.1Mb.4  7 -420.2497
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep7
##         AIC       BIC   logLik
##   -416.0917 -407.8454 215.0459
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -1.520312
##
## Coefficients:
```

```
##                   Value Std.Error   t-value p-value
## (Intercept)   0.0032272 0.0000172 187.11448  0.0000
## thetaC        0.8493976 0.0323282  26.27419  0.0000
## rhoC          0.0010118 0.0057291   0.17661  0.8617
## tmrcaC        0.0044253 0.0000679  65.21370  0.0000
## thetaC:tmrcaC 0.7973603 0.3536114   2.25491  0.0361
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.786
## rhoC           0.436 -0.580
## tmrcaC        -0.057 -0.130 -0.388
## thetaC:tmrcaC -0.886  0.706 -0.483  0.109
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.4816278 -0.6660409 -0.3104259  0.5962878  2.2696281
##
## Residual standard error: 6.080401e-09
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      3.222770      2.256455      1.632443      2.180967
```

5.3.8 Replicate 8

```
rep_8.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.diversi
rep_8.pi.1Mb$avg <- apply(rep_8.pi.1Mb[4:ncol(rep_8.pi.1Mb)], 1, mean)
rep_8.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.rho.1M
rep_8.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.thet
rep_8.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.TMRC
rep_8.tmrca.1Mb$avg <- apply(rep_8.tmrca.1Mb[4:ncol(rep_8.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep8 <- as.data.frame(cbind(rep_8.pi.1Mb$avg, rep_8.theta.1Mb$sample_mean, rep_8.tmrca.1Mb
names(inf.lands.1Mb.rep8) <- c("diversity", "theta", "tmrca", "rho")

cor.test(~theta+tmrca, data = inf.lands.1Mb.rep8, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 318, p-value = 2.389e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.8617391
```

```
cor.test(~theta+rho, data = inf.lands.1Mb.rep8, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
```

```
## data:  theta and rho
## S = 1174, p-value = 0.01624
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.4895652
```

```
cor.test(~rho+tmrca, data = inf.lands.1Mb.rep8, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 698, p-value = 0.0002294
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.6965217
```

```
# centering
inf.lands.1Mb.rep8$thetaC <- inf.lands.1Mb.rep8$theta - mean(inf.lands.1Mb.rep8$theta)
inf.lands.1Mb.rep8$tmrcaC <- inf.lands.1Mb.rep8$tmrca - mean(inf.lands.1Mb.rep8$tmrca)
inf.lands.1Mb.rep8$rhoC <- inf.lands.1Mb.rep8$rho - mean(inf.lands.1Mb.rep8$rho)

inf.lands.1Mb.rep8$bin <- 1:nrow(inf.lands.1Mb.rep8)

# for merging:
inf.lands.1Mb.rep8 <- inf.lands.1Mb.rep8
inf.lands.1Mb.rep8$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep8)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep8)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -408.1140
## m.div.1Mb.2  7 -431.5240
## m.div.1Mb.3  8 -437.7876
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.8179, p-value = 0.2531
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.49032, p-value = 0.464
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep8)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -6.508e-05 -2.315e-05 -6.980e-06  3.657e-05  7.987e-05
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.187e-03  1.665e-05 191.449  < 2e-16 ***
## thetaC         9.816e-01  2.185e-02  44.935  < 2e-16 ***
## rhoC          -3.638e-02  6.222e-03  -5.847 1.24e-05 ***
## tmrcaC         4.835e-03  2.131e-04  22.685 3.19e-15 ***
## thetaC:tmrcaC  1.624e+00  4.018e-01   4.042 0.000696 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.299e-05 on 19 degrees of freedom
## Multiple R-squared:  0.9984, Adjusted R-squared:  0.9981
## F-statistic:  2956 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```
r2.bgs.1Mb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 8] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep8, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep8, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep8, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep8, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -406.1310
## g.div.1Mb.2  8 -404.6132
## g.div.1Mb.3  7 -406.5578
## g.div.1Mb.4  7 -406.1844
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep8
##         AIC       BIC   logLik
##   -406.5578 -398.3115 210.2789
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.7018945
##
## Coefficients:
##                    Value Std.Error   t-value p-value
## (Intercept)    0.0031854 0.0000159 199.91712   0e+00
## thetaC         0.9850537 0.0158532  62.13591   0e+00
## rhoC          -0.0389530 0.0050369  -7.73357   0e+00
## tmrcaC         0.0048569 0.0002006  24.20959   0e+00
## thetaC:tmrcaC  1.6762142 0.3731008   4.49266   2e-04
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.347
## rhoC           0.456  0.137
## tmrcaC        -0.340 -0.343 -0.667
## thetaC:tmrcaC -0.827  0.553 -0.545  0.414
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -1.87875767 -0.65584520 -0.04899389  0.92099803  1.99185455
##
## Residual standard error: 0.00201201
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      3.805533      2.448838      2.910652      4.725060
```

5.3.9 Replicate 9

```
rep_9.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.diversi
rep_9.pi.1Mb$avg <- apply(rep_9.pi.1Mb[4:ncol(rep_9.pi.1Mb)], 1, mean)
rep_9.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.rho.1M
rep_9.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.thet
rep_9.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.TMRC
rep_9.tmrca.1Mb$avg <- apply(rep_9.tmrca.1Mb[4:ncol(rep_9.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep9 <- as.data.frame(cbind(rep_9.pi.1Mb$avg, rep_9.theta.1Mb$sample_mean, rep_9.tmrca.1Mb
names(inf.lands.1Mb.rep9) <- c("diversity", "theta", "tmrca", "rho")

cor.test(~theta+tmrca, data = inf.lands.1Mb.rep9, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 206, p-value = 2.614e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9104348
```

```
cor.test(~theta+rho, data = inf.lands.1Mb.rep9, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
## S = 192, p-value = 2.54e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9165217
```

```
cor.test(~rho+tmrca, data = inf.lands.1Mb.rep9, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 178, p-value = 2.448e-06
```

```
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9226087
```

```
# centering
inf.lands.1Mb.rep9$thetaC <- inf.lands.1Mb.rep9$theta - mean(inf.lands.1Mb.rep9$theta)
inf.lands.1Mb.rep9$tmrcaC <- inf.lands.1Mb.rep9$tmrca - mean(inf.lands.1Mb.rep9$tmrca)
inf.lands.1Mb.rep9$rhoC <- inf.lands.1Mb.rep9$rho - mean(inf.lands.1Mb.rep9$rho)

inf.lands.1Mb.rep9$bin <- 1:nrow(inf.lands.1Mb.rep9)

# for merging:
inf.lands.1Mb.rep9 <- inf.lands.1Mb.rep9
inf.lands.1Mb.rep9$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep9)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep9)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -406.1573
## m.div.1Mb.2  7 -408.0519
## m.div.1Mb.3  8 -408.6284
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
```

```
## data:  m.div.1Mb
## DW = 1.3992, p-value = 0.03732
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.19756, p-value = 0.008
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep9)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -8.630e-05 -2.169e-05 -1.068e-05  2.216e-05  1.198e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.239e-03  1.633e-05 198.380  < 2e-16 ***
## thetaC         9.538e-01  3.045e-02  31.321  < 2e-16 ***
## rhoC          -2.836e-03  1.216e-02  -0.233 0.818000
## tmrcaC         4.134e-03  2.838e-04  14.568 9.19e-12 ***
```

```
## thetaC:tmrcaC  1.118e+00  2.777e-01   4.024 0.000725 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.478e-05 on 19 degrees of freedom
## Multiple R-squared:  0.998,  Adjusted R-squared:  0.9976
## F-statistic:  2400 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 9] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep9, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep9, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), met

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep9, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep9, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -410.8860
## g.div.1Mb.2  8 -414.3732
## g.div.1Mb.3  7 -414.5030
## g.div.1Mb.4  7 -410.7982
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep9
##        AIC       BIC   logLik
##   -414.503 -406.2567 214.2515
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -2.373641
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
```

```
## (Intercept)   0.0032415 0.00001434 226.07447  0.0000
## thetaC        0.9424435 0.03770148  24.99752  0.0000
## rhoC          0.0042696 0.00765516   0.55775  0.5835
## tmrcaC        0.0040464 0.00022530  17.95980  0.0000
## thetaC:tmrcaC 1.0584478 0.24614326   4.30013  0.0004
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.552
## rhoC           0.394 -0.374
## tmrcaC        -0.226 -0.436 -0.480
## thetaC:tmrcaC -0.760  0.308 -0.462  0.352
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.7340864 -0.8630162 -0.1808558  0.9468456  1.8219562
##
## Residual standard error: 4.687009e-11
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      3.416953      2.885606      3.824060      1.644120
```

5.3.10 Replicate 10

```
rep_10.pi.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5.div
rep_10.pi.1Mb$avg <- apply(rep_10.pi.1Mb[4:ncol(rep_10.pi.1Mb)], 1, mean)
rep_10.rho.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5.rho
rep_10.theta.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5.
rep_10.tmrca.1Mb <- read.table("dm_bgs_sims/10r/Homogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5.
rep_10.tmrca.1Mb$avg <- apply(rep_10.tmrca.1Mb[4:ncol(rep_10.tmrca.1Mb)], 1, mean)

inf.lands.1Mb.rep10 <- as.data.frame(cbind(rep_10.pi.1Mb$avg, rep_10.theta.1Mb$sample_mean, rep_10.tmrca
names(inf.lands.1Mb.rep10) <- c("diversity", "theta", "tmrca", "rho")

cor.test(~theta+tmrca, data = inf.lands.1Mb.rep10, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and tmrca
## S = 302, p-value = 2.503e-06
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.8686957
```

```
cor.test(~theta+rho, data = inf.lands.1Mb.rep10, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  theta and rho
```

```
## S = 830, p-value = 0.00101
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.6391304
```

```
cor.test(~rho+tmrca, data = inf.lands.1Mb.rep10, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  rho and tmrca
## S = 796, p-value = 0.000711
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.653913
```

```
# centering
inf.lands.1Mb.rep10$thetaC <- inf.lands.1Mb.rep10$theta - mean(inf.lands.1Mb.rep10$theta)
inf.lands.1Mb.rep10$tmrcaC <- inf.lands.1Mb.rep10$tmrca - mean(inf.lands.1Mb.rep10$tmrca)
inf.lands.1Mb.rep10$rhoC <- inf.lands.1Mb.rep10$rho - mean(inf.lands.1Mb.rep10$rho)

inf.lands.1Mb.rep10$bin <- 1:nrow(inf.lands.1Mb.rep10)

# for merging:
inf.lands.1Mb.rep10 <- inf.lands.1Mb.rep10
inf.lands.1Mb.rep10$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep10)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep10)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -398.4287
## m.div.1Mb.2  7 -403.8113
## m.div.1Mb.3  8 -416.8144
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.0519, p-value = 0.001454
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.51438, p-value = 0.554
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep10)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.202e-04 -2.103e-05 -4.762e-06  1.704e-05  1.463e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.232e-03  2.255e-05 143.313  < 2e-16 ***
## thetaC        8.964e-01  4.162e-02  21.537 8.26e-15 ***
## rhoC          7.052e-03  1.312e-02   0.538   0.5971
## tmrcaC        3.986e-03  2.222e-04  17.942 2.27e-13 ***
## thetaC:tmrcaC 6.957e-01  3.632e-01   1.915   0.0706 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.26e-05 on 19 degrees of freedom
## Multiple R-squared:  0.9974, Adjusted R-squared:  0.9969
## F-statistic:  1824 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
```

```r
r2.bgs.1Mb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 10] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep10, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep10, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep10, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep10, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -427.2196
## g.div.1Mb.2  8 -416.1866
## g.div.1Mb.3  7 -412.5449
## g.div.1Mb.4  7 -426.7420
```

```r
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep10
##         AIC       BIC   logLik
##   -412.5449 -404.2985 213.2724
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## -6.942024
##
## Coefficients:
##                    Value Std.Error   t-value p-value
## (Intercept)    0.0031813 0.0000185 171.77091  0.0000
## thetaC         0.9219060 0.0533941  17.26607  0.0000
## rhoC          -0.0047991 0.0038578  -1.24402  0.2286
## tmrcaC         0.0041277 0.0002063  20.00655  0.0000
## thetaC:tmrcaC  2.2865764 0.3720056   6.14662  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC        -0.790
## rhoC           0.314 -0.127
## tmrcaC         0.212 -0.606 -0.220
## thetaC:tmrcaC -0.402  0.077 -0.364 -0.253
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.37591517 -0.48916830 -0.07381683  0.49448374  1.97637780
##
## Residual standard error: 2.783215e-22
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.951920      1.602199      2.330619      1.459781
```

5.4 all replicates:

```
r2.bgs.50kb$average <- apply(r2.bgs.50kb, 1, mean)
r2.bgs.200kb$average <- apply(r2.bgs.200kb, 1, mean)
r2.bgs.1Mb$average <- apply(r2.bgs.1Mb, 1, mean)

# inferred landscapes
r2.inf.avg <- rbind.data.frame(r2.bgs.50kb$average, r2.bgs.200kb$average, r2.bgs.1Mb$average, make.row.
colnames(r2.inf.avg) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
r2.inf.avg$bin.size <- c(50, 200, 1000)

r2.tab.comb <- rbind.data.frame(r2.inf.avg)

molten.r2 <- melt(r2.tab.comb, id.vars = "bin.size")
r2.plot <- ggplot(data = molten.r2, aes(x = bin.size, y = value, colour = variable))
r2.plot <- r2.plot + geom_line(data = molten.r2)
r2.plot <- r2.plot + geom_point(aes(shape = variable), size = 4)
r2.plot <- r2.plot + scale_x_continuous(breaks = c(50, 200, 1000), trans="log10")
r2.plot <- r2.plot + scale_y_continuous(breaks = pretty_breaks())
r2.plot <- r2.plot + scale_shape_manual(values = c(0, 1, 2, 3, 4))
r2.plot <- r2.plot + labs(title = NULL, x = "Bin Size (kb)", y = "Variance Explained (%)") + theme_bw()
r2.plot <- r2.plot + theme(axis.title = element_text(size = 20), axis.text = element_text(size = 16))
r2.plot
```

6 Simulations with BGS and Gamma Mutation Landscape in SLiM

```
cor.table.theta <- as.data.frame(matrix(ncol = 3, nrow = nreps))
row.names(cor.table.theta) <- reps
colnames(cor.table.theta) <- c("50kb", "200kb", "1Mb")

cor.table.rho <- as.data.frame(matrix(ncol = 3, nrow = nreps))
row.names(cor.table.rho) <- reps
colnames(cor.table.rho) <- c("50kb", "200kb", "1Mb")

cor.table.tmrca <- as.data.frame(matrix(ncol = 3, nrow = nreps))
row.names(cor.table.tmrca) <- reps
colnames(cor.table.tmrca) <- c("50kb", "200kb", "1Mb")
```

6.1 50 kb scale

```
r2.bgs.50kb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.bgs.50kb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.bgs.50kb) <- reps

r2.bgs.50kb.true <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.bgs.50kb.true) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.bgs.50kb.true) <- reps

sim.theta.50kb <- read.table("dm_bgs_sims/MutationMap_50kb.csv", sep = ",", header = T)
```

```
sim.rho.50kb <- read.table("dm_bgs_sims/RecombinationMap_50kb.csv", sep = ",", header = T)
```

6.1.1 Replicate 1

```
rep1.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep1_w50000.csv", sep = ",", header = T)

rep_1.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.di
rep_1.pi.50kb$avg <- apply(rep_1.pi.50kb[4:ncol(rep_1.pi.50kb)], 1, mean)
rep_1.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.rl
rep_1.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5
rep_1.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5
rep_1.tmrca.50kb$avg <- apply(rep_1.tmrca.50kb[4:ncol(rep_1.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep1 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep1.sim.tmrca.50kb$
names(true.lands.50kb.rep1) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep1$thetaC <- true.lands.50kb.rep1$theta - mean(true.lands.50kb.rep1$theta)
true.lands.50kb.rep1$tmrcaC <- true.lands.50kb.rep1$tmrca - mean(true.lands.50kb.rep1$tmrca)
true.lands.50kb.rep1$rhoC <- true.lands.50kb.rep1$rho - mean(true.lands.50kb.rep1$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep1)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.50kb.true[2, 1] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb.true[3, 1] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 1] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 1] <- anova.diversity$VarExp[4] * 100


inf.lands.50kb.rep1 <- as.data.frame(cbind(rep_1.pi.50kb$avg, rep_1.theta.50kb$sample_mean, rep_1.tmrca
names(inf.lands.50kb.rep1) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[1, 1] <- cor.test(sim.theta.50kb$Rate, rep_1.theta.50kb$sample_mean, method = "spearman
cor.table.rho[1, 1] <- cor.test(sim.rho.50kb$Rate, rep_1.rho.50kb$sample_mean, method = "spearman")$est
cor.table.tmrca[1, 1] <- cor.test(rep1.sim.tmrca.50kb$AverageTmrca, rep_1.tmrca.50kb$avg, method = "spea

# centering
inf.lands.50kb.rep1$thetaC <- inf.lands.50kb.rep1$theta - mean(inf.lands.50kb.rep1$theta)
inf.lands.50kb.rep1$tmrcaC <- inf.lands.50kb.rep1$tmrca - mean(inf.lands.50kb.rep1$tmrca)
inf.lands.50kb.rep1$rhoC <- inf.lands.50kb.rep1$rho - mean(inf.lands.50kb.rep1$rho)

inf.lands.50kb.rep1$bin <- 1:nrow(inf.lands.50kb.rep1)

# for merging:
inf.lands.50kb.rep1 <- inf.lands.50kb.rep1
inf.lands.50kb.rep1$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep1)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep1)
```

```r
AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -6711.225
## m.div.50kb.2  7 -6769.917
## m.div.50kb.3  8 -6773.287
```

```r
plot(resid(m.div.50kb.2)~fitted(m.div.50kb.2))
```

```r
dwtest(m.div.50kb.2)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb.2
## DW = 0.87081, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb.2)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb.2
## HMC = 0.44862, p-value = 0.071
```

```r
hist(resid(m.div.50kb.2))
```

```r
summary(m.div.50kb.2)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC +
##     rhoC:tmrcaC, data = inf.lands.50kb.rep1)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -7.520e-04 -7.777e-05  9.490e-06  1.025e-04  8.472e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.408e-03  9.997e-06 340.914  < 2e-16 ***
## thetaC        1.101e+00  5.095e-03 216.158  < 2e-16 ***
## rhoC          3.964e-02  5.571e-03   7.115 4.27e-12 ***
## tmrcaC        4.483e-03  8.032e-05  55.823  < 2e-16 ***
## thetaC:tmrcaC 1.359e+00  2.844e-02  47.792  < 2e-16 ***
## rhoC:tmrcaC   2.809e-01  3.513e-02   7.997 1.02e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001815 on 465 degrees of freedom
## Multiple R-squared:  0.9949, Adjusted R-squared:  0.9948
## F-statistic: 1.8e+04 on 5 and 465 DF,  p-value: < 2.2e-16
```

```r
anova.diversity <- Anova(m.div.50kb.2)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.bgs.50kb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.50kb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 1] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep1, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep1, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep1, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep1, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6931.989
## g.div.50kb.2  8 -6938.978
## g.div.50kb.3  7 -6759.275
## g.div.50kb.4  7 -6765.627
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep1
##         AIC       BIC   logLik
##   -6759.275 -6730.191 3386.638
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.4736601
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0034561 0.000009204 375.4931       0
## thetaC        1.1086892 0.005953385 186.2284       0
## rhoC          0.0521215 0.005177600  10.0667       0
## tmrcaC        0.0039010 0.000044082  88.4950       0
## thetaC:tmrcaC 1.3464303 0.028940376  46.5243       0
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.479
## rhoC           0.007 -0.287
## tmrcaC        -0.099 -0.069 -0.335
```

```
## thetaC:tmrcaC -0.188 -0.187 -0.033  0.524
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.34679876 -0.51809214  0.05652429  0.54157338  4.29208524
##
## Residual standard error: 0.002974804
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.145853      1.284479      1.614090      1.447901
```

6.1.2 Replicate 2

```
rep2.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep2_w50000.csv", sep = ",", header = T)

rep_2.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.di
rep_2.pi.50kb$avg <- apply(rep_2.pi.50kb[4:ncol(rep_2.pi.50kb)], 1, mean)
rep_2.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.rl
rep_2.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5
rep_2.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5
rep_2.tmrca.50kb$avg <- apply(rep_2.tmrca.50kb[4:ncol(rep_2.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep2 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep2.sim.tmrca.50kb$
names(true.lands.50kb.rep2) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep2$thetaC <- true.lands.50kb.rep2$theta - mean(true.lands.50kb.rep2$theta)
true.lands.50kb.rep2$tmrcaC <- true.lands.50kb.rep2$tmrca - mean(true.lands.50kb.rep2$tmrca)
true.lands.50kb.rep2$rhoC <- true.lands.50kb.rep2$rho - mean(true.lands.50kb.rep2$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep2)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.50kb.true[2, 2] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb.true[3, 2] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 2] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 2] <- anova.diversity$VarExp[4] * 100

inf.lands.50kb.rep2 <- as.data.frame(cbind(rep_2.pi.50kb$avg, rep_2.theta.50kb$sample_mean, rep_2.tmrca
names(inf.lands.50kb.rep2) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[2, 1] <- cor.test(sim.theta.50kb$Rate, rep_2.theta.50kb$sample_mean, method = "spearman
cor.table.rho[2, 1] <- cor.test(sim.rho.50kb$Rate, rep_2.rho.50kb$sample_mean, method = "spearman")$est
cor.table.tmrca[2, 1] <- cor.test(rep2.sim.tmrca.50kb$AverageTmrca, rep_2.tmrca.50kb$avg, method = "spe

# centering
inf.lands.50kb.rep2$thetaC <- inf.lands.50kb.rep2$theta - mean(inf.lands.50kb.rep2$theta)
inf.lands.50kb.rep2$tmrcaC <- inf.lands.50kb.rep2$tmrca - mean(inf.lands.50kb.rep2$tmrca)
inf.lands.50kb.rep2$rhoC <- inf.lands.50kb.rep2$rho - mean(inf.lands.50kb.rep2$rho)
```

```
inf.lands.50kb.rep2$bin <- 1:nrow(inf.lands.50kb.rep2)

# for merging:
inf.lands.50kb.rep2 <- inf.lands.50kb.rep2
inf.lands.50kb.rep2$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep2)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep2)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -6649.218
## m.div.50kb.2  7 -6767.980
## m.div.50kb.3  8 -6779.854
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 0.92098, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
```

```
## HMC = 0.4077, p-value = 0.003
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -8.740e-04 -1.030e-04  1.349e-05  1.126e-04  4.971e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.476e-03  9.569e-06  363.28   <2e-16 ***
## thetaC        1.275e+00  6.422e-03  198.57   <2e-16 ***
## rhoC          1.169e-02  5.106e-03    2.29   0.0225 *
## tmrcaC        3.372e-03  3.228e-05  104.47   <2e-16 ***
## thetaC:tmrcaC 1.241e+00  2.277e-02   54.51   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0002066 on 466 degrees of freedom
## Multiple R-squared:  0.9935, Adjusted R-squared:  0.9934
## F-statistic: 1.781e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
```

```r
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.50kb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 2] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep2, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep2, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep2, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep2, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6821.900
## g.div.50kb.2  8 -6850.592
## g.div.50kb.3  7 -6682.201
## g.div.50kb.4  7 -6679.873
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep2
##         AIC       BIC logLik
##   -6682.201 -6653.117 3348.1
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.3935279
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0034782 0.000009638 360.8871   0e+00
## thetaC        1.2721750 0.007190662 176.9204   0e+00
## rhoC          0.0171851 0.004879366   3.5220   5e-04
## tmrcaC        0.0033612 0.000032505 103.4068   0e+00
## thetaC:tmrcaC 1.2081616 0.025105154  48.1240   0e+00
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.398
```

```
## rhoC           0.009 -0.199
## tmrcaC        -0.049 -0.033 -0.345
## thetaC:tmrcaC -0.116 -0.226 -0.055  0.389
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.49301470 -0.46880229  0.07118776  0.57332632  2.12179443
##
## Residual standard error: 0.002103018
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.106155      1.196856      1.344242      1.246553
```

6.1.3 Replicate 3

```
rep3.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep3_w50000.csv", sep = ",", header = T)

rep_3.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.di
rep_3.pi.50kb$avg <- apply(rep_3.pi.50kb[4:ncol(rep_3.pi.50kb)], 1, mean)
rep_3.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.rl
rep_3.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5
rep_3.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5
rep_3.tmrca.50kb$avg <- apply(rep_3.tmrca.50kb[4:ncol(rep_3.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep3 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep3.sim.tmrca.50kb$
names(true.lands.50kb.rep3) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep3$thetaC <- true.lands.50kb.rep3$theta - mean(true.lands.50kb.rep3$theta)
true.lands.50kb.rep3$tmrcaC <- true.lands.50kb.rep3$tmrca - mean(true.lands.50kb.rep3$tmrca)
true.lands.50kb.rep3$rhoC <- true.lands.50kb.rep3$rho - mean(true.lands.50kb.rep3$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep3)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.50kb.true[2, 3] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb.true[3, 3] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 3] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 3] <- anova.diversity$VarExp[4] * 100

inf.lands.50kb.rep3 <- as.data.frame(cbind(rep_3.pi.50kb$avg, rep_3.theta.50kb$sample_mean, rep_3.tmrca
names(inf.lands.50kb.rep3) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[3, 1] <- cor.test(sim.theta.50kb$Rate, rep_3.theta.50kb$sample_mean, method = "spearman
cor.table.rho[3, 1] <- cor.test(sim.rho.50kb$Rate, rep_3.rho.50kb$sample_mean, method = "spearman")$est
cor.table.tmrca[3, 1] <- cor.test(rep3.sim.tmrca.50kb$AverageTmrca, rep_3.tmrca.50kb$avg, method = "spea

# centering
inf.lands.50kb.rep3$thetaC <- inf.lands.50kb.rep3$theta - mean(inf.lands.50kb.rep3$theta)
inf.lands.50kb.rep3$tmrcaC <- inf.lands.50kb.rep3$tmrca - mean(inf.lands.50kb.rep3$tmrca)
```

```r
inf.lands.50kb.rep3$rhoC <- inf.lands.50kb.rep3$rho - mean(inf.lands.50kb.rep3$rho)

inf.lands.50kb.rep3$bin <- 1:nrow(inf.lands.50kb.rep3)

# for merging:
inf.lands.50kb.rep3 <- inf.lands.50kb.rep3
inf.lands.50kb.rep3$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep3)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep3)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -6637.516
## m.div.50kb.2  7 -6780.842
## m.div.50kb.3  8 -6801.784
```

```r
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 0.89865, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
```

```
##
## data:  m.div.50kb
## HMC = 0.50047, p-value = 0.518
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -9.084e-04 -8.949e-05  1.918e-05  1.054e-04  8.099e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.505e-03  9.676e-06 362.237   <2e-16 ***
## thetaC        1.312e+00  6.449e-03 203.462   <2e-16 ***
## rhoC          1.097e-02  5.153e-03   2.129   0.0338 *
## tmrcaC        3.237e-03  3.487e-05  92.838   <2e-16 ***
## thetaC:tmrcaC 1.186e+00  2.390e-02  49.600   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0002091 on 466 degrees of freedom
## Multiple R-squared:  0.9932, Adjusted R-squared:  0.9931
## F-statistic: 1.694e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```r
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.50kb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 3] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep3, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep3, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep3, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep3, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6826.837
## g.div.50kb.2  8 -6854.552
## g.div.50kb.3  7 -6699.473
## g.div.50kb.4  7 -6676.519
```

```r
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep3
##         AIC       BIC   logLik
##   -6699.473 -6670.389 3356.736
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.4942882
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0035078 0.000009699 361.6599   0e+00
## thetaC        1.3054348 0.007242851 180.2377   0e+00
## rhoC          0.0189783 0.004825241   3.9331   1e-04
## tmrcaC        0.0032306 0.000033681  95.9174   0e+00
## thetaC:tmrcaC 1.1145749 0.025369938  43.9329   0e+00
##
##  Correlation:
```

```
##                (Intr) thetaC rhoC   tmrcaC
## thetaC          0.486
## rhoC            0.002 -0.177
## tmrcaC         -0.030  0.002 -0.372
## thetaC:tmrcaC -0.096 -0.152 -0.018  0.312
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -4.41121799 -0.49190945  0.06406712  0.57664950  3.48192284
##
## Residual standard error: 0.003822243
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.059040      1.212741      1.302453      1.145266
```

6.1.4 Replicate 4

```
rep4.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep4_w50000.csv", sep = ",", header = T)

rep_4.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.di
rep_4.pi.50kb$avg <- apply(rep_4.pi.50kb[4:ncol(rep_4.pi.50kb)], 1, mean)
rep_4.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.rl
rep_4.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5
rep_4.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5
rep_4.tmrca.50kb$avg <- apply(rep_4.tmrca.50kb[4:ncol(rep_4.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep4 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep4.sim.tmrca.50kb
names(true.lands.50kb.rep4) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep4$thetaC <- true.lands.50kb.rep4$theta - mean(true.lands.50kb.rep4$theta)
true.lands.50kb.rep4$tmrcaC <- true.lands.50kb.rep4$tmrca - mean(true.lands.50kb.rep4$tmrca)
true.lands.50kb.rep4$rhoC <- true.lands.50kb.rep4$rho - mean(true.lands.50kb.rep4$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep4)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.50kb.true[2, 4] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb.true[3, 4] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 4] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 4] <- anova.diversity$VarExp[4] * 100

inf.lands.50kb.rep4 <- as.data.frame(cbind(rep_4.pi.50kb$avg, rep_4.theta.50kb$sample_mean, rep_4.tmrca
names(inf.lands.50kb.rep4) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[4, 1] <- cor.test(sim.theta.50kb$Rate, rep_4.theta.50kb$sample_mean, method = "spearman
cor.table.rho[4, 1] <- cor.test(sim.rho.50kb$Rate, rep_4.rho.50kb$sample_mean, method = "spearman")$est
cor.table.tmrca[4, 1] <- cor.test(rep4.sim.tmrca.50kb$AverageTmrca, rep_4.tmrca.50kb$avg, method = "spe

# centering
```

```r
inf.lands.50kb.rep4$thetaC <- inf.lands.50kb.rep4$theta - mean(inf.lands.50kb.rep4$theta)
inf.lands.50kb.rep4$tmrcaC <- inf.lands.50kb.rep4$tmrca - mean(inf.lands.50kb.rep4$tmrca)
inf.lands.50kb.rep4$rhoC <- inf.lands.50kb.rep4$rho - mean(inf.lands.50kb.rep4$rho)

inf.lands.50kb.rep4$bin <- 1:nrow(inf.lands.50kb.rep4)

# for merging:
inf.lands.50kb.rep4 <- inf.lands.50kb.rep4
inf.lands.50kb.rep4$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep4)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep4)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -6651.111
## m.div.50kb.2  7 -6649.680
## m.div.50kb.3  8 -6651.347
```

```r
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 0.94293, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.43008, p-value = 0.02
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -9.957e-04 -7.736e-05  4.142e-05  1.167e-04  6.064e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.439e-03  9.716e-06 353.943   <2e-16 ***
## thetaC        1.066e+00  5.131e-03 207.727   <2e-16 ***
## rhoC          4.166e-02  4.872e-03   8.552   <2e-16 ***
## tmrcaC        4.753e-03  6.204e-05  76.620   <2e-16 ***
## thetaC:tmrcaC 1.514e+00  2.990e-02  50.617   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0002061 on 466 degrees of freedom
## Multiple R-squared:  0.9934, Adjusted R-squared:  0.9933
```

```
## F-statistic: 1.742e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.50kb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 4] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep4, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep4, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep4, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep4, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6872.472
## g.div.50kb.2  8 -6875.271
## g.div.50kb.3  7 -6654.881
## g.div.50kb.4  7 -6700.650
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep4
##         AIC       BIC  logLik
##   -6654.881 -6625.797 3334.44
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.1527489
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0034372 0.00000977 351.7078       0
## thetaC        1.0643723 0.00545768 195.0228       0
## rhoC          0.0438922 0.00488900   8.9777       0
## tmrcaC        0.0047503 0.00006318  75.1824       0
## thetaC:tmrcaC 1.5397661 0.03235723  47.5865       0
##
```

```
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.191
## rhoC          -0.009 -0.265
## tmrcaC        -0.080 -0.204 -0.145
## thetaC:tmrcaC -0.232 -0.173  0.053  0.356
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -5.2015408 -0.3778042  0.2017817  0.5787508  2.8135238
##
## Residual standard error: 0.0004990653
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.157157      1.133423      1.236016      1.166786
```

6.1.5 Replicate 5

```
rep5.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep5_w50000.csv", sep = ",", header = T)

rep_5.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.di
rep_5.pi.50kb$avg <- apply(rep_5.pi.50kb[4:ncol(rep_5.pi.50kb)], 1, mean)
rep_5.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.rl
rep_5.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5
rep_5.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5
rep_5.tmrca.50kb$avg <- apply(rep_5.tmrca.50kb[4:ncol(rep_5.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep5 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep5.sim.tmrca.50kb$
names(true.lands.50kb.rep5) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep5$thetaC <- true.lands.50kb.rep5$theta - mean(true.lands.50kb.rep5$theta)
true.lands.50kb.rep5$tmrcaC <- true.lands.50kb.rep5$tmrca - mean(true.lands.50kb.rep5$tmrca)
true.lands.50kb.rep5$rhoC <- true.lands.50kb.rep5$rho - mean(true.lands.50kb.rep5$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep5)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.50kb.true[2, 5] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb.true[3, 5] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 5] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 5] <- anova.diversity$VarExp[4] * 100

inf.lands.50kb.rep5 <- as.data.frame(cbind(rep_5.pi.50kb$avg, rep_5.theta.50kb$sample_mean, rep_5.tmrca
names(inf.lands.50kb.rep5) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[5, 1] <- cor.test(sim.theta.50kb$Rate, rep_5.theta.50kb$sample_mean, method = "spearman
cor.table.rho[5, 1] <- cor.test(sim.rho.50kb$Rate, rep_5.rho.50kb$sample_mean, method = "spearman")$est
cor.table.tmrca[5, 1] <- cor.test(rep5.sim.tmrca.50kb$AverageTmrca, rep_5.tmrca.50kb$avg, method = "spe
```

```r
# centering
inf.lands.50kb.rep5$thetaC <- inf.lands.50kb.rep5$theta - mean(inf.lands.50kb.rep5$theta)
inf.lands.50kb.rep5$tmrcaC <- inf.lands.50kb.rep5$tmrca - mean(inf.lands.50kb.rep5$tmrca)
inf.lands.50kb.rep5$rhoC <- inf.lands.50kb.rep5$rho - mean(inf.lands.50kb.rep5$rho)

inf.lands.50kb.rep5$bin <- 1:nrow(inf.lands.50kb.rep5)

# for merging:
inf.lands.50kb.rep5 <- inf.lands.50kb.rep5
inf.lands.50kb.rep5$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep5)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep5)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -6679.244
## m.div.50kb.2  7 -6802.716
## m.div.50kb.3  8 -6811.009
```

```r
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 0.98276, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.39935, p-value < 2.2e-16
```

```r
hist(resid(m.div.50kb))
```

```r
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep5)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -9.174e-04 -8.788e-05  7.320e-06  9.869e-05  7.862e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.452e-03  9.312e-06 370.741   <2e-16 ***
## thetaC        1.288e+00  6.186e-03 208.139   <2e-16 ***
## rhoC          8.740e-03  4.922e-03   1.776   0.0764 .
## tmrcaC        3.336e-03  3.447e-05  96.761   <2e-16 ***
## thetaC:tmrcaC 1.241e+00  2.398e-02  51.749   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.0002001 on 466 degrees of freedom
## Multiple R-squared:  0.9936, Adjusted R-squared:  0.9936
## F-statistic: 1.81e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.50kb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 5] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep5, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep5, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep5, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep5, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6839.185
## g.div.50kb.2  8 -6876.135
## g.div.50kb.3  7 -6721.049
## g.div.50kb.4  7 -6709.999
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep5
##         AIC       BIC   logLik
##   -6721.049 -6691.965 3367.525
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.4471777
##
## Coefficients:
##                   Value   Std.Error  t-value p-value
## (Intercept)   0.0034552 0.000009444 365.8683  0.0000
## thetaC        1.2814136 0.007104489 180.3667  0.0000
## rhoC          0.0136620 0.004733547   2.8862  0.0041
```

```
## tmrcaC        0.0033220 0.000035485  93.6164  0.0000
## thetaC:tmrcaC 1.1970833 0.027229278  43.9631  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.452
## rhoC           0.005 -0.222
## tmrcaC        -0.079 -0.075 -0.327
## thetaC:tmrcaC -0.159 -0.203 -0.022  0.489
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.43775897 -0.45599876  0.02430307  0.53146119  3.77882686
##
## Residual standard error: 0.002814922
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.106175      1.220884      1.519047      1.385262
```

6.1.6 Replicate 6

```
rep6.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep6_w50000.csv", sep = ",", header = T)

rep_6.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.di
rep_6.pi.50kb$avg <- apply(rep_6.pi.50kb[4:ncol(rep_6.pi.50kb)], 1, mean)
rep_6.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.rl
rep_6.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5
rep_6.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5
rep_6.tmrca.50kb$avg <- apply(rep_6.tmrca.50kb[4:ncol(rep_6.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep6 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep6.sim.tmrca.50kb
names(true.lands.50kb.rep6) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep6$thetaC <- true.lands.50kb.rep6$theta - mean(true.lands.50kb.rep6$theta)
true.lands.50kb.rep6$tmrcaC <- true.lands.50kb.rep6$tmrca - mean(true.lands.50kb.rep6$tmrca)
true.lands.50kb.rep6$rhoC <- true.lands.50kb.rep6$rho - mean(true.lands.50kb.rep6$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep6)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.50kb.true[2, 6] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb.true[3, 6] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 6] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 6] <- anova.diversity$VarExp[4] * 100

inf.lands.50kb.rep6 <- as.data.frame(cbind(rep_6.pi.50kb$avg, rep_6.theta.50kb$sample_mean, rep_6.tmrca
names(inf.lands.50kb.rep6) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[6, 1] <- cor.test(sim.theta.50kb$Rate, rep_6.theta.50kb$sample_mean, method = "spearman
```

```
cor.table.rho[6, 1] <- cor.test(sim.rho.50kb$Rate, rep_6.rho.50kb$sample_mean, method = "spearman")$est
cor.table.tmrca[6, 1] <- cor.test(rep6.sim.tmrca.50kb$AverageTmrca, rep_6.tmrca.50kb$avg, method = "spea

# centering
inf.lands.50kb.rep6$thetaC <- inf.lands.50kb.rep6$theta - mean(inf.lands.50kb.rep6$theta)
inf.lands.50kb.rep6$tmrcaC <- inf.lands.50kb.rep6$tmrca - mean(inf.lands.50kb.rep6$tmrca)
inf.lands.50kb.rep6$rhoC <- inf.lands.50kb.rep6$rho - mean(inf.lands.50kb.rep6$rho)

inf.lands.50kb.rep6$bin <- 1:nrow(inf.lands.50kb.rep6)

# for merging:
inf.lands.50kb.rep6 <- inf.lands.50kb.rep6
inf.lands.50kb.rep6$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep6)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep6)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -6743.728
## m.div.50kb.2  7 -6802.823
## m.div.50kb.3  8 -6805.058
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
```

```
## DW = 0.93612, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.4347, p-value = 0.023
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -8.002e-04 -7.326e-05  1.942e-05  9.821e-05  9.241e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.404e-03  8.832e-06 385.358  < 2e-16 ***
## thetaC        1.105e+00  5.059e-03 218.484  < 2e-16 ***
## rhoC          1.667e-02  3.878e-03   4.299 2.09e-05 ***
## tmrcaC        4.233e-03  4.493e-05  94.200  < 2e-16 ***
## thetaC:tmrcaC 1.423e+00  2.353e-02  60.461  < 2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001868 on 466 degrees of freedom
## Multiple R-squared:  0.9947, Adjusted R-squared:  0.9947
## F-statistic: 2.194e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.50kb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 6] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep6, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep6, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep6, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep6, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6946.467
## g.div.50kb.2  8 -6948.943
## g.div.50kb.3  7 -6762.528
## g.div.50kb.4  7 -6828.222
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep6
##         AIC       BIC   logLik
##   -6762.528 -6733.444 3388.264
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##    power
## 0.271011
##
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0034030 0.000008909 381.9809       0
```

```
## thetaC        1.1044051 0.005512915 200.3305       0
## rhoC          0.0164441 0.003809512   4.3166       0
## tmrcaC        0.0042366 0.000046267  91.5698       0
## thetaC:tmrcaC 1.4328627 0.026527219  54.0148       0
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.338
## rhoC           0.010 -0.168
## tmrcaC        -0.131 -0.232 -0.273
## thetaC:tmrcaC -0.257 -0.299 -0.014  0.493
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -4.1829268 -0.3945907  0.1090670  0.5609412  4.5102796
##
## Residual standard error: 0.0009014532
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.167726      1.160985      1.500100      1.400350
```

6.1.7 Replicate 7

```
rep7.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep7_w50000.csv", sep = ",", header = T)

rep_7.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.di
rep_7.pi.50kb$avg <- apply(rep_7.pi.50kb[4:ncol(rep_7.pi.50kb)], 1, mean)
rep_7.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.rl
rep_7.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5
rep_7.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5
rep_7.tmrca.50kb$avg <- apply(rep_7.tmrca.50kb[4:ncol(rep_7.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep7 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep7.sim.tmrca.50kb$
names(true.lands.50kb.rep7) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep7$thetaC <- true.lands.50kb.rep7$theta - mean(true.lands.50kb.rep7$theta)
true.lands.50kb.rep7$tmrcaC <- true.lands.50kb.rep7$tmrca - mean(true.lands.50kb.rep7$tmrca)
true.lands.50kb.rep7$rhoC <- true.lands.50kb.rep7$rho - mean(true.lands.50kb.rep7$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep7)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.50kb.true[2, 7] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb.true[3, 7] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 7] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 7] <- anova.diversity$VarExp[4] * 100

inf.lands.50kb.rep7 <- as.data.frame(cbind(rep_7.pi.50kb$avg, rep_7.theta.50kb$sample_mean, rep_7.tmrca
names(inf.lands.50kb.rep7) <- c("diversity", "theta", "tmrca", "rho")
```

```
cor.table.theta[7, 1] <- cor.test(sim.theta.50kb$Rate, rep_7.theta.50kb$sample_mean, method = "spearman
cor.table.rho[7, 1] <- cor.test(sim.rho.50kb$Rate, rep_7.rho.50kb$sample_mean, method = "spearman")$est
cor.table.tmrca[7, 1] <- cor.test(rep7.sim.tmrca.50kb$AverageTmrca, rep_7.tmrca.50kb$avg, method = "spea

# centering
inf.lands.50kb.rep7$thetaC <- inf.lands.50kb.rep7$theta - mean(inf.lands.50kb.rep7$theta)
inf.lands.50kb.rep7$tmrcaC <- inf.lands.50kb.rep7$tmrca - mean(inf.lands.50kb.rep7$tmrca)
inf.lands.50kb.rep7$rhoC <- inf.lands.50kb.rep7$rho - mean(inf.lands.50kb.rep7$rho)

inf.lands.50kb.rep7$bin <- 1:nrow(inf.lands.50kb.rep7)

# for merging:
inf.lands.50kb.rep7 <- inf.lands.50kb.rep7
inf.lands.50kb.rep7$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep7)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep7)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -6529.539
## m.div.50kb.2  7 -6563.670
## m.div.50kb.3  8 -6563.389
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
```

```
## data:  m.div.50kb
## DW = 0.80944, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.3513, p-value < 2.2e-16
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep7)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.629e-03 -6.430e-05  2.768e-05  1.120e-04  9.053e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.322e-03  1.142e-05 290.772  < 2e-16 ***
## thetaC        1.111e+00  6.786e-03 163.749  < 2e-16 ***
## rhoC          2.980e-02  5.600e-03   5.321 1.61e-07 ***
## tmrcaC        4.170e-03  6.401e-05  65.143  < 2e-16 ***
```

```
## thetaC:tmrcaC 1.512e+00  3.838e-02  39.399  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0002345 on 466 degrees of freedom
## Multiple R-squared:  0.9912, Adjusted R-squared:  0.9912
## F-statistic: 1.317e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.50kb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 7] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep7, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep7, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep7, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep7, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6771.554
## g.div.50kb.2  8 -6847.669
## g.div.50kb.3  7 -6555.954
## g.div.50kb.4  7 -6550.583
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep7
##         AIC      BIC   logLik
##   -6555.954 -6526.87 3284.977
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.4318893
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
```

```
## (Intercept)   0.0033169 0.00001198 276.85158       0
## thetaC        1.1018671 0.00780478 141.17848       0
## rhoC          0.0317114 0.00550514   5.76032       0
## tmrcaC        0.0042223 0.00007282  57.98213       0
## thetaC:tmrcaC 1.5680797 0.04561954  34.37299       0
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.500
## rhoC          -0.021 -0.250
## tmrcaC        -0.280 -0.329 -0.170
## thetaC:tmrcaC -0.372 -0.336  0.061  0.754
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -6.1376825 -0.2653249  0.1119443  0.4879379  3.3805282
##
## Residual standard error: 0.002942983
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.260102      1.239715      2.733275      2.542584
```

6.1.8 Replicate 8

```
rep8.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep8_w50000.csv", sep = ",", header = T)

rep_8.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.di
rep_8.pi.50kb$avg <- apply(rep_8.pi.50kb[4:ncol(rep_8.pi.50kb)], 1, mean)
rep_8.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.rl
rep_8.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5
rep_8.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5
rep_8.tmrca.50kb$avg <- apply(rep_8.tmrca.50kb[4:ncol(rep_8.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep8 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep8.sim.tmrca.50kb
names(true.lands.50kb.rep8) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep8$thetaC <- true.lands.50kb.rep8$theta - mean(true.lands.50kb.rep8$theta)
true.lands.50kb.rep8$tmrcaC <- true.lands.50kb.rep8$tmrca - mean(true.lands.50kb.rep8$tmrca)
true.lands.50kb.rep8$rhoC <- true.lands.50kb.rep8$rho - mean(true.lands.50kb.rep8$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep8)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.50kb.true[2, 8] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb.true[3, 8] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 8] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 8] <- anova.diversity$VarExp[4] * 100

inf.lands.50kb.rep8 <- as.data.frame(cbind(rep_8.pi.50kb$avg, rep_8.theta.50kb$sample_mean, rep_8.tmrca
```

```
names(inf.lands.50kb.rep8) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[8, 1] <- cor.test(sim.theta.50kb$Rate, rep_8.theta.50kb$sample_mean, method = "spearman
cor.table.rho[8, 1] <- cor.test(sim.rho.50kb$Rate, rep_8.rho.50kb$sample_mean, method = "spearman")$est
cor.table.tmrca[8, 1] <- cor.test(rep8.sim.tmrca.50kb$AverageTmrca, rep_8.tmrca.50kb$avg, method = "spea

# centering
inf.lands.50kb.rep8$thetaC <- inf.lands.50kb.rep8$theta - mean(inf.lands.50kb.rep8$theta)
inf.lands.50kb.rep8$tmrcaC <- inf.lands.50kb.rep8$tmrca - mean(inf.lands.50kb.rep8$tmrca)
inf.lands.50kb.rep8$rhoC <- inf.lands.50kb.rep8$rho - mean(inf.lands.50kb.rep8$rho)

inf.lands.50kb.rep8$bin <- 1:nrow(inf.lands.50kb.rep8)

# for merging:
inf.lands.50kb.rep8 <- inf.lands.50kb.rep8
inf.lands.50kb.rep8$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep8)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep8)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##               df       AIC
## m.div.50kb     6 -6664.906
## m.div.50kb.2   7 -6820.661
## m.div.50kb.3   8 -6830.462
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
```

```
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 0.81053, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.39763, p-value < 2.2e-16
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep8)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -8.849e-04 -1.035e-04  4.480e-06  1.033e-04  8.415e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.424e-03  9.459e-06 361.965   <2e-16 ***
## thetaC        1.242e+00  6.563e-03 189.277   <2e-16 ***
```

```
## rhoC          8.381e-03  5.300e-03   1.581    0.115
## tmrcaC        3.389e-03  3.343e-05 101.373   <2e-16 ***
## thetaC:tmrcaC 1.306e+00  2.597e-02  50.287   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0002031 on 466 degrees of freedom
## Multiple R-squared:  0.9934, Adjusted R-squared:  0.9934
## F-statistic: 1.766e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.50kb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 8] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep8, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep8, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep8, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep8, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6889.362
## g.div.50kb.2  8 -6915.099
## g.div.50kb.3  7 -6750.478
## g.div.50kb.4  7 -6740.369
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep8
##         AIC       BIC   logLik
##   -6750.478 -6721.394 3382.239
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.6457134
##
```

```
## Coefficients:
##                    Value   Std.Error  t-value p-value
## (Intercept)   0.0034302 0.000009656 355.2515  0.0000
## thetaC        1.2470563 0.007332157 170.0804  0.0000
## rhoC          0.0133119 0.004692246   2.8370  0.0048
## tmrcaC        0.0033566 0.000033544 100.0652  0.0000
## thetaC:tmrcaC 1.1613759 0.027531291  42.1838  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.601
## rhoC           0.001 -0.161
## tmrcaC        -0.067 -0.062 -0.360
## thetaC:tmrcaC -0.143 -0.262 -0.027  0.525
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.24874158 -0.53347725  0.04182428  0.53332121  3.79069374
##
## Residual standard error: 0.008905455
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.107570      1.226780      1.655377      1.528324
```

6.1.9 Replicate 9

```
rep9.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep9_w50000.csv", sep = ",", header = T)

rep_9.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.di
rep_9.pi.50kb$avg <- apply(rep_9.pi.50kb[4:ncol(rep_9.pi.50kb)], 1, mean)
rep_9.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.rl
rep_9.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5
rep_9.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5
rep_9.tmrca.50kb$avg <- apply(rep_9.tmrca.50kb[4:ncol(rep_9.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep9 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep9.sim.tmrca.50kb$
names(true.lands.50kb.rep9) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep9$thetaC <- true.lands.50kb.rep9$theta - mean(true.lands.50kb.rep9$theta)
true.lands.50kb.rep9$tmrcaC <- true.lands.50kb.rep9$tmrca - mean(true.lands.50kb.rep9$tmrca)
true.lands.50kb.rep9$rhoC <- true.lands.50kb.rep9$rho - mean(true.lands.50kb.rep9$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep9)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.50kb.true[2, 9] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb.true[3, 9] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 9] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 9] <- anova.diversity$VarExp[4] * 100
```

```
inf.lands.50kb.rep9 <- as.data.frame(cbind(rep_9.pi.50kb$avg, rep_9.theta.50kb$sample_mean, rep_9.tmrca
names(inf.lands.50kb.rep9) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[9, 1] <- cor.test(sim.theta.50kb$Rate, rep_9.theta.50kb$sample_mean, method = "spearman
cor.table.rho[9, 1] <- cor.test(sim.rho.50kb$Rate, rep_9.rho.50kb$sample_mean, method = "spearman")$est
cor.table.tmrca[9, 1] <- cor.test(rep9.sim.tmrca.50kb$AverageTmrca, rep_9.tmrca.50kb$avg, method = "spe

# centering
inf.lands.50kb.rep9$thetaC <- inf.lands.50kb.rep9$theta - mean(inf.lands.50kb.rep9$theta)
inf.lands.50kb.rep9$tmrcaC <- inf.lands.50kb.rep9$tmrca - mean(inf.lands.50kb.rep9$tmrca)
inf.lands.50kb.rep9$rhoC <- inf.lands.50kb.rep9$rho - mean(inf.lands.50kb.rep9$rho)

inf.lands.50kb.rep9$bin <- 1:nrow(inf.lands.50kb.rep9)

# for merging:
inf.lands.50kb.rep9 <- inf.lands.50kb.rep9
inf.lands.50kb.rep9$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep9)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.5
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep9)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##               df       AIC
## m.div.50kb     6 -6772.620
## m.div.50kb.2   7 -6788.142
## m.div.50kb.3   8 -6788.032
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```r
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.2553, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.48698, p-value = 0.353
```

```r
hist(resid(m.div.50kb))
```

```r
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep9)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.028e-03 -7.795e-05  1.040e-05  9.457e-05  8.260e-04
##
## Coefficients:
```

```
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.386e-03  8.797e-06 384.938  < 2e-16 ***
## thetaC        1.197e+00  6.055e-03 197.676  < 2e-16 ***
## rhoC          1.895e-02  4.699e-03   4.032 6.45e-05 ***
## tmrcaC        3.897e-03  4.056e-05  96.094  < 2e-16 ***
## thetaC:tmrcaC 1.553e+00  2.874e-02  54.050  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001812 on 466 degrees of freedom
## Multiple R-squared:  0.9949, Adjusted R-squared:  0.9948
## F-statistic: 2.265e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.50kb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 9] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep9, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep9, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep9, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep9, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6844.364
## g.div.50kb.2  8 -6903.791
## g.div.50kb.3  7 -6804.377
## g.div.50kb.4  7 -6772.655
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep9
##         AIC       BIC   logLik
##   -6804.377 -6775.293 3409.189
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
```

```
##     power
## 0.3817603
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0033889 0.00000899 376.7816       0
## thetaC        1.1924109 0.00675297 176.5759       0
## rhoC          0.0198466 0.00465667   4.2620       0
## tmrcaC        0.0038946 0.00004397  88.5678       0
## thetaC:tmrcaC 1.5343949 0.03313194  46.3116       0
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.467
## rhoC          -0.010 -0.194
## tmrcaC        -0.241 -0.326 -0.296
## thetaC:tmrcaC -0.361 -0.430  0.030  0.663
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -5.23766061 -0.47972109  0.04058785  0.53646388  4.07355376
##
## Residual standard error: 0.001714965
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.313079      1.298832      2.228784      2.083348
```

6.1.10 Replicate 10

```
rep10.sim.tmrca.50kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep10_w50000.csv", sep = ",", header = T)

rep_10.pi.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5
rep_10.pi.50kb$avg <- apply(rep_10.pi.50kb[4:ncol(rep_10.pi.50kb)], 1, mean)
rep_10.rho.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5
rep_10.theta.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep1
rep_10.tmrca.50kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep1
rep_10.tmrca.50kb$avg <- apply(rep_10.tmrca.50kb[4:ncol(rep_10.tmrca.50kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.50kb.rep10 <- as.data.frame(cbind(rep_1.pi.50kb$avg, sim.theta.50kb$Rate, rep10.sim.tmrca.50
names(true.lands.50kb.rep10) <- c("diversity", "theta", "tmrca", "rho")
true.lands.50kb.rep10$thetaC <- true.lands.50kb.rep10$theta - mean(true.lands.50kb.rep10$theta)
true.lands.50kb.rep10$tmrcaC <- true.lands.50kb.rep10$tmrca - mean(true.lands.50kb.rep10$tmrca)
true.lands.50kb.rep10$rhoC <- true.lands.50kb.rep10$rho - mean(true.lands.50kb.rep10$rho, na.rm = T)

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.50kb.rep10)
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.50kb.true[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Var
r2.bgs.50kb.true[2, 10] <- anova.diversity$VarExp[1] * 100
```

```
r2.bgs.50kb.true[3, 10] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb.true[4, 10] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb.true[5, 10] <- anova.diversity$VarExp[4] * 100

inf.lands.50kb.rep10 <- as.data.frame(cbind(rep_10.pi.50kb$avg, rep_10.theta.50kb$sample_mean, rep_10.tr
names(inf.lands.50kb.rep10) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[10, 1] <- cor.test(sim.theta.50kb$Rate, rep_10.theta.50kb$sample_mean, method = "spearma
cor.table.rho[10, 1] <- cor.test(sim.rho.50kb$Rate, rep_10.rho.50kb$sample_mean, method = "spearman")$ea
cor.table.tmrca[10, 1] <- cor.test(rep10.sim.tmrca.50kb$AverageTmrca, rep_10.tmrca.50kb$avg, method = "s

# centering
inf.lands.50kb.rep10$thetaC <- inf.lands.50kb.rep10$theta - mean(inf.lands.50kb.rep10$theta)
inf.lands.50kb.rep10$tmrcaC <- inf.lands.50kb.rep10$tmrca - mean(inf.lands.50kb.rep10$tmrca)
inf.lands.50kb.rep10$rhoC <- inf.lands.50kb.rep10$rho - mean(inf.lands.50kb.rep10$rho)

inf.lands.50kb.rep10$bin <- 1:nrow(inf.lands.50kb.rep10)

# for merging:
inf.lands.50kb.rep10 <- inf.lands.50kb.rep10
inf.lands.50kb.rep10$Replicate <- 1

m.div.50kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.50kb.rep10)
m.div.50kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.50
m.div.50kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.50kb.rep10)

AIC(m.div.50kb, m.div.50kb.2, m.div.50kb.3)
```

```
##              df       AIC
## m.div.50kb    6 -6830.551
## m.div.50kb.2  7 -6838.420
## m.div.50kb.3  8 -6849.626
```

```
plot(resid(m.div.50kb)~fitted(m.div.50kb))
```

```
dwtest(m.div.50kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.50kb
## DW = 1.1672, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.50kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.50kb
## HMC = 0.47978, p-value = 0.28
```

```
hist(resid(m.div.50kb))
```

```
summary(m.div.50kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.50kb.rep10)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -6.172e-04 -6.760e-05  1.266e-05  9.275e-05  8.769e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.443e-03  8.020e-06 429.318  < 2e-16 ***
## thetaC        1.068e+00  4.356e-03 245.081  < 2e-16 ***
## rhoC          3.865e-02  5.193e-03   7.442 4.84e-13 ***
## tmrcaC        4.432e-03  4.816e-05  92.015  < 2e-16 ***
## thetaC:tmrcaC 1.413e+00  2.399e-02  58.877  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001704 on 466 degrees of freedom
## Multiple R-squared:  0.9955, Adjusted R-squared:  0.9955
## F-statistic: 2.585e+04 on 4 and 466 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.50kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.50kb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```
r2.bgs.50kb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.bgs.50kb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.bgs.50kb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.bgs.50kb[5, 10] <- anova.diversity$VarExp[4] * 100

g.div.50kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.50kb.rep10, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.50kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep10, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.50kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep10, weights = varPower(0, ~theta), method = "ML")

g.div.50kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                    data = inf.lands.50kb.rep10, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.50kb.1, g.div.50kb.2, g.div.50kb.3, g.div.50kb.4)
```

```
##              df       AIC
## g.div.50kb.1  8 -6942.468
## g.div.50kb.2  8 -6997.019
## g.div.50kb.3  7 -6868.321
## g.div.50kb.4  7 -6861.577
```

```
summary(g.div.50kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.50kb.rep10
##         AIC       BIC   logLik
##   -6868.321 -6839.237 3441.161
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.3740056
##
## Coefficients:
##                   Value   Std.Error  t-value p-value
## (Intercept)   0.0034424 0.000008084 425.8315       0
## thetaC        1.0686825 0.004873670 219.2768       0
## rhoC          0.0443179 0.004804829   9.2236       0
## tmrcaC        0.0044039 0.000048558  90.6934       0
## thetaC:tmrcaC 1.4173157 0.026661837  53.1590       0
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.393
## rhoC           0.009 -0.275
## tmrcaC        -0.107 -0.100 -0.390
## thetaC:tmrcaC -0.233 -0.151 -0.021  0.463
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.62765088 -0.45676308  0.09593681  0.61915114  3.79738407
##
## Residual standard error: 0.001457433
## Degrees of freedom: 471 total; 466 residual
```

```
vif(g.div.50kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.147519      1.373197      1.611565      1.328148
```

6.2 all replicates:

```
theta.plot <- as.data.frame(cbind(1:nrow(sim.theta.50kb),
                                  sim.theta.50kb$Rate * 1e-7 * 3e+4,
                                  rep_1.theta.50kb$sample_mean,
                                  rep_2.theta.50kb$sample_mean,
                                  rep_3.theta.50kb$sample_mean,
                                  rep_4.theta.50kb$sample_mean,
                                  rep_5.theta.50kb$sample_mean,
                                  rep_6.theta.50kb$sample_mean,
                                  rep_7.theta.50kb$sample_mean,
                                  rep_8.theta.50kb$sample_mean,
                                  rep_9.theta.50kb$sample_mean,
                                  rep_10.theta.50kb$sample_mean))

names(theta.plot) <- c("bin", "sim", reps)
molten.theta <- melt(theta.plot, id.vars = "bin")
theta.map.50kb.bgs <- ggplot(data = molten.theta, aes(x = bin * 50, y = value, colour = variable)) + the
theta.map.50kb.bgs <- theta.map.50kb.bgs + geom_line(data = molten.theta, aes(size = variable)) + scale_
theta.map.50kb.bgs <- theta.map.50kb.bgs + scale_color_manual(values = c("black", brewer.pal(n = 9, name
theta.map.50kb.bgs <- theta.map.50kb.bgs + scale_x_continuous(breaks = pretty_breaks()) + scale_y_contin
theta.map.50kb.bgs <- theta.map.50kb.bgs + labs(title = NULL, x = NULL, y = NULL)
theta.map.50kb.bgs <- theta.map.50kb.bgs + theme(text = element_text(size = 20), axis.title.x = element_
theta.map.50kb.bgs
```

6.3 200 kb scale

```
r2.bgs.200kb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.bgs.200kb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.bgs.200kb) <- reps

r2.bgs.200kb.true <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.bgs.200kb.true) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.bgs.200kb.true) <- reps

sim.theta.200kb <- read.table("dm_bgs_sims/MutationMap_200kb.csv", sep = ",", header = T)
sim.rho.200kb <- read.table("dm_bgs_sims/RecombinationMap_200kb.csv", sep = ",", header = T)
```

6.3.1 Replicate 1

```
rep1.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep1_w200000.csv", sep = ",", header = T)

rep_1.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.di
rep_1.pi.200kb$avg <- apply(rep_1.pi.200kb[4:ncol(rep_1.pi.200kb)], 1, mean)
rep_1.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.r
rep_1.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5
rep_1.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5
rep_1.tmrca.200kb$avg <- apply(rep_1.tmrca.200kb[4:ncol(rep_1.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep1 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep1.sim.tmrca.20
```

```
names(true.lands.200kb.rep1) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep1$thetaC <- true.lands.200kb.rep1$theta - mean(true.lands.200kb.rep1$theta)
true.lands.200kb.rep1$tmrcaC <- true.lands.200kb.rep1$tmrca - mean(true.lands.200kb.rep1$tmrca)
true.lands.200kb.rep1$rhoC <- true.lands.200kb.rep1$rho - mean(true.lands.200kb.rep1$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep1)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Var
r2.bgs.200kb.true[2, 1] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb.true[3, 1] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 1] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 1] <- anova.diversity$VarExp[4] * 100


inf.lands.200kb.rep1 <- as.data.frame(cbind(rep_1.pi.200kb$avg, rep_1.theta.200kb$sample_mean, rep_1.tm
names(inf.lands.200kb.rep1) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[1, 2] <- cor.test(sim.theta.200kb$Rate, rep_1.theta.200kb$sample_mean, method = "spearma
cor.table.rho[1, 2] <- cor.test(sim.rho.200kb$Rate, rep_1.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[1, 2] <- cor.test(rep1.sim.tmrca.200kb$AverageTmrca, rep_1.tmrca.200kb$avg, method = "sp

# centering
inf.lands.200kb.rep1$thetaC <- inf.lands.200kb.rep1$theta - mean(inf.lands.200kb.rep1$theta)
inf.lands.200kb.rep1$tmrcaC <- inf.lands.200kb.rep1$tmrca - mean(inf.lands.200kb.rep1$tmrca)
inf.lands.200kb.rep1$rhoC <- inf.lands.200kb.rep1$rho - mean(inf.lands.200kb.rep1$rho)

inf.lands.200kb.rep1$bin <- 1:nrow(inf.lands.200kb.rep1)

# for merging:
inf.lands.200kb.rep1 <- inf.lands.200kb.rep1
inf.lands.200kb.rep1$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep1)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep1)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1724.393
## m.div.200kb.2  7 -1740.570
## m.div.200kb.3  8 -1739.520
```

```
plot(resid(m.div.200kb.2)~fitted(m.div.200kb.2))
```

```
dwtest(m.div.200kb.2)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb.2
## DW = 1.167, p-value = 1.002e-06
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb.2)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb.2
## HMC = 0.38434, p-value = 0.024
```

```
hist(resid(m.div.200kb.2))
```

```r
summary(m.div.200kb.2)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC +
##     rhoC:tmrcaC, data = inf.lands.200kb.rep1)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -5.278e-04 -5.856e-05  9.240e-06  7.099e-05  3.337e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.415e-03  1.769e-05 193.085  < 2e-16 ***
## thetaC        1.094e+00  1.019e-02 107.387  < 2e-16 ***
## rhoC          3.098e-02  1.007e-02   3.077  0.00262 **
## tmrcaC        4.599e-03  1.787e-04  25.740  < 2e-16 ***
## thetaC:tmrcaC 1.336e+00  6.156e-02  21.699  < 2e-16 ***
## rhoC:tmrcaC   3.240e-01  7.502e-02   4.319 3.41e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001466 on 112 degrees of freedom
## Multiple R-squared:  0.9957, Adjusted R-squared:  0.9955
## F-statistic:  5132 on 5 and 112 DF,  p-value: < 2.2e-16
```

```r
anova.diversity <- Anova(m.div.200kb.2)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.bgs.200kb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.200kb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 1] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep1, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep1, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep1, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep1, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1766.542
## g.div.200kb.2  8 -1764.701
## g.div.200kb.3  7 -1743.945
## g.div.200kb.4  7 -1752.612
```

```
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep1
##         AIC       BIC   logLik
##   -1743.945 -1724.551 878.9727
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.7361752
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0034707 0.00001499 231.50685       0
## thetaC        1.1083561 0.01144751  96.82072       0
## rhoC          0.0495004 0.00837259   5.91220       0
## tmrcaC        0.0038521 0.00007797  49.40405       0
## thetaC:tmrcaC 1.3100299 0.05918971  22.13273       0
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.563
## rhoC           0.004 -0.373
## tmrcaC        -0.076  0.027 -0.377
```

```
## thetaC:tmrcaC -0.149 -0.092 -0.098  0.550
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.34897044 -0.50845520 -0.03176751  0.64399094  2.60457466
##
## Residual standard error: 0.01046373
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.190126      1.391290      1.695926      1.471880
```

6.3.2 Replicate 2

```
rep2.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep2_w200000.csv", sep = ",", header = T)

rep_2.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.di
rep_2.pi.200kb$avg <- apply(rep_2.pi.200kb[4:ncol(rep_2.pi.200kb)], 1, mean)
rep_2.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.r
rep_2.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5
rep_2.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5
rep_2.tmrca.200kb$avg <- apply(rep_2.tmrca.200kb[4:ncol(rep_2.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep2 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep2.sim.tmrca.20
names(true.lands.200kb.rep2) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep2$thetaC <- true.lands.200kb.rep2$theta - mean(true.lands.200kb.rep2$theta)
true.lands.200kb.rep2$tmrcaC <- true.lands.200kb.rep2$tmrca - mean(true.lands.200kb.rep2$tmrca)
true.lands.200kb.rep2$rhoC <- true.lands.200kb.rep2$rho - mean(true.lands.200kb.rep2$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep2)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Var
r2.bgs.200kb.true[2, 2] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb.true[3, 2] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 2] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 2] <- anova.diversity$VarExp[4] * 100

inf.lands.200kb.rep2 <- as.data.frame(cbind(rep_2.pi.200kb$avg, rep_2.theta.200kb$sample_mean, rep_2.tm
names(inf.lands.200kb.rep2) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[2, 2] <- cor.test(sim.theta.200kb$Rate, rep_2.theta.200kb$sample_mean, method = "spearma
cor.table.rho[2, 2] <- cor.test(sim.rho.200kb$Rate, rep_2.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[2, 2] <- cor.test(rep2.sim.tmrca.200kb$AverageTmrca, rep_2.tmrca.200kb$avg, method = "sp

# centering
inf.lands.200kb.rep2$thetaC <- inf.lands.200kb.rep2$theta - mean(inf.lands.200kb.rep2$theta)
inf.lands.200kb.rep2$tmrcaC <- inf.lands.200kb.rep2$tmrca - mean(inf.lands.200kb.rep2$tmrca)
inf.lands.200kb.rep2$rhoC <- inf.lands.200kb.rep2$rho - mean(inf.lands.200kb.rep2$rho)
```

```r
inf.lands.200kb.rep2$bin <- 1:nrow(inf.lands.200kb.rep2)

# for merging:
inf.lands.200kb.rep2 <- inf.lands.200kb.rep2
inf.lands.200kb.rep2$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep2)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep2)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1700.943
## m.div.200kb.2  7 -1724.292
## m.div.200kb.3  8 -1728.465
```

```r
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.4796, p-value = 0.001381
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
```

```
## HMC = 0.31601, p-value = 0.001
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -8.069e-04 -8.170e-05  1.681e-05  8.234e-05  4.681e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.490e-03  1.608e-05 217.015   <2e-16 ***
## thetaC        1.270e+00  1.302e-02  97.564   <2e-16 ***
## rhoC          1.464e-02  9.533e-03   1.536    0.127
## tmrcaC        3.319e-03  6.107e-05  54.351   <2e-16 ***
## thetaC:tmrcaC 1.234e+00  4.611e-02  26.756   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001742 on 113 degrees of freedom
## Multiple R-squared:  0.9939, Adjusted R-squared:  0.9937
## F-statistic:  4640 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
```

```r
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.200kb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 2] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep2, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep2, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep2, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep2, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1726.678
## g.div.200kb.2  8 -1710.004
## g.div.200kb.3  7 -1704.923
## g.div.200kb.4  7 -1719.743
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep2
##         AIC       BIC   logLik
##   -1704.923 -1685.528 859.4613
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.3678003
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0034912 0.00001612 216.52518  0.0000
## thetaC        1.2689894 0.01421513  89.27034  0.0000
## rhoC          0.0205454 0.00923486   2.22477  0.0281
## tmrcaC        0.0033095 0.00005994  55.21633  0.0000
## thetaC:tmrcaC 1.1986237 0.05096523  23.51846  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.314
```

```
## rhoC           0.006 -0.302
## tmrcaC        -0.020  0.079 -0.408
## thetaC:tmrcaC -0.087 -0.224 -0.018  0.173
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -4.90801162 -0.54193854  0.09032308  0.52513685  2.21280365
##
## Residual standard error: 0.0014917
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.168684      1.316209      1.240963      1.095538
```

6.3.3 Replicate 3

```
rep3.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep3_w200000.csv", sep = ",", header = T)

rep_3.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.di
rep_3.pi.200kb$avg <- apply(rep_3.pi.200kb[4:ncol(rep_3.pi.200kb)], 1, mean)
rep_3.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.r
rep_3.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5
rep_3.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5
rep_3.tmrca.200kb$avg <- apply(rep_3.tmrca.200kb[4:ncol(rep_3.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep3 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep3.sim.tmrca.20
names(true.lands.200kb.rep3) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep3$thetaC <- true.lands.200kb.rep3$theta - mean(true.lands.200kb.rep3$theta)
true.lands.200kb.rep3$tmrcaC <- true.lands.200kb.rep3$tmrca - mean(true.lands.200kb.rep3$tmrca)
true.lands.200kb.rep3$rhoC <- true.lands.200kb.rep3$rho - mean(true.lands.200kb.rep3$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep3)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Var
r2.bgs.200kb.true[2, 3] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb.true[3, 3] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 3] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 3] <- anova.diversity$VarExp[4] * 100

inf.lands.200kb.rep3 <- as.data.frame(cbind(rep_3.pi.200kb$avg, rep_3.theta.200kb$sample_mean, rep_3.tm
names(inf.lands.200kb.rep3) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[3, 2] <- cor.test(sim.theta.200kb$Rate, rep_3.theta.200kb$sample_mean, method = "spearma
cor.table.rho[3, 2] <- cor.test(sim.rho.200kb$Rate, rep_3.rho.200kb$sample_mean, method = "spearman")$e
cor.table.tmrca[3, 2] <- cor.test(rep3.sim.tmrca.200kb$AverageTmrca, rep_3.tmrca.200kb$avg, method = "s

# centering
inf.lands.200kb.rep3$thetaC <- inf.lands.200kb.rep3$theta - mean(inf.lands.200kb.rep3$theta)
inf.lands.200kb.rep3$tmrcaC <- inf.lands.200kb.rep3$tmrca - mean(inf.lands.200kb.rep3$tmrca)
```

```r
inf.lands.200kb.rep3$rhoC <- inf.lands.200kb.rep3$rho - mean(inf.lands.200kb.rep3$rho)

inf.lands.200kb.rep3$bin <- 1:nrow(inf.lands.200kb.rep3)

# for merging:
inf.lands.200kb.rep3 <- inf.lands.200kb.rep3
inf.lands.200kb.rep3$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep3)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep3)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1707.314
## m.div.200kb.2  7 -1754.542
## m.div.200kb.3  8 -1764.055
```

```r
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.3035, p-value = 3.818e-05
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
```

```
##
## data:  m.div.200kb
## HMC = 0.44743, p-value = 0.198
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -5.233e-04 -6.723e-05  2.374e-05  9.394e-05  6.412e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.519e-03  1.563e-05 225.098   <2e-16 ***
## thetaC        1.304e+00  1.266e-02 103.074   <2e-16 ***
## rhoC          1.509e-02  9.182e-03   1.643    0.103
## tmrcaC        3.196e-03  6.524e-05  48.993   <2e-16 ***
## thetaC:tmrcaC 1.178e+00  4.759e-02  24.752   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001695 on 113 degrees of freedom
## Multiple R-squared:  0.9941, Adjusted R-squared:  0.9939
## F-statistic:  4792 on 4 and 113 DF,  p-value: < 2.2e-16
```

```r
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.200kb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 3] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep3, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep3, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep3, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep3, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1726.196
## g.div.200kb.2  8 -1733.453
## g.div.200kb.3  7 -1719.642
## g.div.200kb.4  7 -1713.673
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep3
##         AIC       BIC   logLik
##   -1719.642 -1700.247 866.8209
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.5394141
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0035215 0.00001564 225.20991  0.0000
## thetaC        1.3024097 0.01384062  94.10053  0.0000
## rhoC          0.0250918 0.00888843   2.82298  0.0056
## tmrcaC        0.0032052 0.00006146  52.14774  0.0000
## thetaC:tmrcaC 1.0588155 0.05105686  20.73797  0.0000
##
##  Correlation:
```

```
##                (Intr) thetaC rhoC   tmrcaC
## thetaC          0.444
## rhoC           -0.003 -0.311
## tmrcaC         -0.004  0.137 -0.447
## thetaC:tmrcaC -0.064 -0.175  0.023  0.090
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -4.0617345 -0.4658902  0.1165599  0.5723830  3.1881386
##
## Residual standard error: 0.003946494
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.142896      1.358502      1.266648      1.046199
```

6.3.4 Replicate 4

```
rep4.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep4_w200000.csv", sep = ",", header = T)

rep_4.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.di
rep_4.pi.200kb$avg <- apply(rep_4.pi.200kb[4:ncol(rep_4.pi.200kb)], 1, mean)
rep_4.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.r
rep_4.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5
rep_4.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5
rep_4.tmrca.200kb$avg <- apply(rep_4.tmrca.200kb[4:ncol(rep_4.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep4 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep4.sim.tmrca.20
names(true.lands.200kb.rep4) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep4$thetaC <- true.lands.200kb.rep4$theta - mean(true.lands.200kb.rep4$theta)
true.lands.200kb.rep4$tmrcaC <- true.lands.200kb.rep4$tmrca - mean(true.lands.200kb.rep4$tmrca)
true.lands.200kb.rep4$rhoC <- true.lands.200kb.rep4$rho - mean(true.lands.200kb.rep4$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep4)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Var
r2.bgs.200kb.true[2, 4] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb.true[3, 4] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 4] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 4] <- anova.diversity$VarExp[4] * 100

inf.lands.200kb.rep4 <- as.data.frame(cbind(rep_4.pi.200kb$avg, rep_4.theta.200kb$sample_mean, rep_4.tm
names(inf.lands.200kb.rep4) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[4, 2] <- cor.test(sim.theta.200kb$Rate, rep_4.theta.200kb$sample_mean, method = "spearma
cor.table.rho[4, 2] <- cor.test(sim.rho.200kb$Rate, rep_4.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[4, 2] <- cor.test(rep4.sim.tmrca.200kb$AverageTmrca, rep_4.tmrca.200kb$avg, method = "sp

# centering
```

```r
inf.lands.200kb.rep4$thetaC <- inf.lands.200kb.rep4$theta - mean(inf.lands.200kb.rep4$theta)
inf.lands.200kb.rep4$tmrcaC <- inf.lands.200kb.rep4$tmrca - mean(inf.lands.200kb.rep4$tmrca)
inf.lands.200kb.rep4$rhoC <- inf.lands.200kb.rep4$rho - mean(inf.lands.200kb.rep4$rho)

inf.lands.200kb.rep4$bin <- 1:nrow(inf.lands.200kb.rep4)

# for merging:
inf.lands.200kb.rep4 <- inf.lands.200kb.rep4
inf.lands.200kb.rep4$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep4)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep4)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1713.378
## m.div.200kb.2  7 -1712.955
## m.div.200kb.3  8 -1711.494
```

```r
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.3999, p-value = 0.0002898
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.54457, p-value = 0.753
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -9.505e-04 -7.323e-05  1.451e-05  9.198e-05  3.959e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0034540  0.0000156 221.409  < 2e-16 ***
## thetaC        1.0510911  0.0098390 106.829  < 2e-16 ***
## rhoC          0.0457967  0.0086872   5.272 6.56e-07 ***
## tmrcaC        0.0050692  0.0001243  40.793  < 2e-16 ***
## thetaC:tmrcaC 1.4832766  0.0629216  23.573  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001652 on 113 degrees of freedom
## Multiple R-squared:  0.9944, Adjusted R-squared:  0.9942
```

```
## F-statistic:  5034 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.200kb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 4] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep4, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep4, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep4, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep4, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1734.204
## g.div.200kb.2  8 -1714.992
## g.div.200kb.3  7 -1711.444
## g.div.200kb.4  7 -1731.676
```

```
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep4
##         AIC      BIC   logLik
##   -1711.444 -1692.05 862.7222
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##      power
## 0.03207734
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0034535 0.00001562 221.09509       0
## thetaC        1.0511893 0.00995135 105.63281       0
## rhoC          0.0461182 0.00870646   5.29701       0
## tmrcaC        0.0050657 0.00012452  40.68200       0
## thetaC:tmrcaC 1.4907778 0.06401702  23.28721       0
##
```

```
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.065
## rhoC          -0.019 -0.338
## tmrcaC        -0.038 -0.172 -0.220
## thetaC:tmrcaC -0.228 -0.170  0.085  0.171
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -5.95926937 -0.46066423  0.09235141  0.57624590  2.39702801
##
## Residual standard error: 0.0001948346
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.229295      1.247833      1.162730      1.058662
```

6.3.5 Replicate 5

```
rep5.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep5_w200000.csv", sep = ",", header = T)

rep_5.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.di
rep_5.pi.200kb$avg <- apply(rep_5.pi.200kb[4:ncol(rep_5.pi.200kb)], 1, mean)
rep_5.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.r
rep_5.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5
rep_5.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5
rep_5.tmrca.200kb$avg <- apply(rep_5.tmrca.200kb[4:ncol(rep_5.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep5 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep5.sim.tmrca.20
names(true.lands.200kb.rep5) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep5$thetaC <- true.lands.200kb.rep5$theta - mean(true.lands.200kb.rep5$theta)
true.lands.200kb.rep5$tmrcaC <- true.lands.200kb.rep5$tmrca - mean(true.lands.200kb.rep5$tmrca)
true.lands.200kb.rep5$rhoC <- true.lands.200kb.rep5$rho - mean(true.lands.200kb.rep5$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep5)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Var
r2.bgs.200kb.true[2, 5] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb.true[3, 5] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 5] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 5] <- anova.diversity$VarExp[4] * 100

inf.lands.200kb.rep5 <- as.data.frame(cbind(rep_5.pi.200kb$avg, rep_5.theta.200kb$sample_mean, rep_5.tm
names(inf.lands.200kb.rep5) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[5, 2] <- cor.test(sim.theta.200kb$Rate, rep_5.theta.200kb$sample_mean, method = "spearma
cor.table.rho[5, 2] <- cor.test(sim.rho.200kb$Rate, rep_5.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[5, 2] <- cor.test(rep5.sim.tmrca.200kb$AverageTmrca, rep_5.tmrca.200kb$avg, method = "sp
```

```r
# centering
inf.lands.200kb.rep5$thetaC <- inf.lands.200kb.rep5$theta - mean(inf.lands.200kb.rep5$theta)
inf.lands.200kb.rep5$tmrcaC <- inf.lands.200kb.rep5$tmrca - mean(inf.lands.200kb.rep5$tmrca)
inf.lands.200kb.rep5$rhoC <- inf.lands.200kb.rep5$rho - mean(inf.lands.200kb.rep5$rho)

inf.lands.200kb.rep5$bin <- 1:nrow(inf.lands.200kb.rep5)

# for merging:
inf.lands.200kb.rep5 <- inf.lands.200kb.rep5
inf.lands.200kb.rep5$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep5)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep5)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1719.223
## m.div.200kb.2  7 -1752.807
## m.div.200kb.3  8 -1753.702
```

```r
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.3237, p-value = 6.006e-05
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.38497, p-value = 0.036
```

```r
hist(resid(m.div.200kb))
```

```r
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep5)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -5.064e-04 -7.206e-05  1.342e-05  6.558e-05  5.195e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.461e-03  1.497e-05 231.248   <2e-16 ***
## thetaC        1.289e+00  1.202e-02 107.250   <2e-16 ***
## rhoC          5.476e-03  8.877e-03   0.617    0.539
## tmrcaC        3.298e-03  6.213e-05  53.084   <2e-16 ***
## thetaC:tmrcaC 1.196e+00  4.528e-02  26.410   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.0001612 on 113 degrees of freedom
## Multiple R-squared:  0.9947, Adjusted R-squared:  0.9945
## F-statistic:  5272 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.200kb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 5] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep5, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep5, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep5, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep5, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1755.478
## g.div.200kb.2  8 -1740.007
## g.div.200kb.3  7 -1725.582
## g.div.200kb.4  7 -1749.740
```

```
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep5
##         AIC       BIC   logLik
##   -1725.582 -1706.187 869.7909
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.4273153
##
## Coefficients:
##                   Value  Std.Error   t-value p-value
## (Intercept)   0.0034638 0.00001505 230.12028  0.0000
## thetaC        1.2857552 0.01337872  96.10447  0.0000
## rhoC          0.0119472 0.00874229   1.36660  0.1745
```

```
## tmrcaC        0.0032822 0.00006202  52.92532  0.0000
## thetaC:tmrcaC 1.1470417 0.05046751  22.72832  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.368
## rhoC           0.004 -0.344
## tmrcaC        -0.053  0.016 -0.393
## thetaC:tmrcaC -0.148 -0.190 -0.011  0.342
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.75640462 -0.47240720  0.07496887  0.41643045  2.69009455
##
## Residual standard error: 0.001969146
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.190932      1.375450      1.374444      1.190572
```

6.3.6 Replicate 6

```
rep6.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep6_w200000.csv", sep = ",", header = T)

rep_6.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.di
rep_6.pi.200kb$avg <- apply(rep_6.pi.200kb[4:ncol(rep_6.pi.200kb)], 1, mean)
rep_6.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.r
rep_6.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5
rep_6.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5
rep_6.tmrca.200kb$avg <- apply(rep_6.tmrca.200kb[4:ncol(rep_6.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep6 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep6.sim.tmrca.20
names(true.lands.200kb.rep6) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep6$thetaC <- true.lands.200kb.rep6$theta - mean(true.lands.200kb.rep6$theta)
true.lands.200kb.rep6$tmrcaC <- true.lands.200kb.rep6$tmrca - mean(true.lands.200kb.rep6$tmrca)
true.lands.200kb.rep6$rhoC <- true.lands.200kb.rep6$rho - mean(true.lands.200kb.rep6$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep6)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Var
r2.bgs.200kb.true[2, 6] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb.true[3, 6] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 6] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 6] <- anova.diversity$VarExp[4] * 100

inf.lands.200kb.rep6 <- as.data.frame(cbind(rep_6.pi.200kb$avg, rep_6.theta.200kb$sample_mean, rep_6.tm
names(inf.lands.200kb.rep6) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[6, 2] <- cor.test(sim.theta.200kb$Rate, rep_6.theta.200kb$sample_mean, method = "spearma
```

```
cor.table.rho[6, 2] <- cor.test(sim.rho.200kb$Rate, rep_6.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[6, 2] <- cor.test(rep6.sim.tmrca.200kb$AverageTmrca, rep_6.tmrca.200kb$avg, method = "sp

# centering
inf.lands.200kb.rep6$thetaC <- inf.lands.200kb.rep6$theta - mean(inf.lands.200kb.rep6$theta)
inf.lands.200kb.rep6$tmrcaC <- inf.lands.200kb.rep6$tmrca - mean(inf.lands.200kb.rep6$tmrca)
inf.lands.200kb.rep6$rhoC <- inf.lands.200kb.rep6$rho - mean(inf.lands.200kb.rep6$rho)

inf.lands.200kb.rep6$bin <- 1:nrow(inf.lands.200kb.rep6)

# for merging:
inf.lands.200kb.rep6 <- inf.lands.200kb.rep6
inf.lands.200kb.rep6$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep6)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep6)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1721.284
## m.div.200kb.2  7 -1738.537
## m.div.200kb.3  8 -1737.848
```

```
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
```

```
## data:  m.div.200kb
## DW = 1.1428, p-value = 6.441e-07
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.37066, p-value = 0.023
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -5.169e-04 -6.868e-05  1.496e-05  7.251e-05  3.947e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.420e-03  1.512e-05  226.13   <2e-16 ***
## thetaC        1.101e+00  1.019e-02  107.98   <2e-16 ***
## rhoC          1.707e-02  7.262e-03    2.35   0.0205 *
## tmrcaC        4.264e-03  8.815e-05   48.37   <2e-16 ***
```

```
## thetaC:tmrcaC 1.417e+00  5.145e-02   27.55   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001598 on 113 degrees of freedom
## Multiple R-squared:  0.995,  Adjusted R-squared:  0.9949
## F-statistic:  5662 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.200kb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 6] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep6, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep6, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep6, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep6, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1750.916
## g.div.200kb.2  8 -1746.207
## g.div.200kb.3  7 -1722.532
## g.div.200kb.4  7 -1738.724
```

```
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep6
##         AIC       BIC  logLik
##   -1722.532 -1703.137 868.266
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.2643174
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
```

```
## (Intercept)   0.0034189 0.00001529 223.64371  0.0000
## thetaC        1.1060883 0.01089476 101.52477  0.0000
## rhoC          0.0170656 0.00717873   2.37724  0.0191
## tmrcaC        0.0042514 0.00008988  47.30080  0.0000
## thetaC:tmrcaC 1.4159442 0.05773371  24.52543  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.297
## rhoC           0.011 -0.220
## tmrcaC        -0.115 -0.201 -0.326
## thetaC:tmrcaC -0.265 -0.294 -0.019  0.420
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -3.8683483 -0.4416453  0.1205680  0.5114539  2.3563921
##
## Residual standard error: 0.0007276808
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.202416      1.242707      1.436129      1.292643
```

6.3.7 Replicate 7

```
rep7.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep7_w200000.csv", sep = ",", header = T)

rep_7.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.di
rep_7.pi.200kb$avg <- apply(rep_7.pi.200kb[4:ncol(rep_7.pi.200kb)], 1, mean)
rep_7.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.r
rep_7.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5
rep_7.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5
rep_7.tmrca.200kb$avg <- apply(rep_7.tmrca.200kb[4:ncol(rep_7.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep7 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep7.sim.tmrca.20
names(true.lands.200kb.rep7) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep7$thetaC <- true.lands.200kb.rep7$theta - mean(true.lands.200kb.rep7$theta)
true.lands.200kb.rep7$tmrcaC <- true.lands.200kb.rep7$tmrca - mean(true.lands.200kb.rep7$tmrca)
true.lands.200kb.rep7$rhoC <- true.lands.200kb.rep7$rho - mean(true.lands.200kb.rep7$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep7)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Var
r2.bgs.200kb.true[2, 7] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb.true[3, 7] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 7] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 7] <- anova.diversity$VarExp[4] * 100

inf.lands.200kb.rep7 <- as.data.frame(cbind(rep_7.pi.200kb$avg, rep_7.theta.200kb$sample_mean, rep_7.tm
```

```r
names(inf.lands.200kb.rep7) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[7, 2] <- cor.test(sim.theta.200kb$Rate, rep_7.theta.200kb$sample_mean, method = "spearma
cor.table.rho[7, 2] <- cor.test(sim.rho.200kb$Rate, rep_1.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[7, 2] <- cor.test(rep7.sim.tmrca.200kb$AverageTmrca, rep_7.tmrca.200kb$avg, method = "sp

# centering
inf.lands.200kb.rep7$thetaC <- inf.lands.200kb.rep7$theta - mean(inf.lands.200kb.rep7$theta)
inf.lands.200kb.rep7$tmrcaC <- inf.lands.200kb.rep7$tmrca - mean(inf.lands.200kb.rep7$tmrca)
inf.lands.200kb.rep7$rhoC <- inf.lands.200kb.rep7$rho - mean(inf.lands.200kb.rep7$rho)

inf.lands.200kb.rep7$bin <- 1:nrow(inf.lands.200kb.rep7)

# for merging:
inf.lands.200kb.rep7 <- inf.lands.200kb.rep7
inf.lands.200kb.rep7$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep7)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep7)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1675.014
## m.div.200kb.2  7 -1686.944
## m.div.200kb.3  8 -1685.792
```

```r
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.4537, p-value = 0.0008495
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.2727, p-value < 2.2e-16
```

```r
hist(resid(m.div.200kb))
```

```r
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep7)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.129e-03 -4.377e-05  1.575e-05  7.812e-05  4.801e-04
##
## Coefficients:
```

```
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.341e-03  1.908e-05 175.137   <2e-16 ***
## thetaC        1.104e+00  1.358e-02  81.322   <2e-16 ***
## rhoC          2.579e-02  1.041e-02   2.477   0.0147 *
## tmrcaC        4.301e-03  1.323e-04  32.515   <2e-16 ***
## thetaC:tmrcaC 1.493e+00  8.184e-02  18.247   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001944 on 113 degrees of freedom
## Multiple R-squared:  0.9923, Adjusted R-squared:  0.9921
## F-statistic:  3652 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.200kb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 7] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep7, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep7, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep7, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep7, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1699.268
## g.div.200kb.2  8 -1683.987
## g.div.200kb.3  7 -1674.195
## g.div.200kb.4  7 -1698.505
```

```
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep7
##         AIC       BIC   logLik
##   -1674.195 -1654.801 844.0976
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
```

```
##     power
## 0.2235436
##
## Coefficients:
##                   Value  Std.Error   t-value p-value
## (Intercept)   0.0033379 0.00001961 170.23653  0.0000
## thetaC        1.1020624 0.01463631  75.29644  0.0000
## rhoC          0.0269409 0.01057491   2.54762  0.0122
## tmrcaC        0.0043212 0.00014132  30.57810  0.0000
## thetaC:tmrcaC 1.5314588 0.09352710  16.37449  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.309
## rhoC          -0.017 -0.333
## tmrcaC        -0.248 -0.333 -0.236
## thetaC:tmrcaC -0.396 -0.376  0.054  0.631
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -5.6176121 -0.2278995  0.1068782  0.4264629  2.6382834
##
## Residual standard error: 0.0007080388
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.430415      1.377155      2.059089      1.818743
```

6.3.8 Replicate 8

```
rep8.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep8_w200000.csv", sep = ",", header = T)

rep_8.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.di
rep_8.pi.200kb$avg <- apply(rep_8.pi.200kb[4:ncol(rep_8.pi.200kb)], 1, mean)
rep_8.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.r
rep_8.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5
rep_8.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5
rep_8.tmrca.200kb$avg <- apply(rep_8.tmrca.200kb[4:ncol(rep_8.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep8 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep8.sim.tmrca.20
names(true.lands.200kb.rep8) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep8$thetaC <- true.lands.200kb.rep8$theta - mean(true.lands.200kb.rep8$theta)
true.lands.200kb.rep8$tmrcaC <- true.lands.200kb.rep8$tmrca - mean(true.lands.200kb.rep8$tmrca)
true.lands.200kb.rep8$rhoC <- true.lands.200kb.rep8$rho - mean(true.lands.200kb.rep8$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep8)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.200kb.true[2, 8] <- anova.diversity$VarExp[1] * 100
```

```r
r2.bgs.200kb.true[3, 8] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 8] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 8] <- anova.diversity$VarExp[4] * 100

inf.lands.200kb.rep8 <- as.data.frame(cbind(rep_8.pi.200kb$avg, rep_8.theta.200kb$sample_mean, rep_8.tm
names(inf.lands.200kb.rep8) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[8, 2] <- cor.test(sim.theta.200kb$Rate, rep_8.theta.200kb$sample_mean, method = "spearma
cor.table.rho[8, 2] <- cor.test(sim.rho.200kb$Rate, rep_8.rho.200kb$sample_mean, method = "spearman")$ea
cor.table.tmrca[8, 2] <- cor.test(rep8.sim.tmrca.200kb$AverageTmrca, rep_8.tmrca.200kb$avg, method = "sp

# centering
inf.lands.200kb.rep8$thetaC <- inf.lands.200kb.rep8$theta - mean(inf.lands.200kb.rep8$theta)
inf.lands.200kb.rep8$tmrcaC <- inf.lands.200kb.rep8$tmrca - mean(inf.lands.200kb.rep8$tmrca)
inf.lands.200kb.rep8$rhoC <- inf.lands.200kb.rep8$rho - mean(inf.lands.200kb.rep8$rho)

inf.lands.200kb.rep8$bin <- 1:nrow(inf.lands.200kb.rep8)

# for merging:
inf.lands.200kb.rep8 <- inf.lands.200kb.rep8
inf.lands.200kb.rep8$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep8)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep8)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1711.201
## m.div.200kb.2  7 -1752.777
## m.div.200kb.3  8 -1754.136
```

```r
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.4172, p-value = 0.0004208
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.36722, p-value = 0.018
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep8)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.995e-04 -7.501e-05  1.296e-05  7.253e-05  5.573e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.435e-03  1.548e-05 221.813   <2e-16 ***
## thetaC        1.237e+00  1.300e-02  95.133   <2e-16 ***
## rhoC          9.601e-03  9.644e-03   0.996    0.322
## tmrcaC        3.336e-03  6.085e-05  54.824   <2e-16 ***
## thetaC:tmrcaC 1.303e+00  5.161e-02  25.255   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001668 on 113 degrees of freedom
## Multiple R-squared:  0.9943, Adjusted R-squared:  0.9941
## F-statistic:  4932 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
```

```r
r2.bgs.200kb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 8] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep8, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep8, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep8, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep8, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1755.740
## g.div.200kb.2  8 -1748.646
## g.div.200kb.3  7 -1738.511
## g.div.200kb.4  7 -1746.257
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep8
##         AIC       BIC   logLik
##   -1738.511 -1719.117 876.2557
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.9408611
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0034413 0.00001606 214.34353  0.0000
## thetaC        1.2493957 0.01435473  87.03720  0.0000
## rhoC          0.0146563 0.00812362   1.80415  0.0739
## tmrcaC        0.0032989 0.00005854  56.35435  0.0000
## thetaC:tmrcaC 1.0505342 0.05573402  18.84906  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.669
## rhoC          -0.030 -0.307
## tmrcaC        -0.024  0.044 -0.398
## thetaC:tmrcaC -0.138 -0.320  0.016  0.489
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -3.28659664 -0.55199225  0.07202289  0.46452382  2.73812395
##
## Residual standard error: 0.03959627
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.256856      1.355579      1.701715      1.595889
```

6.3.9 Replicate 9

```
rep9.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep9_w200000.csv", sep = ",", header = T)

rep_9.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.di
rep_9.pi.200kb$avg <- apply(rep_9.pi.200kb[4:ncol(rep_9.pi.200kb)], 1, mean)
rep_9.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.r
rep_9.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5
rep_9.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5
rep_9.tmrca.200kb$avg <- apply(rep_9.tmrca.200kb[4:ncol(rep_9.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep9 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep9.sim.tmrca.20
names(true.lands.200kb.rep9) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep9$thetaC <- true.lands.200kb.rep9$theta - mean(true.lands.200kb.rep9$theta)
true.lands.200kb.rep9$tmrcaC <- true.lands.200kb.rep9$tmrca - mean(true.lands.200kb.rep9$tmrca)
true.lands.200kb.rep9$rhoC <- true.lands.200kb.rep9$rho - mean(true.lands.200kb.rep9$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep9)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Var
r2.bgs.200kb.true[2, 9] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb.true[3, 9] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 9] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 9] <- anova.diversity$VarExp[4] * 100

inf.lands.200kb.rep9 <- as.data.frame(cbind(rep_9.pi.200kb$avg, rep_9.theta.200kb$sample_mean, rep_9.tm
names(inf.lands.200kb.rep9) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[9, 2] <- cor.test(sim.theta.200kb$Rate, rep_9.theta.200kb$sample_mean, method = "spearma
cor.table.rho[9, 2] <- cor.test(sim.rho.200kb$Rate, rep_9.rho.200kb$sample_mean, method = "spearman")$e
cor.table.tmrca[9, 2] <- cor.test(rep9.sim.tmrca.200kb$AverageTmrca, rep_9.tmrca.200kb$avg, method = "s

# centering
inf.lands.200kb.rep9$thetaC <- inf.lands.200kb.rep9$theta - mean(inf.lands.200kb.rep9$theta)
inf.lands.200kb.rep9$tmrcaC <- inf.lands.200kb.rep9$tmrca - mean(inf.lands.200kb.rep9$tmrca)
inf.lands.200kb.rep9$rhoC <- inf.lands.200kb.rep9$rho - mean(inf.lands.200kb.rep9$rho)

inf.lands.200kb.rep9$bin <- 1:nrow(inf.lands.200kb.rep9)
```

```r
# for merging:
inf.lands.200kb.rep9 <- inf.lands.200kb.rep9
inf.lands.200kb.rep9$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep9)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep9)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1753.774
## m.div.200kb.2  7 -1752.849
## m.div.200kb.3  8 -1750.851
```

```r
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.9167, p-value = 0.2774
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
## HMC = 0.31553, p-value = 0.001
```

```r
hist(resid(m.div.200kb))
```

```r
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep9)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -3.962e-04 -7.380e-05 -6.080e-06  5.746e-05  4.650e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.408e-03  1.349e-05 252.580  < 2e-16 ***
## thetaC        1.189e+00  1.137e-02 104.514  < 2e-16 ***
## rhoC          2.129e-02  8.117e-03   2.623  0.00991 **
## tmrcaC        3.851e-03  7.068e-05  54.487  < 2e-16 ***
## thetaC:tmrcaC 1.520e+00  5.234e-02  29.035  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001392 on 113 degrees of freedom
## Multiple R-squared:  0.9961, Adjusted R-squared:  0.9959
## F-statistic:  7167 on 4 and 113 DF,  p-value: < 2.2e-16
```

```r
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.bgs.200kb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.200kb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 9] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep9, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep9, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep9, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep9, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1763.451
## g.div.200kb.2  8 -1750.490
## g.div.200kb.3  7 -1752.426
## g.div.200kb.4  7 -1765.407
```

```
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep9
##         AIC       BIC   logLik
##   -1752.426 -1733.031 883.2129
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.1388582
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0034089 0.00001363 250.13516  0.0000
## thetaC        1.1896415 0.01189557 100.00708  0.0000
## rhoC          0.0210014 0.00826443   2.54118  0.0124
## tmrcaC        0.0038498 0.00007232  53.23173  0.0000
## thetaC:tmrcaC 1.5109710 0.05625278  26.86038  0.0000
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.246
## rhoC          -0.026 -0.309
## tmrcaC        -0.140 -0.243 -0.369
```

```
## thetaC:tmrcaC -0.338 -0.452  0.086  0.405
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.76264494 -0.52309076 -0.04576363  0.42768957  3.47227913
##
## Residual standard error: 0.0003107247
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.472047      1.461443      1.584082      1.445190
```

6.3.10 Replicate 10

```
rep10.sim.tmrca.200kb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep10_w200000.csv", sep = ",", header = T

rep_10.pi.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_
rep_10.pi.200kb$avg <- apply(rep_10.pi.200kb[4:ncol(rep_10.pi.200kb)], 1, mean)
rep_10.rho.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10
rep_10.theta.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep
rep_10.tmrca.200kb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep
rep_10.tmrca.200kb$avg <- apply(rep_10.tmrca.200kb[4:ncol(rep_10.tmrca.200kb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.200kb.rep10 <- as.data.frame(cbind(rep_1.pi.200kb$avg, sim.theta.200kb$Rate, rep10.sim.tmrca
names(true.lands.200kb.rep10) <- c("diversity", "theta", "tmrca", "rho")
true.lands.200kb.rep10$thetaC <- true.lands.200kb.rep10$theta - mean(true.lands.200kb.rep10$theta)
true.lands.200kb.rep10$tmrcaC <- true.lands.200kb.rep10$tmrca - mean(true.lands.200kb.rep10$tmrca)
true.lands.200kb.rep10$rhoC <- true.lands.200kb.rep10$rho - mean(true.lands.200kb.rep10$rho, na.rm = T)

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.200kb.rep10)
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.200kb.true[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$Va
r2.bgs.200kb.true[2, 10] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb.true[3, 10] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb.true[4, 10] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb.true[5, 10] <- anova.diversity$VarExp[4] * 100

inf.lands.200kb.rep10 <- as.data.frame(cbind(rep_10.pi.200kb$avg, rep_10.theta.200kb$sample_mean, rep_1
names(inf.lands.200kb.rep10) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[10, 2] <- cor.test(sim.theta.200kb$Rate, rep_10.theta.200kb$sample_mean, method = "spea
cor.table.rho[10, 2] <- cor.test(sim.rho.200kb$Rate, rep_10.rho.200kb$sample_mean, method = "spearman")
cor.table.tmrca[10, 2] <- cor.test(rep10.sim.tmrca.200kb$AverageTmrca, rep_10.tmrca.200kb$avg, method =

# centering
inf.lands.200kb.rep10$thetaC <- inf.lands.200kb.rep10$theta - mean(inf.lands.200kb.rep10$theta)
inf.lands.200kb.rep10$tmrcaC <- inf.lands.200kb.rep10$tmrca - mean(inf.lands.200kb.rep10$tmrca)
inf.lands.200kb.rep10$rhoC <- inf.lands.200kb.rep10$rho - mean(inf.lands.200kb.rep10$rho)
```

```r
inf.lands.200kb.rep10$bin <- 1:nrow(inf.lands.200kb.rep10)

# for merging:
inf.lands.200kb.rep10 <- inf.lands.200kb.rep10
inf.lands.200kb.rep10$Replicate <- 1

m.div.200kb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.200kb.rep10)
m.div.200kb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.2
m.div.200kb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.200kb.rep10)

AIC(m.div.200kb, m.div.200kb.2, m.div.200kb.3)
```

```
##               df       AIC
## m.div.200kb    6 -1734.618
## m.div.200kb.2  7 -1735.599
## m.div.200kb.3  8 -1737.127
```

```r
plot(resid(m.div.200kb)~fitted(m.div.200kb))
```

```r
dwtest(m.div.200kb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.200kb
## DW = 1.2969, p-value = 3.058e-05
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.200kb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.200kb
```

```
## HMC = 0.39016, p-value = 0.042
```

```
hist(resid(m.div.200kb))
```

```
summary(m.div.200kb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.200kb.rep10)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.675e-04 -5.994e-05  2.700e-07  8.305e-05  4.753e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.461e-03  1.419e-05 244.019  < 2e-16 ***
## thetaC        1.063e+00  9.242e-03 115.002  < 2e-16 ***
## rhoC          3.280e-02  1.017e-02   3.224  0.00165 **
## tmrcaC        4.528e-03  1.033e-04  43.821  < 2e-16 ***
## thetaC:tmrcaC 1.330e+00  5.244e-02  25.371  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000151 on 113 degrees of freedom
## Multiple R-squared:  0.9954, Adjusted R-squared:  0.9953
## F-statistic:  6149 on 4 and 113 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.200kb)
apiss <- anova.diversity$"Sum Sq"
```

```r
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.200kb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3
r2.bgs.200kb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.bgs.200kb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.bgs.200kb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.bgs.200kb[5, 10] <- anova.diversity$VarExp[4] * 100

g.div.200kb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                  data = inf.lands.200kb.rep10, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), m

g.div.200kb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep10, weights = varPower(0, ~theta), cor = corAR1(0, ~bin),

g.div.200kb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep10, weights = varPower(0, ~theta), method = "ML")

g.div.200kb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.200kb.rep10, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.200kb.1, g.div.200kb.2, g.div.200kb.3, g.div.200kb.4)
```

```
##               df       AIC
## g.div.200kb.1  8 -1746.826
## g.div.200kb.2  8 -1760.403
## g.div.200kb.3  7 -1740.621
## g.div.200kb.4  7 -1734.608
```

```r
summary(g.div.200kb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.200kb.rep10
##         AIC       BIC   logLik
##   -1740.621 -1721.226 877.3105
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.4149818
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0034619 0.00001437 240.98955   0e+00
## thetaC        1.0667145 0.01021663 104.40964   0e+00
## rhoC          0.0388858 0.00943343   4.12213   1e-04
## tmrcaC        0.0044840 0.00010213  43.90449   0e+00
## thetaC:tmrcaC 1.3172353 0.05899803  22.32677   0e+00
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.362
```

```
## rhoC           0.007 -0.354
## tmrcaC        -0.087 -0.026 -0.469
## thetaC:tmrcaC -0.228 -0.117 -0.017  0.381
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -3.925337793 -0.433792136  0.008606498  0.591307894  3.278902328
##
## Residual standard error: 0.001603282
## Degrees of freedom: 118 total; 113 residual
```

```
vif(g.div.200kb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.210381      1.593301      1.623969      1.219199
```

6.4 all replicates:

```
theta.plot <- as.data.frame(cbind(1:nrow(sim.theta.200kb),
                                  sim.theta.200kb$Rate * 1e-7 * 3e+4,
                                  rep_1.theta.200kb$sample_mean,
                                  rep_2.theta.200kb$sample_mean,
                                  rep_3.theta.200kb$sample_mean,
                                  rep_4.theta.200kb$sample_mean,
                                  rep_5.theta.200kb$sample_mean,
                                  rep_6.theta.200kb$sample_mean,
                                  rep_7.theta.200kb$sample_mean,
                                  rep_8.theta.200kb$sample_mean,
                                  rep_9.theta.200kb$sample_mean,
                                  rep_10.theta.200kb$sample_mean))

names(theta.plot) <- c("bin", "sim", reps)
molten.theta <- melt(theta.plot, id.vars = "bin")
theta.map.200kb.bgs <- ggplot(data = molten.theta, aes(x = bin * 200, y = value, colour = variable)) + 
theta.map.200kb.bgs <- theta.map.200kb.bgs + geom_line(data = molten.theta, aes(size = variable)) + scal
theta.map.200kb.bgs <- theta.map.200kb.bgs + scale_color_manual(values = c("black", brewer.pal(n = 9, na
theta.map.200kb.bgs <- theta.map.200kb.bgs + scale_x_continuous(breaks = pretty_breaks()) + scale_y_cont
theta.map.200kb.bgs <- theta.map.200kb.bgs + labs(title = NULL, x = NULL, y = NULL)
theta.map.200kb.bgs <- theta.map.200kb.bgs + theme(text = element_text(size = 20), axis.title.x = elemen
theta.map.200kb.bgs
```

6.5 1 Mb scale

```
r2.bgs.1Mb <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.bgs.1Mb) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.bgs.1Mb) <- reps

r2.bgs.1Mb.true <- as.data.frame(matrix(ncol = nreps, nrow = 5))
row.names(r2.bgs.1Mb.true) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
colnames(r2.bgs.1Mb.true) <- reps

sim.theta.1Mb <- read.table("dm_bgs_sims/MutationMap_1000kb.csv", sep = ",", header = T)
sim.rho.1Mb <- read.table("dm_bgs_sims/RecombinationMap_1000kb.csv", sep = ",", header = T)
```

6.5.1 Replicate 1

```
rep1.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep1_w1000000.csv", sep = ",", header = T)

rep_1.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.dive
rep_1.pi.1Mb$avg <- apply(rep_1.pi.1Mb[4:ncol(rep_1.pi.1Mb)], 1, mean)
rep_1.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.rho
rep_1.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.
rep_1.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep1/dm2L_bgs_rep1_5.
rep_1.tmrca.1Mb$avg <- apply(rep_1.tmrca.1Mb[4:ncol(rep_1.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep1 <- as.data.frame(cbind(rep_1.pi.1Mb$avg, sim.theta.1Mb$Rate, rep1.sim.tmrca.1Mb$Ave
```

```r
names(true.lands.1Mb.rep1) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep1$thetaC <- true.lands.1Mb.rep1$theta - mean(true.lands.1Mb.rep1$theta)
true.lands.1Mb.rep1$tmrcaC <- true.lands.1Mb.rep1$tmrca - mean(true.lands.1Mb.rep1$tmrca)
true.lands.1Mb.rep1$rhoC <- true.lands.1Mb.rep1$rho - mean(true.lands.1Mb.rep1$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep1)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp
r2.bgs.1Mb.true[2, 1] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb.true[3, 1] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 1] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 1] <- anova.diversity$VarExp[4] * 100


inf.lands.1Mb.rep1 <- as.data.frame(cbind(rep_1.pi.1Mb$avg, rep_1.theta.1Mb$sample_mean, rep_1.tmrca.1M
names(inf.lands.1Mb.rep1) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[1, 3] <- cor.test(sim.theta.1Mb$Rate, rep_1.theta.1Mb$sample_mean, method = "spearman")$
cor.table.rho[1, 3] <- cor.test(sim.rho.200kb$Rate, rep_1.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[1, 3] <- cor.test(rep1.sim.tmrca.1Mb$AverageTmrca, rep_1.tmrca.1Mb$avg, method = "spear

# centering
inf.lands.1Mb.rep1$thetaC <- inf.lands.1Mb.rep1$theta - mean(inf.lands.1Mb.rep1$theta)
inf.lands.1Mb.rep1$tmrcaC <- inf.lands.1Mb.rep1$tmrca - mean(inf.lands.1Mb.rep1$tmrca)
inf.lands.1Mb.rep1$rhoC <- inf.lands.1Mb.rep1$rho - mean(inf.lands.1Mb.rep1$rho)

inf.lands.1Mb.rep1$bin <- 1:nrow(inf.lands.1Mb.rep1)

# for merging:
inf.lands.1Mb.rep1 <- inf.lands.1Mb.rep1
inf.lands.1Mb.rep1$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep1)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep1)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -350.1912
## m.div.1Mb.2  7 -353.4005
## m.div.1Mb.3  8 -351.4502
```

```r
plot(resid(m.div.1Mb.2)~fitted(m.div.1Mb.2))
```

```
dwtest(m.div.1Mb.2)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb.2
## DW = 2.0504, p-value = 0.4914
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb.2)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb.2
## HMC = 0.57197, p-value = 0.686
```

```
hist(resid(m.div.1Mb.2))
```

```
summary(m.div.1Mb.2)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC +
##     rhoC:tmrcaC, data = inf.lands.1Mb.rep1)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.588e-04 -7.924e-05 -2.075e-05  3.932e-05  2.998e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.342e-03  5.732e-05  58.303  < 2e-16 ***
## thetaC        9.934e-01  7.341e-02  13.532 7.13e-11 ***
## rhoC          3.751e-02  3.528e-02   1.063   0.3018
## tmrcaC        5.305e-03  6.288e-04   8.436 1.14e-07 ***
## thetaC:tmrcaC 2.105e+00  9.052e-01   2.325   0.0319 *
## rhoC:tmrcaC   5.481e-01  2.624e-01   2.089   0.0512 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001325 on 18 degrees of freedom
## Multiple R-squared:  0.994,  Adjusted R-squared:  0.9923
## F-statistic: 594.7 on 5 and 18 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb.2)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.bgs.1Mb[1, 1] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 1] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 1] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 1] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 1] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep1, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep1, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep1, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep1, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -347.5891
## g.div.1Mb.2  8 -348.1816
## g.div.1Mb.3  7 -348.6342
## g.div.1Mb.4  7 -348.6639
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep1
##         AIC       BIC   logLik
##   -348.6342 -340.3878 181.3171
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.4951664
##
## Coefficients:
##                    Value Std.Error  t-value p-value
## (Intercept)   0.0034361 0.0000408 84.27643  0.0000
## thetaC        1.0132773 0.0824669 12.28708  0.0000
## rhoC          0.0755056 0.0307041  2.45914  0.0237
## tmrcaC        0.0040537 0.0002377 17.05650  0.0000
## thetaC:tmrcaC 2.0534919 1.0009642  2.05151  0.0543
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.678
## rhoC          -0.286 -0.638
## tmrcaC        -0.346 -0.298 -0.275
```

```
## thetaC:tmrcaC -0.685 -0.871  0.418  0.505
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.4032965 -0.4764999 -0.2622762  0.2758051  3.3133189
##
## Residual standard error: 0.002225736
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      6.680534      2.811830      2.189128      5.727340
```

6.5.2 Replicate 2

```
rep2.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep2_w1000000.csv", sep = ",", header = T)

rep_2.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.div
rep_2.pi.1Mb$avg <- apply(rep_2.pi.1Mb[4:ncol(rep_2.pi.1Mb)], 1, mean)
rep_2.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.rho
rep_2.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.
rep_2.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep2/dm2L_bgs_rep2_5.
rep_2.tmrca.1Mb$avg <- apply(rep_2.tmrca.1Mb[4:ncol(rep_2.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep2 <- as.data.frame(cbind(rep_2.pi.1Mb$avg, sim.theta.1Mb$Rate, rep2.sim.tmrca.1Mb$Ave
names(true.lands.1Mb.rep2) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep2$thetaC <- true.lands.1Mb.rep2$theta - mean(true.lands.1Mb.rep2$theta)
true.lands.1Mb.rep2$tmrcaC <- true.lands.1Mb.rep2$tmrca - mean(true.lands.1Mb.rep2$tmrca)
true.lands.1Mb.rep2$rhoC <- true.lands.1Mb.rep2$rho - mean(true.lands.1Mb.rep2$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep2)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp
r2.bgs.1Mb.true[2, 2] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb.true[3, 2] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 2] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 2] <- anova.diversity$VarExp[4] * 100

inf.lands.1Mb.rep2 <- as.data.frame(cbind(rep_2.pi.1Mb$avg, rep_2.theta.1Mb$sample_mean, rep_2.tmrca.1M
names(inf.lands.1Mb.rep2) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[2, 3] <- cor.test(sim.theta.1Mb$Rate, rep_2.theta.1Mb$sample_mean, method = "spearman")$
cor.table.rho[2, 3] <- cor.test(sim.rho.200kb$Rate, rep_2.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[2, 3] <- cor.test(rep2.sim.tmrca.1Mb$AverageTmrca, rep_2.tmrca.1Mb$avg, method = "spear

# centering
inf.lands.1Mb.rep2$thetaC <- inf.lands.1Mb.rep2$theta - mean(inf.lands.1Mb.rep2$theta)
inf.lands.1Mb.rep2$tmrcaC <- inf.lands.1Mb.rep2$tmrca - mean(inf.lands.1Mb.rep2$tmrca)
inf.lands.1Mb.rep2$rhoC <- inf.lands.1Mb.rep2$rho - mean(inf.lands.1Mb.rep2$rho)
```

```r
inf.lands.1Mb.rep2$bin <- 1:nrow(inf.lands.1Mb.rep2)

# for merging:
inf.lands.1Mb.rep2 <- inf.lands.1Mb.rep2
inf.lands.1Mb.rep2$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep2)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep2)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -340.4890
## m.div.1Mb.2  7 -338.4960
## m.div.1Mb.3  8 -340.4696
```

```r
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.2615, p-value = 0.5747
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
```

```
## HMC = 0.50859, p-value = 0.513
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.085e-04 -1.217e-04 -2.989e-05  5.085e-05  4.677e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.476e-03  3.592e-05  96.792  < 2e-16 ***
## thetaC        1.305e+00  5.560e-02  23.465 1.71e-15 ***
## rhoC          8.429e-03  3.289e-02   0.256   0.8005
## tmrcaC        3.602e-03  2.238e-04  16.094 1.59e-12 ***
## thetaC:tmrcaC 5.490e-01  2.724e-01   2.016   0.0582 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001759 on 19 degrees of freedom
## Multiple R-squared:  0.9897, Adjusted R-squared:  0.9876
## F-statistic: 457.9 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
```

```r
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 2] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 2] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 2] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 2] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 2] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep2, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep2, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep2, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep2, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -338.4110
## g.div.1Mb.2  8 -339.3155
## g.div.1Mb.3  7 -340.4491
## g.div.1Mb.4  7 -338.4931
```

```r
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep2
##         AIC       BIC   logLik
##   -340.4491 -332.2027 177.2246
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##    power
## 1.281934
##
## Coefficients:
##                    Value  Std.Error  t-value p-value
## (Intercept)   0.0034897 0.00003930 88.78892  0.0000
## thetaC        1.3432539 0.06488550 20.70191  0.0000
## rhoC          0.0248667 0.02714294  0.91614  0.3711
## tmrcaC        0.0034959 0.00019680 17.76351  0.0000
## thetaC:tmrcaC 0.7237951 0.27108673  2.66998  0.0151
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.495
```

```
## rhoC          -0.089 -0.609
## tmrcaC         0.091  0.581 -0.543
## thetaC:tmrcaC -0.021 -0.448  0.052 -0.310
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.3253500 -0.4724077 -0.2657178  0.3816227  2.4010863
##
## Residual standard error: 0.2910646
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.340049      1.981495      1.704582      1.430447
```

6.5.3 Replicate 3

```
rep3.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep3_w1000000.csv", sep = ",", header = T)

rep_3.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.div
rep_3.pi.1Mb$avg <- apply(rep_3.pi.1Mb[4:ncol(rep_3.pi.1Mb)], 1, mean)
rep_3.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.rho
rep_3.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.
rep_3.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep3/dm2L_bgs_rep3_5.
rep_3.tmrca.1Mb$avg <- apply(rep_3.tmrca.1Mb[4:ncol(rep_3.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep3 <- as.data.frame(cbind(rep_2.pi.1Mb$avg, sim.theta.1Mb$Rate, rep3.sim.tmrca.1Mb$Ave
names(true.lands.1Mb.rep3) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep3$thetaC <- true.lands.1Mb.rep3$theta - mean(true.lands.1Mb.rep3$theta)
true.lands.1Mb.rep3$tmrcaC <- true.lands.1Mb.rep3$tmrca - mean(true.lands.1Mb.rep3$tmrca)
true.lands.1Mb.rep3$rhoC <- true.lands.1Mb.rep3$rho - mean(true.lands.1Mb.rep3$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep3)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp
r2.bgs.1Mb.true[2, 3] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb.true[3, 3] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 3] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 3] <- anova.diversity$VarExp[4] * 100

inf.lands.1Mb.rep3 <- as.data.frame(cbind(rep_3.pi.1Mb$avg, rep_3.theta.1Mb$sample_mean, rep_3.tmrca.1M
names(inf.lands.1Mb.rep3) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[3, 3] <- cor.test(sim.theta.1Mb$Rate, rep_3.theta.1Mb$sample_mean, method = "spearman")$
cor.table.rho[3, 3] <- cor.test(sim.rho.200kb$Rate, rep_3.rho.200kb$sample_mean, method = "spearman")$e
cor.table.tmrca[3, 3] <- cor.test(rep3.sim.tmrca.1Mb$AverageTmrca, rep_3.tmrca.1Mb$avg, method = "spear

# centering
inf.lands.1Mb.rep3$thetaC <- inf.lands.1Mb.rep3$theta - mean(inf.lands.1Mb.rep3$theta)
inf.lands.1Mb.rep3$tmrcaC <- inf.lands.1Mb.rep3$tmrca - mean(inf.lands.1Mb.rep3$tmrca)
```

```r
inf.lands.1Mb.rep3$rhoC <- inf.lands.1Mb.rep3$rho - mean(inf.lands.1Mb.rep3$rho)

inf.lands.1Mb.rep3$bin <- 1:nrow(inf.lands.1Mb.rep3)

# for merging:
inf.lands.1Mb.rep3 <- inf.lands.1Mb.rep3
inf.lands.1Mb.rep3$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep3)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep3)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -345.5820
## m.div.1Mb.2  7 -346.9896
## m.div.1Mb.3  8 -348.3194
```

```r
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.5052, p-value = 0.8416
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
```

```
##
## data:  m.div.1Mb
## HMC = 0.52869, p-value = 0.539
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.532e-04 -9.316e-05  1.061e-05  5.944e-05  4.466e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.517e-03  3.231e-05 108.850  < 2e-16 ***
## thetaC         1.323e+00  4.594e-02  28.788  < 2e-16 ***
## rhoC          -7.069e-03  3.016e-02  -0.234  0.81719
## tmrcaC         3.670e-03  2.069e-04  17.732 2.81e-13 ***
## thetaC:tmrcaC  7.505e-01  1.650e-01   4.549  0.00022 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001582 on 19 degrees of freedom
## Multiple R-squared:  0.9905, Adjusted R-squared:  0.9885
## F-statistic: 495.2 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 3] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 3] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 3] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 3] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 3] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep3, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep3, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep3, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep3, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -345.3677
## g.div.1Mb.2  8 -347.1405
## g.div.1Mb.3  7 -346.0258
## g.div.1Mb.4  7 -344.4834
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep3
##         AIC       BIC   logLik
##   -346.0258 -337.7794 180.0129
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##    power
## 1.225993
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)   0.0035246 0.00003454 102.03989  0.0000
## thetaC        1.3580663 0.05237639  25.92898  0.0000
## rhoC          0.0073311 0.02720431   0.26948  0.7905
## tmrcaC        0.0035822 0.00018732  19.12301  0.0000
## thetaC:tmrcaC 0.8272546 0.18843140   4.39022  0.0003
##
##  Correlation:
```

```
##                (Intr) thetaC rhoC   tmrcaC
## thetaC          0.526
## rhoC           -0.045 -0.573
## tmrcaC          0.033  0.393 -0.627
## thetaC:tmrcaC   0.009 -0.072  0.033 -0.049
##
## Standardized residuals:
##          Min           Q1          Med           Q3          Max
## -1.54924816 -0.55078534 -0.07541142  0.29979596  2.70092421
##
## Residual standard error: 0.1890319
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.497949      2.078778      1.653891      1.006384
```

6.5.4 Replicate 4

```
rep4.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep4_w1000000.csv", sep = ",", header = T)

rep_4.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.div
rep_4.pi.1Mb$avg <- apply(rep_4.pi.1Mb[4:ncol(rep_4.pi.1Mb)], 1, mean)
rep_4.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.rho
rep_4.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.
rep_4.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep4/dm2L_bgs_rep4_5.
rep_4.tmrca.1Mb$avg <- apply(rep_4.tmrca.1Mb[4:ncol(rep_4.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep4 <- as.data.frame(cbind(rep_4.pi.1Mb$avg, sim.theta.1Mb$Rate, rep4.sim.tmrca.1Mb$Ave
names(true.lands.1Mb.rep4) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep4$thetaC <- true.lands.1Mb.rep4$theta - mean(true.lands.1Mb.rep4$theta)
true.lands.1Mb.rep4$tmrcaC <- true.lands.1Mb.rep4$tmrca - mean(true.lands.1Mb.rep4$tmrca)
true.lands.1Mb.rep4$rhoC <- true.lands.1Mb.rep4$rho - mean(true.lands.1Mb.rep4$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep4)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp
r2.bgs.1Mb.true[2, 4] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb.true[3, 4] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 4] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 4] <- anova.diversity$VarExp[4] * 100

inf.lands.1Mb.rep4 <- as.data.frame(cbind(rep_4.pi.1Mb$avg, rep_4.theta.1Mb$sample_mean, rep_4.tmrca.1M
names(inf.lands.1Mb.rep4) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[4, 3] <- cor.test(sim.theta.1Mb$Rate, rep_4.theta.1Mb$sample_mean, method = "spearman")$
cor.table.rho[4, 3] <- cor.test(sim.rho.200kb$Rate, rep_4.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[4, 3] <- cor.test(rep4.sim.tmrca.1Mb$AverageTmrca, rep_4.tmrca.1Mb$avg, method = "spear

# centering
```

```r
inf.lands.1Mb.rep4$thetaC <- inf.lands.1Mb.rep4$theta - mean(inf.lands.1Mb.rep4$theta)
inf.lands.1Mb.rep4$tmrcaC <- inf.lands.1Mb.rep4$tmrca - mean(inf.lands.1Mb.rep4$tmrca)
inf.lands.1Mb.rep4$rhoC <- inf.lands.1Mb.rep4$rho - mean(inf.lands.1Mb.rep4$rho)

inf.lands.1Mb.rep4$bin <- 1:nrow(inf.lands.1Mb.rep4)

# for merging:
inf.lands.1Mb.rep4 <- inf.lands.1Mb.rep4
inf.lands.1Mb.rep4$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep4)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep4)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -342.6901
## m.div.1Mb.2  7 -343.4354
## m.div.1Mb.3  8 -341.7061
```

```r
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 1.8262, p-value = 0.2127
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.67938, p-value = 0.887
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep4)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.769e-04 -7.914e-05 -1.131e-05  4.493e-05  4.419e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.458e-03  3.683e-05  93.894  < 2e-16 ***
## thetaC        1.003e+00  4.174e-02  24.022 1.11e-15 ***
## rhoC          5.858e-02  3.134e-02   1.869  0.07708 .
## tmrcaC        5.291e-03  4.059e-04  13.037 6.31e-11 ***
## thetaC:tmrcaC 1.875e+00  6.290e-01   2.981  0.00768 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.000168 on 19 degrees of freedom
## Multiple R-squared:  0.9896, Adjusted R-squared:  0.9874
```

```
## F-statistic: 452.5 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 4] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 4] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 4] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 4] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 4] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep4, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), metl

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep4, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep4, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep4, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -343.6387
## g.div.1Mb.2  8 -338.7066
## g.div.1Mb.3  7 -340.7065
## g.div.1Mb.4  7 -345.6369
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep4
##         AIC       BIC   logLik
##   -340.7065 -332.4601 177.3532
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##      power
## 0.07850966
##
## Coefficients:
##                    Value Std.Error  t-value p-value
## (Intercept)   0.0034577 0.0000368 94.00350  0.0000
## thetaC        1.0040631 0.0422359 23.77274  0.0000
## rhoC          0.0592684 0.0315630  1.87778  0.0758
## tmrcaC        0.0052834 0.0004080 12.95035  0.0000
## thetaC:tmrcaC 1.8738540 0.6211915  3.01655  0.0071
##
```

```
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.166
## rhoC          -0.003 -0.578
## tmrcaC        -0.195 -0.025 -0.433
## thetaC:tmrcaC -0.361 -0.367  0.013  0.537
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.8895276 -0.5306627 -0.0678023  0.2944035  2.9182328
##
## Residual standard error: 0.000234782
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.908607      2.160469      1.976808      1.713511
```

6.5.5 Replicate 5

```
rep5.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep5_w1000000.csv", sep = ",", header = T)

rep_5.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.div
rep_5.pi.1Mb$avg <- apply(rep_5.pi.1Mb[4:ncol(rep_5.pi.1Mb)], 1, mean)
rep_5.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.rho
rep_5.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.
rep_5.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep5/dm2L_bgs_rep5_5.
rep_5.tmrca.1Mb$avg <- apply(rep_5.tmrca.1Mb[4:ncol(rep_5.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep5 <- as.data.frame(cbind(rep_4.pi.1Mb$avg, sim.theta.1Mb$Rate, rep5.sim.tmrca.1Mb$Ave
names(true.lands.1Mb.rep5) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep5$thetaC <- true.lands.1Mb.rep5$theta - mean(true.lands.1Mb.rep5$theta)
true.lands.1Mb.rep5$tmrcaC <- true.lands.1Mb.rep5$tmrca - mean(true.lands.1Mb.rep5$tmrca)
true.lands.1Mb.rep5$rhoC <- true.lands.1Mb.rep5$rho - mean(true.lands.1Mb.rep5$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep5)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp
r2.bgs.1Mb.true[2, 5] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb.true[3, 5] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 5] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 5] <- anova.diversity$VarExp[4] * 100

inf.lands.1Mb.rep5 <- as.data.frame(cbind(rep_5.pi.1Mb$avg, rep_5.theta.1Mb$sample_mean, rep_5.tmrca.1M
names(inf.lands.1Mb.rep5) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[5, 3] <- cor.test(sim.theta.1Mb$Rate, rep_5.theta.1Mb$sample_mean, method = "spearman")$
cor.table.rho[5, 3] <- cor.test(sim.rho.200kb$Rate, rep_5.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[5, 3] <- cor.test(rep5.sim.tmrca.1Mb$AverageTmrca, rep_5.tmrca.1Mb$avg, method = "spear
```

```r
# centering
inf.lands.1Mb.rep5$thetaC <- inf.lands.1Mb.rep5$theta - mean(inf.lands.1Mb.rep5$theta)
inf.lands.1Mb.rep5$tmrcaC <- inf.lands.1Mb.rep5$tmrca - mean(inf.lands.1Mb.rep5$tmrca)
inf.lands.1Mb.rep5$rhoC <- inf.lands.1Mb.rep5$rho - mean(inf.lands.1Mb.rep5$rho)

inf.lands.1Mb.rep5$bin <- 1:nrow(inf.lands.1Mb.rep5)

# for merging:
inf.lands.1Mb.rep5 <- inf.lands.1Mb.rep5
inf.lands.1Mb.rep5$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep5)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep5)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -346.9226
## m.div.1Mb.2  7 -345.5662
## m.div.1Mb.3  8 -347.9807
```

```r
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.3054, p-value = 0.6445
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.58123, p-value = 0.734
```

```r
hist(resid(m.div.1Mb))
```

```r
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep5)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.555e-04 -1.124e-04 -2.460e-06  3.966e-05  4.422e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.463e-03  3.141e-05 110.243  < 2e-16 ***
## thetaC         1.311e+00  4.668e-02  28.076  < 2e-16 ***
## rhoC          -1.091e-02  2.788e-02  -0.391  0.69992
## tmrcaC         3.365e-03  1.899e-04  17.720 2.84e-13 ***
## thetaC:tmrcaC  7.466e-01  2.170e-01   3.441  0.00274 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.0001538 on 19 degrees of freedom
## Multiple R-squared:  0.9917, Adjusted R-squared:  0.9899
## F-statistic: 564.1 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 5] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 5] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 5] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 5] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 5] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep5, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), metl

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep5, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep5, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep5, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -345.3251
## g.div.1Mb.2  8 -345.5873
## g.div.1Mb.3  7 -345.9151
## g.div.1Mb.4  7 -345.0317
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep5
##         AIC       BIC   logLik
##   -345.9151 -337.6687 179.9576
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##    power
## 1.118503
##
## Coefficients:
##                   Value  Std.Error   t-value p-value
## (Intercept)   0.0034688 0.00003433 101.02877  0.0000
## thetaC        1.3283474 0.05424610  24.48743  0.0000
## rhoC          0.0084012 0.02639289   0.31831  0.7537
```

```
## tmrcaC        0.0032478 0.00017056  19.04241  0.0000
## thetaC:tmrcaC 0.8841209 0.20550441   4.30220  0.0004
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.479
## rhoC          -0.052 -0.647
## tmrcaC         0.056  0.455 -0.535
## thetaC:tmrcaC -0.040 -0.216 -0.002 -0.176
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.2131276 -0.7213807 -0.1575783  0.3239893  2.7480924
##
## Residual standard error: 0.1012663
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      1.898829      2.087085      1.486258      1.121324
```

6.5.6 Replicate 6

```
rep6.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep6_w1000000.csv", sep = ",", header = T)

rep_6.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.div
rep_6.pi.1Mb$avg <- apply(rep_6.pi.1Mb[4:ncol(rep_6.pi.1Mb)], 1, mean)
rep_6.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.rho
rep_6.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.
rep_6.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep6/dm2L_bgs_rep6_5.
rep_6.tmrca.1Mb$avg <- apply(rep_6.tmrca.1Mb[4:ncol(rep_6.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep6 <- as.data.frame(cbind(rep_6.pi.1Mb$avg, sim.theta.1Mb$Rate, rep6.sim.tmrca.1Mb$Ave
names(true.lands.1Mb.rep6) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep6$thetaC <- true.lands.1Mb.rep6$theta - mean(true.lands.1Mb.rep6$theta)
true.lands.1Mb.rep6$tmrcaC <- true.lands.1Mb.rep6$tmrca - mean(true.lands.1Mb.rep6$tmrca)
true.lands.1Mb.rep6$rhoC <- true.lands.1Mb.rep6$rho - mean(true.lands.1Mb.rep6$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep6)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp
r2.bgs.1Mb.true[2, 6] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb.true[3, 6] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 6] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 6] <- anova.diversity$VarExp[4] * 100

inf.lands.1Mb.rep6 <- as.data.frame(cbind(rep_6.pi.1Mb$avg, rep_6.theta.1Mb$sample_mean, rep_6.tmrca.1M
names(inf.lands.1Mb.rep6) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[6, 3] <- cor.test(sim.theta.1Mb$Rate, rep_6.theta.1Mb$sample_mean, method = "spearman")
```

```r
cor.table.rho[6, 3] <- cor.test(sim.rho.200kb$Rate, rep_6.rho.200kb$sample_mean, method = "spearman")$es
cor.table.tmrca[6, 3] <- cor.test(rep6.sim.tmrca.1Mb$AverageTmrca, rep_6.tmrca.1Mb$avg, method = "spearm

# centering
inf.lands.1Mb.rep6$thetaC <- inf.lands.1Mb.rep6$theta - mean(inf.lands.1Mb.rep6$theta)
inf.lands.1Mb.rep6$tmrcaC <- inf.lands.1Mb.rep6$tmrca - mean(inf.lands.1Mb.rep6$tmrca)
inf.lands.1Mb.rep6$rhoC <- inf.lands.1Mb.rep6$rho - mean(inf.lands.1Mb.rep6$rho)

inf.lands.1Mb.rep6$bin <- 1:nrow(inf.lands.1Mb.rep6)

# for merging:
inf.lands.1Mb.rep6 <- inf.lands.1Mb.rep6
inf.lands.1Mb.rep6$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep6)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1Mb
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep6)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -341.5943
## m.div.1Mb.2  7 -342.4713
## m.div.1Mb.3  8 -343.2357
```

```r
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
```

```
## DW = 2.1636, p-value = 0.4919
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.65357, p-value = 0.846
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep6)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.957e-04 -9.541e-05 -4.075e-05  2.041e-05  5.490e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.457e-03  3.707e-05  93.270  < 2e-16 ***
## thetaC        1.115e+00  4.504e-02  24.757 6.38e-16 ***
## rhoC          2.697e-02  2.765e-02   0.975    0.342
## tmrcaC        4.032e-03  2.537e-04  15.894 1.98e-12 ***
## thetaC:tmrcaC 7.433e-01  4.689e-01   1.585    0.129
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001719 on 19 degrees of freedom
## Multiple R-squared:  0.9901, Adjusted R-squared:  0.988
## F-statistic: 475.1 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 6] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 6] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 6] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 6] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 6] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep6, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep6, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep6, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep6, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -341.9099
## g.div.1Mb.2  8 -343.7706
## g.div.1Mb.3  7 -342.0432
## g.div.1Mb.4  7 -343.8311
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep6
##         AIC       BIC   logLik
##   -342.0432 -333.7968 178.0216
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##    power
## 3.130632
##
## Coefficients:
##                    Value Std.Error  t-value p-value
## (Intercept)   0.0035759 0.0000507 70.55075  0.0000
```

```
## thetaC        1.2887546 0.0524764 24.55877  0.0000
## rhoC          0.0174945 0.0150458  1.16275  0.2593
## tmrcaC        0.0042388 0.0002674 15.85209  0.0000
## thetaC:tmrcaC 0.9962534 0.3461699  2.87793  0.0096
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.898
## rhoC          -0.264 -0.504
## tmrcaC         0.062  0.099 -0.443
## thetaC:tmrcaC -0.417 -0.518 -0.211  0.635
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.4121973 -0.7867277 -0.4832574  0.1781136  2.8300575
##
## Residual standard error: 11371.55
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      4.311521      2.306290      2.925110      5.296967
```

6.5.7 Replicate 7

```
rep7.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep7_w1000000.csv", sep = ",", header = T)

rep_7.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.div
rep_7.pi.1Mb$avg <- apply(rep_7.pi.1Mb[4:ncol(rep_7.pi.1Mb)], 1, mean)
rep_7.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.rho
rep_7.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.
rep_7.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep7/dm2L_bgs_rep7_5.
rep_7.tmrca.1Mb$avg <- apply(rep_7.tmrca.1Mb[4:ncol(rep_7.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep7 <- as.data.frame(cbind(rep_7.pi.1Mb$avg, sim.theta.1Mb$Rate, rep7.sim.tmrca.1Mb$Ave
names(true.lands.1Mb.rep7) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep7$thetaC <- true.lands.1Mb.rep7$theta - mean(true.lands.1Mb.rep7$theta)
true.lands.1Mb.rep7$tmrcaC <- true.lands.1Mb.rep7$tmrca - mean(true.lands.1Mb.rep7$tmrca)
true.lands.1Mb.rep7$rhoC <- true.lands.1Mb.rep7$rho - mean(true.lands.1Mb.rep7$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep7)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp
r2.bgs.1Mb.true[2, 7] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb.true[3, 7] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 7] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 7] <- anova.diversity$VarExp[4] * 100

inf.lands.1Mb.rep7 <- as.data.frame(cbind(rep_7.pi.1Mb$avg, rep_7.theta.1Mb$sample_mean, rep_7.tmrca.1M
names(inf.lands.1Mb.rep7) <- c("diversity", "theta", "tmrca", "rho")
```

```
cor.table.theta[7, 3] <- cor.test(sim.theta.1Mb$Rate, rep_7.theta.1Mb$sample_mean, method = "spearman")$
cor.table.rho[7, 3] <- cor.test(sim.rho.200kb$Rate, rep_7.rho.200kb$sample_mean, method = "spearman")$ea
cor.table.tmrca[7, 3] <- cor.test(rep7.sim.tmrca.1Mb$AverageTmrca, rep_7.tmrca.1Mb$avg, method = "spearm

# centering
inf.lands.1Mb.rep7$thetaC <- inf.lands.1Mb.rep7$theta - mean(inf.lands.1Mb.rep7$theta)
inf.lands.1Mb.rep7$tmrcaC <- inf.lands.1Mb.rep7$tmrca - mean(inf.lands.1Mb.rep7$tmrca)
inf.lands.1Mb.rep7$rhoC <- inf.lands.1Mb.rep7$rho - mean(inf.lands.1Mb.rep7$rho)

inf.lands.1Mb.rep7$bin <- 1:nrow(inf.lands.1Mb.rep7)

# for merging:
inf.lands.1Mb.rep7 <- inf.lands.1Mb.rep7
inf.lands.1Mb.rep7$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep7)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep7)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -345.4274
## m.div.1Mb.2  7 -343.4288
## m.div.1Mb.3  8 -342.2231
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
```

```
##
## data:  m.div.1Mb
## DW = 2.1895, p-value = 0.5182
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.49501, p-value = 0.49
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep7)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.052e-04 -8.794e-05 -3.079e-05  1.908e-05  3.884e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.351e-03  5.037e-05  66.526  < 2e-16 ***
## thetaC        1.086e+00  5.359e-02  20.259 2.52e-14 ***
## rhoC          2.736e-02  3.012e-02   0.908   0.3752
```

```
## tmrcaC        4.505e-03  5.883e-04   7.658 3.19e-07 ***
## thetaC:tmrcaC 1.624e+00  8.673e-01   1.873   0.0766 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001587 on 19 degrees of freedom
## Multiple R-squared:  0.9912, Adjusted R-squared:  0.9893
## F-statistic: 534.8 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 7] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 7] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 7] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 7] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 7] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep7, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep7, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep7, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep7, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -343.8782
## g.div.1Mb.2  8 -342.3197
## g.div.1Mb.3  7 -343.7752
## g.div.1Mb.4  7 -345.6709
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep7
##         AIC       BIC   logLik
##   -343.7752 -335.5288 178.8876
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##      power
## -0.6249336
##
## Coefficients:
```

```
##                   Value Std.Error  t-value p-value
## (Intercept)   0.0033351 0.0000523 63.73232  0.0000
## thetaC        1.0522911 0.0496791 21.18178  0.0000
## rhoC          0.0194373 0.0287308  0.67653  0.5069
## tmrcaC        0.0048610 0.0006046  8.04002  0.0000
## thetaC:tmrcaC 2.1186592 0.9123559  2.32219  0.0315
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.429
## rhoC           0.161 -0.143
## tmrcaC        -0.656 -0.631 -0.462
## thetaC:tmrcaC -0.770 -0.725 -0.247  0.861
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.3170115 -0.6371261 -0.1942405  0.1843185  2.7685824
##
## Residual standard error: 3.639405e-06
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      3.127304      2.123346      5.921703      4.937703
```

6.5.8 Replicate 8

```
rep8.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep8_w1000000.csv", sep = ",", header = T)

rep_8.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.div
rep_8.pi.1Mb$avg <- apply(rep_8.pi.1Mb[4:ncol(rep_8.pi.1Mb)], 1, mean)
rep_8.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.rho
rep_8.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.
rep_8.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep8/dm2L_bgs_rep8_5.
rep_8.tmrca.1Mb$avg <- apply(rep_8.tmrca.1Mb[4:ncol(rep_8.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep8 <- as.data.frame(cbind(rep_8.pi.1Mb$avg, sim.theta.1Mb$Rate, rep8.sim.tmrca.1Mb$Ave
names(true.lands.1Mb.rep8) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep8$thetaC <- true.lands.1Mb.rep8$theta - mean(true.lands.1Mb.rep8$theta)
true.lands.1Mb.rep8$tmrcaC <- true.lands.1Mb.rep8$tmrca - mean(true.lands.1Mb.rep8$tmrca)
true.lands.1Mb.rep8$rhoC <- true.lands.1Mb.rep8$rho - mean(true.lands.1Mb.rep8$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep8)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp
r2.bgs.1Mb.true[2, 8] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb.true[3, 8] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 8] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 8] <- anova.diversity$VarExp[4] * 100
```

```r
inf.lands.1Mb.rep8 <- as.data.frame(cbind(rep_8.pi.1Mb$avg, rep_8.theta.1Mb$sample_mean, rep_8.tmrca.1M
names(inf.lands.1Mb.rep8) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[8, 3] <- cor.test(sim.theta.1Mb$Rate, rep_8.theta.1Mb$sample_mean, method = "spearman")$
cor.table.rho[8, 3] <- cor.test(sim.rho.200kb$Rate, rep_8.rho.200kb$sample_mean, method = "spearman")$e
cor.table.tmrca[8, 3] <- cor.test(rep8.sim.tmrca.1Mb$AverageTmrca, rep_8.tmrca.1Mb$avg, method = "spear

# centering
inf.lands.1Mb.rep8$thetaC <- inf.lands.1Mb.rep8$theta - mean(inf.lands.1Mb.rep8$theta)
inf.lands.1Mb.rep8$tmrcaC <- inf.lands.1Mb.rep8$tmrca - mean(inf.lands.1Mb.rep8$tmrca)
inf.lands.1Mb.rep8$rhoC <- inf.lands.1Mb.rep8$rho - mean(inf.lands.1Mb.rep8$rho)

inf.lands.1Mb.rep8$bin <- 1:nrow(inf.lands.1Mb.rep8)

# for merging:
inf.lands.1Mb.rep8 <- inf.lands.1Mb.rep8
inf.lands.1Mb.rep8$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep8)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep8)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -345.1858
## m.div.1Mb.2  7 -343.4070
## m.div.1Mb.3  8 -345.1559
```

```r
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.3323, p-value = 0.6935
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.65145, p-value = 0.855
```

```r
hist(resid(m.div.1Mb))
```

```r
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep8)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.941e-04 -9.009e-05 -1.798e-05  2.516e-05  5.049e-04
##
## Coefficients:
```

```
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.428e-03  3.304e-05 103.752  < 2e-16 ***
## thetaC         1.264e+00  5.490e-02  23.029 2.42e-15 ***
## rhoC          -6.271e-03  3.593e-02  -0.175   0.8633
## tmrcaC         3.564e-03  2.162e-04  16.485 1.03e-12 ***
## thetaC:tmrcaC  8.224e-01  2.892e-01   2.844   0.0104 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001595 on 19 degrees of freedom
## Multiple R-squared:  0.9909, Adjusted R-squared:  0.989
## F-statistic: 515.7 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 8] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
r2.bgs.1Mb[2, 8] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 8] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 8] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 8] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep8, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), metl

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep8, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep8, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep8, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -344.1779
## g.div.1Mb.2  8 -346.5652
## g.div.1Mb.3  7 -346.9859
## g.div.1Mb.4  7 -343.9414
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep8
##         AIC       BIC  logLik
##   -346.9859 -338.7395 180.493
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
```

```
##    power
## 1.844674
##
## Coefficients:
##                   Value  Std.Error  t-value p-value
## (Intercept)   0.0034418 0.00003793 90.73488  0.0000
## thetaC        1.2870866 0.06260981 20.55727  0.0000
## rhoC          0.0283001 0.02906012  0.97385  0.3424
## tmrcaC        0.0033891 0.00018982 17.85469  0.0000
## thetaC:tmrcaC 1.0438448 0.28464931  3.66713  0.0016
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.665
## rhoC          -0.170 -0.687
## tmrcaC         0.167  0.532 -0.679
## thetaC:tmrcaC -0.163 -0.449  0.284 -0.114
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.4834747 -0.6354580 -0.2279145  0.1515981  2.7918383
##
## Residual standard error: 7.052248
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.256212      2.567824      1.937620      1.291249
```

6.5.9 Replicate 9

```
rep9.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep9_w1000000.csv", sep = ",", header = T)

rep_9.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.div
rep_9.pi.1Mb$avg <- apply(rep_9.pi.1Mb[4:ncol(rep_9.pi.1Mb)], 1, mean)
rep_9.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.rho
rep_9.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.
rep_9.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep9/dm2L_bgs_rep9_5.
rep_9.tmrca.1Mb$avg <- apply(rep_9.tmrca.1Mb[4:ncol(rep_9.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep9 <- as.data.frame(cbind(rep_9.pi.1Mb$avg, sim.theta.1Mb$Rate, rep9.sim.tmrca.1Mb$Ave
names(true.lands.1Mb.rep9) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep9$thetaC <- true.lands.1Mb.rep9$theta - mean(true.lands.1Mb.rep9$theta)
true.lands.1Mb.rep9$tmrcaC <- true.lands.1Mb.rep9$tmrca - mean(true.lands.1Mb.rep9$tmrca)
true.lands.1Mb.rep9$rhoC <- true.lands.1Mb.rep9$rho - mean(true.lands.1Mb.rep9$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep9)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp
r2.bgs.1Mb.true[2, 9] <- anova.diversity$VarExp[1] * 100
```

```
r2.bgs.1Mb.true[3, 9] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 9] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 9] <- anova.diversity$VarExp[4] * 100

inf.lands.1Mb.rep9 <- as.data.frame(cbind(rep_9.pi.1Mb$avg, rep_9.theta.1Mb$sample_mean, rep_9.tmrca.1M
names(inf.lands.1Mb.rep9) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[9, 3] <- cor.test(sim.theta.1Mb$Rate, rep_9.theta.1Mb$sample_mean, method = "spearman")$
cor.table.rho[9, 3] <- cor.test(sim.rho.200kb$Rate, rep_9.rho.200kb$sample_mean, method = "spearman")$ea
cor.table.tmrca[9, 3] <- cor.test(rep9.sim.tmrca.1Mb$AverageTmrca, rep_9.tmrca.1Mb$avg, method = "spear

# centering
inf.lands.1Mb.rep9$thetaC <- inf.lands.1Mb.rep9$theta - mean(inf.lands.1Mb.rep9$theta)
inf.lands.1Mb.rep9$tmrcaC <- inf.lands.1Mb.rep9$tmrca - mean(inf.lands.1Mb.rep9$tmrca)
inf.lands.1Mb.rep9$rhoC <- inf.lands.1Mb.rep9$rho - mean(inf.lands.1Mb.rep9$rho)

inf.lands.1Mb.rep9$bin <- 1:nrow(inf.lands.1Mb.rep9)

# for merging:
inf.lands.1Mb.rep9 <- inf.lands.1Mb.rep9
inf.lands.1Mb.rep9$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep9)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep9)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -344.7058
## m.div.1Mb.2  7 -345.4784
## m.div.1Mb.3  8 -344.8491
```

```
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.6206, p-value = 0.8798
## alternative hypothesis: true autocorrelation is greater than 0
```

```
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.64661, p-value = 0.856
```

```
hist(resid(m.div.1Mb))
```

```
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep9)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.942e-04 -7.503e-05 -2.451e-05  2.644e-05  5.209e-04
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.453e-03  3.751e-05  92.060  < 2e-16 ***
## thetaC         1.216e+00  5.336e-02  22.783 2.94e-15 ***
## rhoC          -6.052e-03  3.346e-02  -0.181   0.8584
## tmrcaC         3.807e-03  2.487e-04  15.304 3.87e-12 ***
## thetaC:tmrcaC  9.750e-01  3.713e-01   2.626   0.0166 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001611 on 19 degrees of freedom
## Multiple R-squared:  0.9903, Adjusted R-squared:  0.9883
## F-statistic: 486.3 on 4 and 19 DF,  p-value: < 2.2e-16
```

```
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)

r2.bgs.1Mb[1, 9] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3] +
```

```r
r2.bgs.1Mb[2, 9] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 9] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 9] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 9] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep9, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), meth

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep9, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), me

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep9, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep9, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -347.5106
## g.div.1Mb.2  8 -349.8511
## g.div.1Mb.3  7 -346.0547
## g.div.1Mb.4  7 -343.4336
```

```r
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep9
##         AIC       BIC   logLik
##   -346.0547 -337.8083 180.0274
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##    power
## 1.916343
##
## Coefficients:
##                     Value Std.Error  t-value p-value
## (Intercept)    0.0034744 0.0000413 84.07203  0.0000
## thetaC         1.2936613 0.0586924 22.04139  0.0000
## rhoC          -0.0029089 0.0268341 -0.10840  0.9148
## tmrcaC         0.0037363 0.0002296 16.27063  0.0000
## thetaC:tmrcaC  1.1021921 0.3485437  3.16228  0.0051
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.742
## rhoC          -0.235 -0.684
## tmrcaC         0.020  0.220 -0.494
## thetaC:tmrcaC -0.342 -0.372  0.181  0.448
```

```
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -1.37197717 -0.54593737 -0.36858902  0.09647638  2.59397338
##
## Residual standard error: 10.76124
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.137369      2.606907      2.189951      1.876093
```

6.5.10 Replicate 10

```
rep10.sim.tmrca.1Mb <- read.table("dm_bgs_sims/10r/dm2L_bgs_rep10_w1000000.csv", sep = ",", header = T)

rep_10.pi.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5.
rep_10.pi.1Mb$avg <- apply(rep_10.pi.1Mb[4:ncol(rep_10.pi.1Mb)], 1, mean)
rep_10.rho.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10_5
rep_10.theta.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10
rep_10.tmrca.1Mb <- read.table("dm_bgs_sims/10r/NonHomogeneous/iSMC/NonHomogeneous/rep10/dm2L_bgs_rep10
rep_10.tmrca.1Mb$avg <- apply(rep_10.tmrca.1Mb[4:ncol(rep_10.tmrca.1Mb)], 1, mean)

# first the linear modeling with the true (simulated) landscapes
true.lands.1Mb.rep10 <- as.data.frame(cbind(rep_10.pi.1Mb$avg, sim.theta.1Mb$Rate, rep10.sim.tmrca.1Mb$A
names(true.lands.1Mb.rep10) <- c("diversity", "theta", "tmrca", "rho")
true.lands.1Mb.rep10$thetaC <- true.lands.1Mb.rep10$theta - mean(true.lands.1Mb.rep10$theta)
true.lands.1Mb.rep10$tmrcaC <- true.lands.1Mb.rep10$tmrca - mean(true.lands.1Mb.rep10$tmrca)
true.lands.1Mb.rep10$rhoC <- true.lands.1Mb.rep10$rho - mean(true.lands.1Mb.rep10$rho, na.rm = T)

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = true.lands.1Mb.rep10)
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
r2.bgs.1Mb.true[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarE
r2.bgs.1Mb.true[2, 10] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb.true[3, 10] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb.true[4, 10] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb.true[5, 10] <- anova.diversity$VarExp[4] * 100

inf.lands.1Mb.rep10 <- as.data.frame(cbind(rep_10.pi.1Mb$avg, rep_10.theta.1Mb$sample_mean, rep_10.tmrca
names(inf.lands.1Mb.rep10) <- c("diversity", "theta", "tmrca", "rho")

cor.table.theta[10, 3] <- cor.test(sim.theta.1Mb$Rate, rep_10.theta.1Mb$sample_mean, method = "spearman
cor.table.rho[10, 3] <- cor.test(sim.rho.200kb$Rate, rep_10.rho.200kb$sample_mean, method = "spearman")$
cor.table.tmrca[10, 3] <- cor.test(rep10.sim.tmrca.1Mb$AverageTmrca, rep_10.tmrca.1Mb$avg, method = "sp

# centering
inf.lands.1Mb.rep10$thetaC <- inf.lands.1Mb.rep10$theta - mean(inf.lands.1Mb.rep10$theta)
inf.lands.1Mb.rep10$tmrcaC <- inf.lands.1Mb.rep10$tmrca - mean(inf.lands.1Mb.rep10$tmrca)
inf.lands.1Mb.rep10$rhoC <- inf.lands.1Mb.rep10$rho - mean(inf.lands.1Mb.rep10$rho)

inf.lands.1Mb.rep10$bin <- 1:nrow(inf.lands.1Mb.rep10)
```

```r
# for merging:
inf.lands.1Mb.rep10 <- inf.lands.1Mb.rep10
inf.lands.1Mb.rep10$Replicate <- 1

m.div.1Mb <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC, data = inf.lands.1Mb.rep10)
m.div.1Mb.2 <- lm(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC + rhoC:tmrcaC, data = inf.lands.1M
m.div.1Mb.3 <- lm(diversity ~ (thetaC + rhoC + tmrcaC) ^ 2, data = inf.lands.1Mb.rep10)

AIC(m.div.1Mb, m.div.1Mb.2, m.div.1Mb.3)
```

```
##             df       AIC
## m.div.1Mb    6 -349.8858
## m.div.1Mb.2  7 -353.7635
## m.div.1Mb.3  8 -351.7726
```

```r
plot(resid(m.div.1Mb)~fitted(m.div.1Mb))
```

```r
dwtest(m.div.1Mb)
```

```
##
##  Durbin-Watson test
##
## data:  m.div.1Mb
## DW = 2.0448, p-value = 0.4006
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
hmctest(m.div.1Mb)
```

```
##
##  Harrison-McCabe test
##
## data:  m.div.1Mb
## HMC = 0.61476, p-value = 0.788
```

```r
hist(resid(m.div.1Mb))
```

```r
summary(m.div.1Mb)
```

```
##
## Call:
## lm(formula = diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
##     data = inf.lands.1Mb.rep10)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.796e-04 -9.261e-05 -2.968e-05  6.834e-05  3.971e-04
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.457e-03  3.521e-05  98.178  < 2e-16 ***
## thetaC        1.015e+00  4.199e-02  24.171 9.92e-16 ***
## rhoC          4.850e-02  3.237e-02   1.498   0.1505
## tmrcaC        4.783e-03  3.222e-04  14.844 6.62e-12 ***
## thetaC:tmrcaC 1.310e+00  5.763e-01   2.273   0.0348 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0001446 on 19 degrees of freedom
## Multiple R-squared:  0.9925, Adjusted R-squared:  0.9909
## F-statistic: 624.7 on 4 and 19 DF,  p-value: < 2.2e-16
```

```r
anova.diversity <- Anova(m.div.1Mb)
apiss <- anova.diversity$"Sum Sq"
anova.diversity$VarExp <- apiss / sum(apiss)
```

```
r2.bgs.1Mb[1, 10] <- (anova.diversity$VarExp[1] + anova.diversity$VarExp[2] + anova.diversity$VarExp[3]
r2.bgs.1Mb[2, 10] <- anova.diversity$VarExp[1] * 100
r2.bgs.1Mb[3, 10] <- anova.diversity$VarExp[2] * 100
r2.bgs.1Mb[4, 10] <- anova.diversity$VarExp[3] * 100
r2.bgs.1Mb[5, 10] <- anova.diversity$VarExp[4] * 100

g.div.1Mb.1 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                   data = inf.lands.1Mb.rep10, weights = varPower(0, ~tmrcaC), cor = corAR1(0, ~bin), me

g.div.1Mb.2 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep10, weights = varPower(0, ~theta), cor = corAR1(0, ~bin), m

g.div.1Mb.3 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep10, weights = varPower(0, ~theta), method = "ML")

g.div.1Mb.4 <- gls(diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC,
                     data = inf.lands.1Mb.rep10, weights = varPower(0, ~tmrcaC), method = "ML")

AIC(g.div.1Mb.1, g.div.1Mb.2, g.div.1Mb.3, g.div.1Mb.4)
```

```
##             df       AIC
## g.div.1Mb.1  8 -347.2457
## g.div.1Mb.2  8 -348.1811
## g.div.1Mb.3  7 -348.8246
## g.div.1Mb.4  7 -349.0469
```

```
summary(g.div.1Mb.3)
```

```
## Generalized least squares fit by maximum likelihood
##   Model: diversity ~ thetaC + rhoC + tmrcaC + thetaC:tmrcaC
##   Data: inf.lands.1Mb.rep10
##         AIC       BIC   logLik
##   -348.8246 -340.5782 181.4123
##
## Variance function:
##  Structure: Power of variance covariate
##  Formula: ~theta
##  Parameter estimates:
##     power
## 0.7454269
##
## Coefficients:
##                    Value Std.Error  t-value p-value
## (Intercept)   0.0034586 0.0000353 97.87553  0.0000
## thetaC        1.0353207 0.0442649 23.38920  0.0000
## rhoC          0.0439740 0.0290366  1.51444  0.1464
## tmrcaC        0.0047829 0.0003101 15.42616  0.0000
## thetaC:tmrcaC 1.3035906 0.5316560  2.45194  0.0241
##
##  Correlation:
##               (Intr) thetaC rhoC   tmrcaC
## thetaC         0.520
## rhoC          -0.109 -0.611
## tmrcaC        -0.182  0.060 -0.527
```

```
## thetaC:tmrcaC -0.498 -0.504  0.190  0.390
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.3571009 -0.7800013 -0.1949018  0.4014887  2.8468855
##
## Residual standard error: 0.009233511
## Degrees of freedom: 24 total; 19 residual
```

```
vif(g.div.1Mb.3)
```

```
##        thetaC          rhoC        tmrcaC thetaC:tmrcaC
##      2.161747      2.693198      2.158165      1.824506
```

6.6 all replicates:

```
theta.plot <- as.data.frame(cbind(1:nrow(sim.theta.1Mb),
                                  sim.theta.1Mb$Rate * 1e-7 * 3e+4,
                                  rep_1.theta.1Mb$sample_mean,
                                  rep_2.theta.1Mb$sample_mean,
                                  rep_3.theta.1Mb$sample_mean,
                                  rep_4.theta.1Mb$sample_mean,
                                  rep_5.theta.1Mb$sample_mean,
                                  rep_6.theta.1Mb$sample_mean,
                                  rep_7.theta.1Mb$sample_mean,
                                  rep_8.theta.1Mb$sample_mean,
                                  rep_9.theta.1Mb$sample_mean,
                                  rep_10.theta.1Mb$sample_mean))

names(theta.plot) <- c("bin", "sim", reps)
molten.theta <- melt(theta.plot, id.vars = "bin")
theta.map.1Mb.bgs <- ggplot(data = molten.theta, aes(x = bin * 1000, y = value, colour = variable)) + tl
theta.map.1Mb.bgs <- theta.map.1Mb.bgs + geom_line(data = molten.theta, aes(size = variable)) + scale_si
theta.map.1Mb.bgs <- theta.map.1Mb.bgs + scale_color_manual(values = c("black", brewer.pal(n = 9, name =
theta.map.1Mb.bgs <- theta.map.1Mb.bgs + scale_x_continuous(breaks = pretty_breaks()) + scale_y_continuo
theta.map.1Mb.bgs <- theta.map.1Mb.bgs + labs(title = NULL, x = "Position (kb)", y = NULL)
theta.map.1Mb.bgs <- theta.map.1Mb.bgs + theme(text = element_text(size = 20), axis.title.x = element_te
theta.map.1Mb.bgs
```

7 Plots

```
write.table(cor.table.theta, "cor.mutation.bgs.tsv", quote = F, sep = "\t")
write.table(cor.table.tmrca, "cor.tmrca.bgs.tsv", quote = F, sep = "\t")
write.table(cor.table.rho, "cor.rho.bgs.tsv", quote = F, sep = "\t")

cor.table.comb <- rbind.data.frame(t(cor.table.theta), t(cor.table.rho), t(cor.table.tmrca))
cor.table.comb$bin.size <- rep(c(50, 200, 1000), 3)
cor.table.comb$rate <- c(rep("Theta", 3), rep("Rho", 3), rep("TMRCA", 3))

molten.cor <- melt(cor.table.comb, id.vars = c("bin.size", "rate"))
cor.bgs.plot <- ggplot(data = molten.cor, aes(x = bin.size, y = value))
cor.bgs.plot <- cor.bgs.plot + geom_point(size=3, alpha = 0.6) + facet_wrap(~rate)
cor.bgs.plot <- cor.bgs.plot + scale_x_continuous(breaks = c(50, 200, 1000), trans="log10")
cor.bgs.plot <- cor.bgs.plot + scale_y_continuous(breaks = pretty_breaks())
cor.bgs.plot <- cor.bgs.plot + labs(title = NULL, x = "Bin Size (kb)", y = "Spearman's Rho")
cor.bgs.plot <- cor.bgs.plot + theme_bw() + ylim(0, 1)
cor.bgs.plot <- cor.bgs.plot + theme(axis.title = element_text(size = 20), axis.text = element_text(size
                                     legend.position = "bottom")
cor.bgs.plot
```

```
ggsave("Figure4.pdf", cor.bgs.plot, device = "pdf", height = 5, width = 12)

# true (simulated) landscapes
r2.bgs.50kb.true$average <- apply(r2.bgs.50kb.true, 1, mean)
r2.bgs.200kb.true$average <- apply(r2.bgs.200kb.true, 1, mean)
r2.bgs.1Mb.true$average <- apply(r2.bgs.1Mb.true, 1, mean)

r2.true.avg <- rbind.data.frame(r2.bgs.50kb.true$average, r2.bgs.200kb.true$average, r2.bgs.1Mb.true$ave
colnames(r2.true.avg) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
r2.true.avg$bin.size <- c(50, 200, 1000)

# inferred landscapes
r2.bgs.50kb$average <- apply(r2.bgs.50kb, 1, mean)
r2.bgs.200kb$average <- apply(r2.bgs.200kb, 1, mean)
r2.bgs.1Mb$average <- apply(r2.bgs.1Mb, 1, mean)

r2.inf.avg <- rbind.data.frame(r2.bgs.50kb$average, r2.bgs.200kb$average, r2.bgs.1Mb$average, make.row.n
colnames(r2.inf.avg) <- c("Total", "Theta", "Rho", "TMRCA", "Theta:TMRCA")
r2.inf.avg$bin.size <- c(50, 200, 1000)

r2.tab.comb <- rbind.data.frame(r2.dm.tab, r2.true.avg, r2.inf.avg)
r2.tab.comb$type <- c(rep("iSMC-inferred (Drosophila)", 3), rep("True landscapes (simulations)", 3), rep

molten.r2 <- melt(r2.tab.comb, id.vars = c("bin.size", "type"))
r2.bgs.plot <- ggplot(data = molten.r2, aes(x = bin.size, y = value, shape = variable))
r2.bgs.plot <- r2.bgs.plot + geom_line(data = molten.r2)
r2.bgs.plot <- r2.bgs.plot + geom_point(aes(shape = variable), size = 4) + facet_wrap(~type)
```

```
r2.bgs.plot <- r2.bgs.plot + scale_x_continuous(breaks = c(50, 200, 1000), trans="log10")
r2.bgs.plot <- r2.bgs.plot + scale_y_continuous(breaks = pretty_breaks())
r2.bgs.plot <- r2.bgs.plot + scale_shape_manual(values = c(0, 1, 2, 3, 4))
r2.bgs.plot <- r2.bgs.plot + labs(title = NULL, x = "Bin Size (kb)", y = "Variance Explained (%)")
r2.bgs.plot <- r2.bgs.plot + theme_bw()
r2.bgs.plot <- r2.bgs.plot + theme(axis.title = element_text(size = 20), axis.text = element_text(size =
                                   legend.position = "bottom")
r2.bgs.plot
```

```
# combine coal and BGS R2 plots
r2.comb.plot <- plot_grid(r2.coal.plot, r2.bgs.plot, nrow = 2, ncol = 1, labels = "AUTO")
cowplot::save_plot("Figure5.pdf", r2.comb.plot, device = "pdf", base_height = 9, base_width = 12)

# bench theta
theta.maps.bgs <- plot_grid(theta.map.50kb.bgs, theta.map.200kb.bgs, theta.map.1Mb.bgs, nrow = 3, ncol =

theta.maps.comb <- plot_grid(theta.maps.coal, theta.maps.bgs, labels = "AUTO", nrow = 1, ncol = 2)
save_plot("Figure2.pdf", theta.maps.comb, base_height = 16, base_width = 12)

# usualise true join landscapes
dm_2L_exome <- read.table("dm_bgs_sims/GeneAnnotations/exons_chr2L.csv", header = F, sep = ",")
names(dm_2L_exome) <- c("exon_start", "exon_end")

genes_barcode <- ggplot(data = dm_2L_exome) + geom_segment(aes(x = exon_start, xend = exon_end, size = 1

## 50kb
theta.land.50kb <- cbind.data.frame(1:nrow(sim.theta.50kb), sim.theta.50kb$Rate)
```

```
names(theta.land.50kb) <- c("bin", "rate")

rho.land.50kb <- cbind.data.frame(1:nrow(sim.rho.50kb), sim.rho.50kb$Rate)
names(rho.land.50kb) <- c("bin", "rate")

pi.lands.50kb <- cbind.data.frame(1:nrow(sim.theta.50kb), rep_1.pi.50kb$avg, rep_2.pi.50kb$avg, rep_3.pi
                                  rep_6.pi.50kb$avg, rep_7.pi.50kb$avg, rep_8.pi.50kb$avg, rep_9.pi.50kb
names(pi.lands.50kb) <- c("bin", reps)


tau.lands.50kb <- cbind.data.frame(1:nrow(sim.theta.50kb), rep1.sim.tmrca.50kb$AverageTmrca, rep2.sim.tm
                                   rep4.sim.tmrca.50kb$AverageTmrca, rep5.sim.tmrca.50kb$AverageTmrca, re
                                   rep7.sim.tmrca.50kb$AverageTmrca, rep8.sim.tmrca.50kb$AverageTmrca, re
                                   rep10.sim.tmrca.50kb$AverageTmrca)
names(tau.lands.50kb) <- c("bin", reps)

rho.map.50kb <- ggplot(data = rho.land.50kb, aes(x = bin * 5e+4, y = rate))
rho.map.50kb <- rho.map.50kb + geom_line(data = rho.land.50kb) + theme_bw()
rho.map.50kb <- rho.map.50kb + scale_y_continuous(breaks = pretty_breaks(), labels = scale.3d)
rho.map.50kb <- rho.map.50kb + labs(title = NULL, x = NULL, y = expression(rho))
rho.map.50kb <- rho.map.50kb + theme(axis.title.x=element_blank(),
                                     axis.text.x=element_blank(),
                                     axis.ticks.x=element_blank(),
                                     axis.title.y=element_text(size = 20), legend.position = "none")
rho.map.50kb
```

```
theta.map.50kb <- ggplot(data = theta.land.50kb, aes(x = bin * 5e+4, y = rate))
theta.map.50kb <- theta.map.50kb + geom_line(data = theta.land.50kb) + theme_bw()
theta.map.50kb <- theta.map.50kb + scale_y_continuous(breaks = pretty_breaks(), labels = scale.3d)
theta.map.50kb <- theta.map.50kb + labs(title = NULL, x = NULL, y = expression(theta))
theta.map.50kb <- theta.map.50kb + geom_segment(data = dm_2L_exome, aes(x = exon_start, xend = exon_end
theta.map.50kb <- theta.map.50kb + theme(axis.title.x=element_blank(),
                                         axis.text.x=element_blank(),
                                         axis.ticks.x=element_blank(),
                                         axis.title.y=element_text(size = 20), legend.position = "none")
theta.map.50kb
```

```
molten.tau <- melt(tau.lands.50kb, id.vars = "bin")
tau.map.50kb <- ggplot(data = molten.tau, aes(x = bin * 5e+4, y = value, colour = variable))
tau.map.50kb <- tau.map.50kb + geom_line(data = molten.tau) + theme_bw()
tau.map.50kb <- tau.map.50kb + scale_y_continuous(breaks = pretty_breaks())
tau.map.50kb <- tau.map.50kb + labs(title = NULL, x = NULL, y = expression(tau))
tau.map.50kb <- tau.map.50kb + scale_color_manual(values = c(brewer.pal(n = 9, name = "Greys"), "Black"
tau.map.50kb <- tau.map.50kb + theme(axis.title.x=element_blank(),
                                     axis.text.x=element_blank(),
                                     axis.ticks.x=element_blank(),
                                     axis.title.y=element_text(size = 20), legend.position = "none")
tau.map.50kb
```

```
molten.pi <- melt(pi.lands.50kb, id.vars = "bin")
pi.map.50kb <- ggplot(data = molten.pi, aes(x = bin * 5e+4, y = value, colour = variable))
pi.map.50kb <- pi.map.50kb + geom_line(data = molten.pi) + theme_bw()
pi.map.50kb <- pi.map.50kb + scale_y_continuous(breaks = pretty_breaks())
pi.map.50kb <- pi.map.50kb + labs(title = NULL, x = NULL, y = expression(pi))
pi.map.50kb <- pi.map.50kb + scale_color_manual(values = c(brewer.pal(n = 9, name = "Greys"), "Black"))
pi.map.50kb <- pi.map.50kb + theme(axis.title.x=element_blank(),
                                   axis.text.x=element_blank(),
                                   axis.ticks.x=element_blank(),
                                   axis.title.y=element_text(size = 20), legend.position = "none")
pi.map.50kb
```

```
true.maps.50kb <- plot_grid(pi.map.50kb, tau.map.50kb, rho.map.50kb, theta.map.50kb, ncol = 1)
true.maps.50kb
```

```
## 200kb
theta.land.200kb <- cbind.data.frame(1:nrow(sim.theta.200kb), sim.theta.200kb$Rate)
names(theta.land.200kb) <- c("bin", "rate")

rho.land.200kb <- cbind.data.frame(1:nrow(sim.rho.200kb), sim.rho.200kb$Rate)
names(rho.land.200kb) <- c("bin", "rate")

pi.lands.200kb <- cbind.data.frame(1:nrow(sim.theta.200kb), rep_1.pi.200kb$avg, rep_2.pi.200kb$avg, rep_
                                   rep_6.pi.200kb$avg, rep_7.pi.200kb$avg, rep_8.pi.200kb$avg, rep_9.pi.20
names(pi.lands.200kb) <- c("bin", reps)


tau.lands.200kb <- cbind.data.frame(1:nrow(sim.theta.200kb), rep1.sim.tmrca.200kb$AverageTmrca, rep2.sim
                                    rep4.sim.tmrca.200kb$AverageTmrca, rep5.sim.tmrca.200kb$AverageTmrca,
                                    rep7.sim.tmrca.200kb$AverageTmrca, rep8.sim.tmrca.200kb$AverageTmrca,
                                    rep10.sim.tmrca.200kb$AverageTmrca)
names(tau.lands.200kb) <- c("bin", reps)


rho.map.200kb <- ggplot(data = rho.land.200kb, aes(x = bin * 2e+5, y = rate))
rho.map.200kb <- rho.map.200kb + geom_line(data = rho.land.200kb) + theme_bw()
rho.map.200kb <- rho.map.200kb + scale_y_continuous(breaks = pretty_breaks(), labels = scale.3d)
rho.map.200kb <- rho.map.200kb + labs(title = NULL, x = NULL, y = NULL)
rho.map.200kb <- rho.map.200kb + theme(axis.title.x=element_blank(),
                                       axis.text.x=element_blank(),
                                       axis.ticks.x=element_blank(),
                                       axis.title.y=element_text(size = 20), legend.position = "none")
```

```
rho.map.200kb
```

```
theta.map.200kb <- ggplot(data = theta.land.200kb, aes(x = bin * 2e+5, y = rate))
theta.map.200kb <- theta.map.200kb + geom_line(data = theta.land.200kb) + theme_bw()
theta.map.200kb <- theta.map.200kb + scale_y_continuous(breaks = pretty_breaks(), labels = scale.3d)
theta.map.200kb <- theta.map.200kb + labs(title = NULL, x = NULL, y = NULL)
theta.map.200kb <- theta.map.200kb + geom_segment(data = dm_2L_exome, aes(x = exon_start, xend = exon_e
theta.map.200kb <- theta.map.200kb + theme(axis.title.x=element_blank(),
                                           axis.text.x=element_blank(),
                                           axis.ticks.x=element_blank(),
                                           axis.title.y=element_text(size = 20), legend.position = "none")
theta.map.200kb
```

```
molten.tau <- melt(tau.lands.200kb, id.vars = "bin")
tau.map.200kb <- ggplot(data = molten.tau, aes(x = bin * 2e+5, y = value, colour = variable))
tau.map.200kb <- tau.map.200kb + geom_line(data = molten.tau) + theme_bw()
tau.map.200kb <- tau.map.200kb + scale_y_continuous(breaks = pretty_breaks())
tau.map.200kb <- tau.map.200kb + labs(title = NULL, x = NULL, y = NULL)
tau.map.200kb <- tau.map.200kb + scale_color_manual(values = c(brewer.pal(n = 9, name = "Greys"), "Blac
tau.map.200kb <- tau.map.200kb + theme(axis.title.x=element_blank(),
                                       axis.text.x=element_blank(),
                                       axis.ticks.x=element_blank(),
                                       axis.title.y=element_text(size = 20), legend.position = "none")
tau.map.200kb
```

```
molten.pi <- melt(pi.lands.200kb, id.vars = "bin")
pi.map.200kb <- ggplot(data = molten.pi, aes(x = bin * 2e+5, y = value, colour = variable))
pi.map.200kb <- pi.map.200kb + geom_line(data = molten.pi) + theme_bw()
pi.map.200kb <- pi.map.200kb + scale_y_continuous(breaks = pretty_breaks())
pi.map.200kb <- pi.map.200kb + labs(title = NULL, x = NULL, y = NULL)
pi.map.200kb <- pi.map.200kb + scale_color_manual(values = c(brewer.pal(n = 9, name = "Greys"), "Black"
pi.map.200kb <- pi.map.200kb + theme(axis.title.x=element_blank(),
                                     axis.text.x=element_blank(),
                                     axis.ticks.x=element_blank(),
                                     axis.title.y=element_text(size = 20), legend.position = "none")
pi.map.200kb
```

```
true.maps.200kb <- plot_grid(pi.map.200kb, tau.map.200kb, rho.map.200kb, theta.map.200kb, ncol = 1)
true.maps.200kb
```

```
## 1Mb
theta.land.1Mb <- cbind.data.frame(1:nrow(sim.theta.1Mb), sim.theta.1Mb$Rate)
names(theta.land.1Mb) <- c("bin", "rate")

rho.land.1Mb <- cbind.data.frame(1:nrow(sim.rho.1Mb), sim.rho.1Mb$Rate)
names(rho.land.1Mb) <- c("bin", "rate")

pi.lands.1Mb <- cbind.data.frame(1:nrow(sim.theta.1Mb), rep_1.pi.1Mb$avg, rep_2.pi.1Mb$avg, rep_3.pi.1M
                                 rep_6.pi.1Mb$avg, rep_7.pi.1Mb$avg, rep_8.pi.1Mb$avg, rep_9.pi.1Mb$avg
names(pi.lands.1Mb) <- c("bin", reps)


tau.lands.1Mb <- cbind.data.frame(1:nrow(sim.theta.1Mb), rep1.sim.tmrca.1Mb$AverageTmrca, rep2.sim.tmrca
                                  rep4.sim.tmrca.1Mb$AverageTmrca, rep5.sim.tmrca.1Mb$AverageTmrca, rep
                                  rep8.sim.tmrca.1Mb$AverageTmrca, rep9.sim.tmrca.1Mb$AverageTmrca, rep
names(tau.lands.1Mb) <- c("bin", reps)


rho.map.1Mb <- ggplot(data = rho.land.1Mb, aes(x = bin * 1e+6, y = rate))
rho.map.1Mb <- rho.map.1Mb + geom_line(data = rho.land.1Mb)  + theme_bw()
rho.map.1Mb <- rho.map.1Mb + scale_y_continuous(breaks = pretty_breaks(), labels = scale.3d)
rho.map.1Mb <- rho.map.1Mb + labs(title = NULL, x = NULL, y = NULL)
rho.map.1Mb <- rho.map.1Mb + theme(axis.title.x=element_blank(),
                                   axis.text.x=element_blank(),
                                   axis.ticks.x=element_blank(),
                                   axis.title.y=element_text(size = 20), legend.position = "none")
rho.map.1Mb
```

```
theta.map.1Mb <- ggplot(data = theta.land.1Mb, aes(x = bin * 1e+6, y = rate))
theta.map.1Mb <- theta.map.1Mb + geom_line(data = theta.land.1Mb)  + theme_bw()
theta.map.1Mb <- theta.map.1Mb + scale_y_continuous(breaks = pretty_breaks(), labels = scale.3d)
theta.map.1Mb <- theta.map.1Mb + labs(title = NULL, x = NULL, y = NULL)
theta.map.1Mb <- theta.map.1Mb + geom_segment(data = dm_2L_exome, aes(x = exon_start, xend = exon_end, y
theta.map.1Mb <- theta.map.1Mb + theme(axis.title.x=element_blank(),
                                       axis.text.x=element_blank(),
                                       axis.ticks.x=element_blank(),
                                       axis.title.y=element_text(size = 20), legend.position = "none")
theta.map.1Mb
```

```
molten.tau <- melt(tau.lands.1Mb, id.vars = "bin")
tau.map.1Mb <- ggplot(data = molten.tau, aes(x = bin * 1e+6, y = value, colour = variable))
tau.map.1Mb <- tau.map.1Mb + geom_line(data = molten.tau) + theme_bw()
tau.map.1Mb <- tau.map.1Mb + scale_y_continuous(breaks = pretty_breaks())
tau.map.1Mb <- tau.map.1Mb + labs(title = NULL, x = NULL, y = NULL)
tau.map.1Mb <- tau.map.1Mb + scale_color_manual(values = c(brewer.pal(n = 9, name = "Greys"), "Black"))
tau.map.1Mb <- tau.map.1Mb + theme(axis.title.x=element_blank(),
                                   axis.text.x=element_blank(),
                                   axis.ticks.x=element_blank(),
                                   axis.title.y=element_text(size = 20), legend.position = "none")
tau.map.1Mb
```

```
molten.pi <- melt(pi.lands.1Mb, id.vars = "bin")
pi.map.1Mb <- ggplot(data = molten.pi, aes(x = bin * 1e+6, y = value, colour = variable))
pi.map.1Mb <- pi.map.1Mb + geom_line(data = molten.pi) + theme_bw()
pi.map.1Mb <- pi.map.1Mb + scale_y_continuous(breaks = pretty_breaks())
pi.map.1Mb <- pi.map.1Mb + labs(title = NULL, x = NULL, y = NULL)
pi.map.1Mb <- pi.map.1Mb + scale_color_manual(values = c(brewer.pal(n = 9, name = "Greys"), "Black"))
pi.map.1Mb <- pi.map.1Mb + theme(axis.title.x=element_blank(),
                                 axis.text.x=element_blank(),
                                 axis.ticks.x=element_blank(),
                                 axis.title.y=element_text(size = 20), legend.position = "none")
pi.map.1Mb
```

```
true.maps.1Mb <- plot_grid(pi.map.1Mb, tau.map.1Mb, rho.map.1Mb, theta.map.1Mb, ncol = 1)
true.maps.1Mb
```

```
true.maps <- plot_grid(true.maps.50kb, true.maps.200kb, true.maps.1Mb, labels = "AUTO", ncol = 3)
true.maps
```

```
save_plot("Figure7.pdf", true.maps, base_height = 10, base_width = 20)
```

```
mean(apply(tau.lands.50kb[2:ncol(tau.lands.50kb)], 2, sd) / apply(tau.lands.50kb[2:ncol(tau.lands.50kb)
```

```
## [1] 0.3626566
```

```
mean(apply(tau.lands.200kb[2:ncol(tau.lands.200kb)], 2, sd) / apply(tau.lands.200kb[2:ncol(tau.lands.20
```

```
## [1] 0.3352754
```

```
mean(apply(tau.lands.1Mb[2:ncol(tau.lands.1Mb)], 2, sd) / apply(tau.lands.1Mb[2:ncol(tau.lands.1Mb)], 2
```

```
## [1] 0.3202002
```

```
sd(apply(tau.lands.50kb[2:ncol(tau.lands.50kb)], 2, sd) / apply(tau.lands.50kb[2:ncol(tau.lands.50kb)],
```

```
## [1] 0.007548117
```

```
sd(apply(tau.lands.200kb[2:ncol(tau.lands.200kb)], 2, sd) / apply(tau.lands.200kb[2:ncol(tau.lands.200k
```

```
## [1] 0.006994296
```

```
sd(apply(tau.lands.1Mb[2:ncol(tau.lands.1Mb)], 2, sd) / apply(tau.lands.1Mb[2:ncol(tau.lands.1Mb)], 2, n
```

```
## [1] 0.009614134
```

```
mean(apply(pi.lands.50kb[2:ncol(pi.lands.50kb)], 2, sd) / apply(pi.lands.50kb[2:ncol(pi.lands.50kb)], 2
```

```
## [1] 0.715368
```

```
mean(apply(pi.lands.200kb[2:ncol(pi.lands.200kb)], 2, sd) / apply(pi.lands.200kb[2:ncol(pi.lands.200kb)
```

```
## [1] 0.6213829
```

```
mean(apply(pi.lands.1Mb[2:ncol(pi.lands.1Mb)], 2, sd) / apply(pi.lands.1Mb[2:ncol(pi.lands.1Mb)], 2, mea
```

```
## [1] 0.437516
```

```
sd(apply(pi.lands.50kb[2:ncol(pi.lands.50kb)], 2, sd) / apply(pi.lands.50kb[2:ncol(pi.lands.50kb)], 2, m
```

```
## [1] 0.005943989
```

```
sd(apply(pi.lands.200kb[2:ncol(pi.lands.200kb)], 2, sd) / apply(pi.lands.200kb[2:ncol(pi.lands.200kb)],
```

```
## [1] 0.006465506
```

```
sd(apply(pi.lands.1Mb[2:ncol(pi.lands.1Mb)], 2, sd) / apply(pi.lands.1Mb[2:ncol(pi.lands.1Mb)], 2, mean
```

```
## [1] 0.01168644
```