Science Advances
AAAS

Supplementary Materials for

Dynamics of sex chromosome evolution in a rapid radiation of cichlid fishes

Athimed El Taher, Fabrizia Ronco, Michael Matschiner, Walter Salzburger, Astrid Böhne*

*Corresponding author. Email: a.boehne@leibniz-zfmk.de

Published 1 September 2021, *Sci. Adv.* **7**, eabe8215 (2021)
DOI: 10.1126/sciadv.abe8215

The PDF file includes:

Figs. S1 to S11

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S6

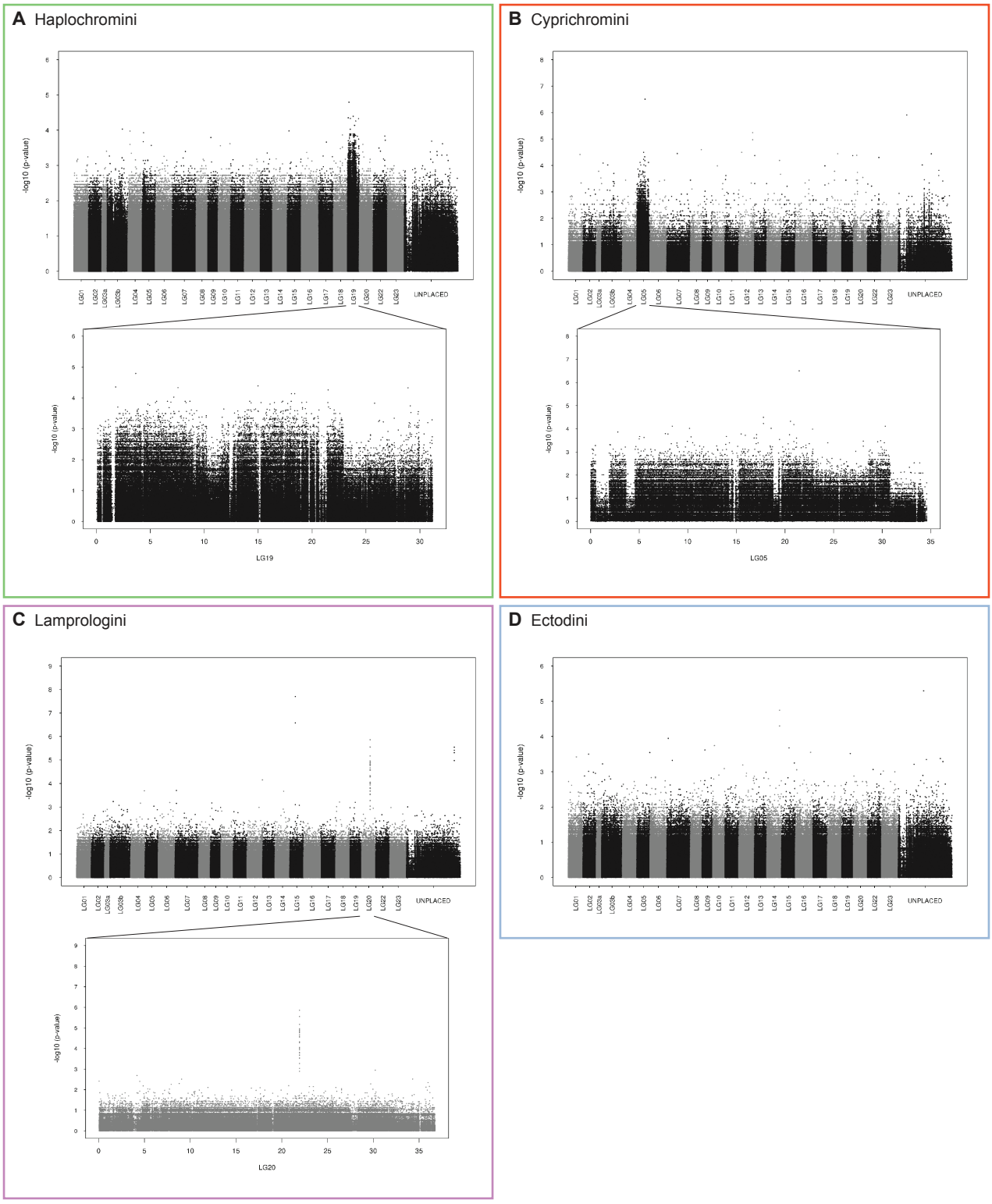


Fig. S1. Results of tribe-wise GWAS for an association with sex (approach 1). Manhattan plots of a GWAS analysis for an association with sex for the tribes Haplochromini (**A**), Cyprichromini (**B**), Lamprologini (**C**), and Ectodini (**D**) using the reference genome of the Nile tilapia (*Oreochromis niloticus*). Interchanging black and grey colors delimit chromosomes. Unplaced scaffolds were concatenated into an "UNPLACED" chromosome for visualisation. Insets show zooms of regions with an accumulation of SNPs that show an association with sex, i.e. sex chromosomes.

Fig. S2

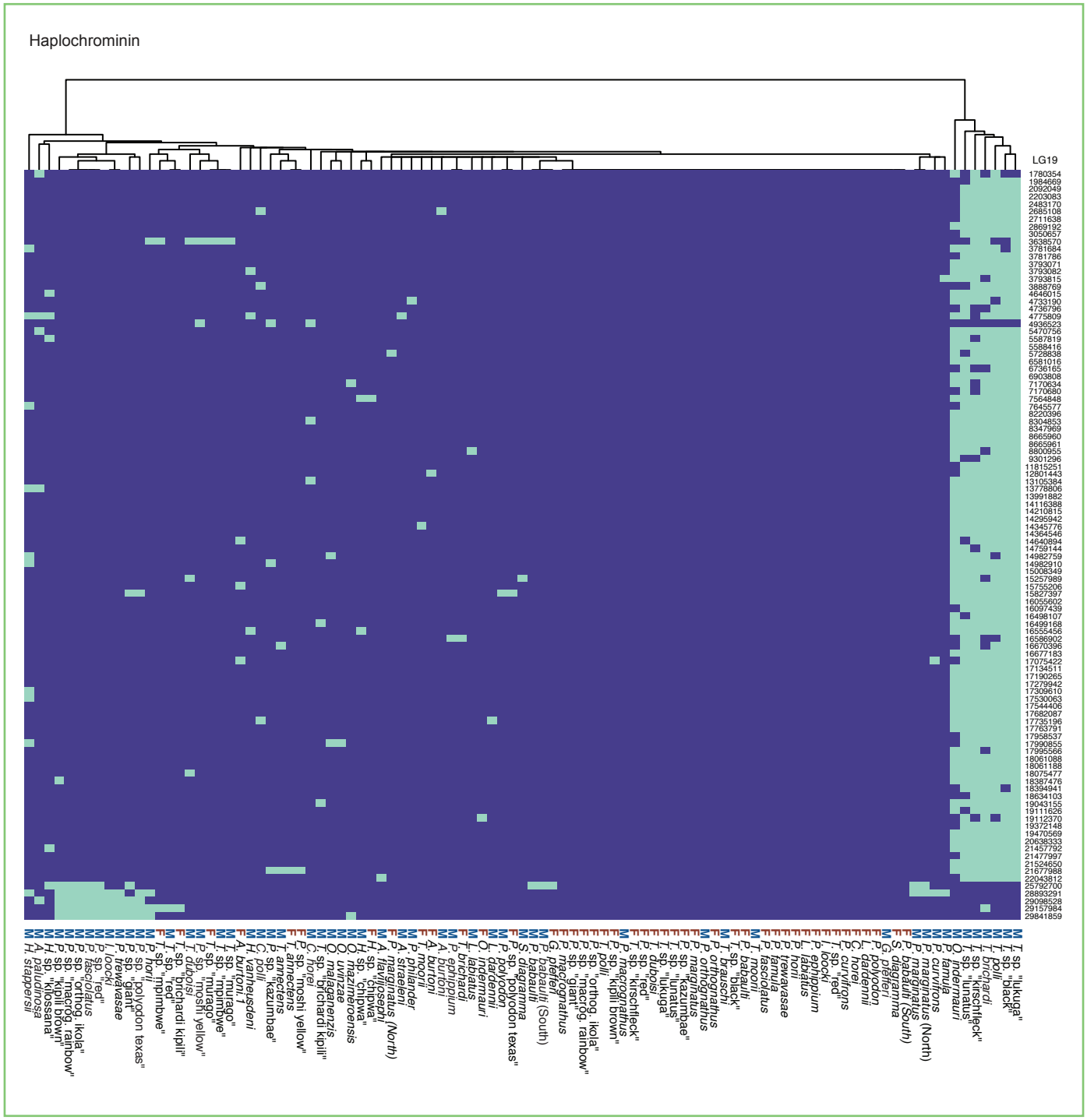
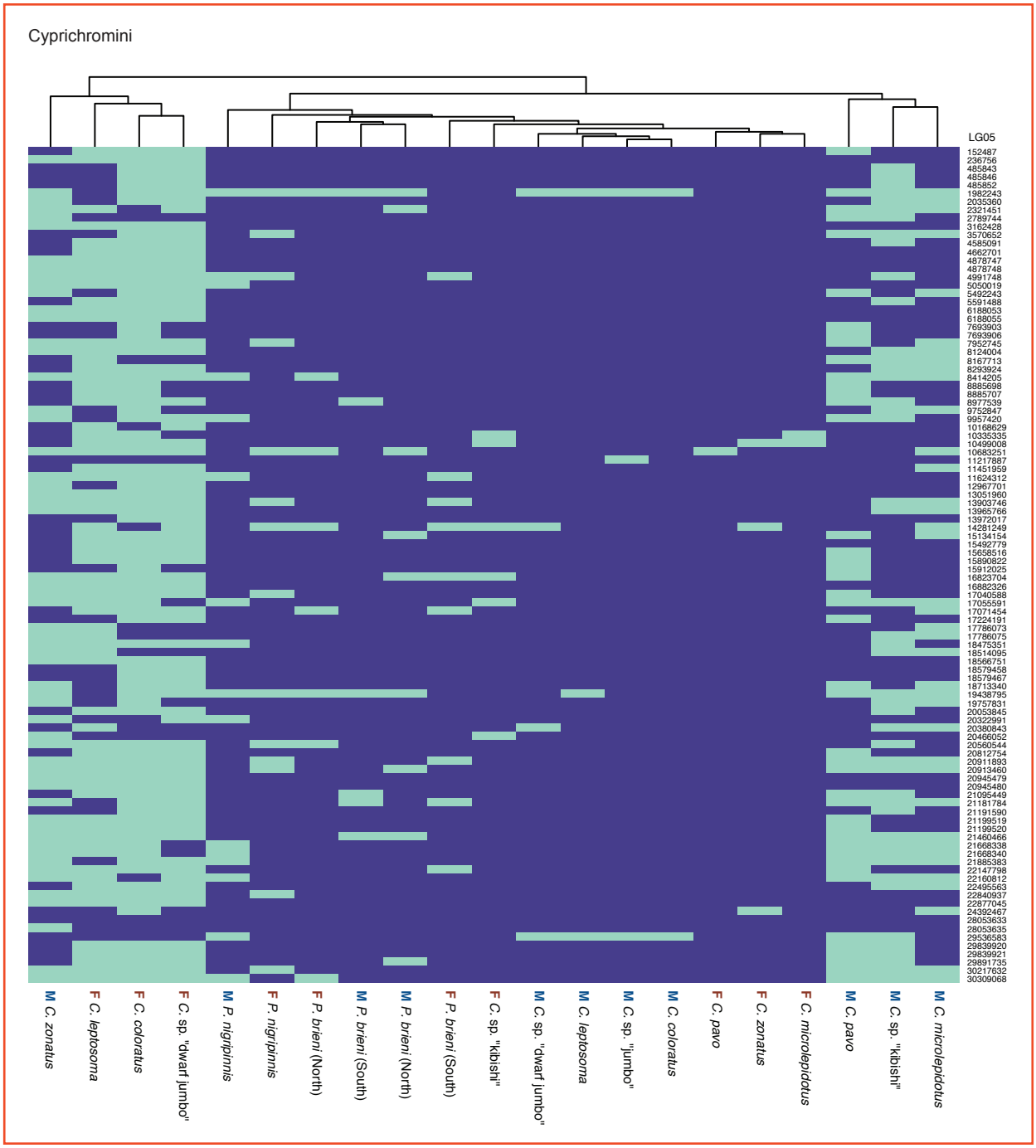

Haplochrominin
LG19
1780354
1984669
2092049
2203083
2483170
2685108
2711638
2869192
3050657
3638570
3781884
3781786
3793071
3793082
3793815
3888769
4646015
4733190
4736796
4775809
4936523
5470756
5587819
5588416
5728838
6581016
6736165
6903808
7170634
7170680
7564848
7645577
8220396
8304853
8347969
8665960
8665961
8800955
9301296
11815251
12801443
13105384
13778806
13991882
14116388
14210815
14295942
14345776
14364546
14640894
14759144
14982759
14982910
15008349
15257989
15755206
15827397
16055602
16097439
16498107
16499168
16555456
16586902
16670396
16677183
17075422
17134511
17190265
17279942
17309610
17530063
17544406
17682087
17735196
17763791
17958537
17990855
17995566
18061088
18061188
18075477
18387476
18394941
18634103
19043155
19111626
19112370
19372148
19470569
20638333
21457792
21477997
21524650
21677988
22043812
25792700
28893291
29098528
29157984
29841859
T. sp. "ukuga"
T. sp. "black"
T. brichardi
T. polli
T. sp. "kirschfleck"
T. sp. "lunatus"
O. indermauri
P. tamula
P. marginatus (North)
P. curvifrons
P. marginatus
P. babaulti (South)
S. diagramma
G. pleiferi
P. polyodon
L. dardennii
C. horei
P. curvifrons
T. sp. "red"
L. loocki
L. labiatus
L. ephippium
P. honi
P. trewavasae
P. fasciolatus
P. tamula
T. mbipi
T. sp. "black"
P. babaulti
T. brauschi
T. sp. "black"
P. orthognathus
P. marginatus
P. orthognathus
P. sp. "kazumbae"
P. sp. "lunatus"
T. sp. "lukuga"
P. sp. "lunatus"
P. duboisi
T. sp. "red"
T. sp. "kirschfleck"
P. sp. "macrog. rainbow"
P. sp. "kipili brown"
P. polli
P. sp. "orthog. ikola"
P. sp. "giant"
P. sp. "macrog. rainbow"
P. macrophthalmus
G. pleiferi (South)
P. babaulti (South)
P. polyodon
P. sp. "polyodon texas"
S. diagramma
P. babaulti
P. macrophthalmus
G. pleiferi
O. linnelli
P. polyodon
P. sp. "polyodon texas"
T. labiatus
T. brichardi
P. ephippium
A. burtoni
P. orthognathus
A. burtoni
T. mooreii
A. burtoni
P. phlanderi
A. straeleni
P. marginatus (North)
A. flavijosephi
O. linnelli
T. sp. "chipwa"
O. uvinzae
O. malagarazensis
C. horei
T. sp. "brichardi kipili"
P. sp. "moshi yellow"
T. annectens
P. sp. "kazumbae"
G. polli
A. burtoni 7
A. vanheusdeni
T. sp. "murago"
T. sp. "murago we"
P. sp. "moshi yellow"
T. sp. "murago"
T. duboisi
T. sp. "red"
T. sp. "brichardi kipili"
T. sp. "mpimbwe"
P. horii
P. sp. "polyodon texas"
P. sp. "giant"
P. trewavasae
P. sp. "red"
P. loocki
P. fasciolatus
P. sp. "orthog. ikola"
P. sp. "macrog. rainbow"
P. sp. "kipili brown"
H. sp. "kilossana"
H. stappersii
A. paludinosa

Fig. S2 (continued)



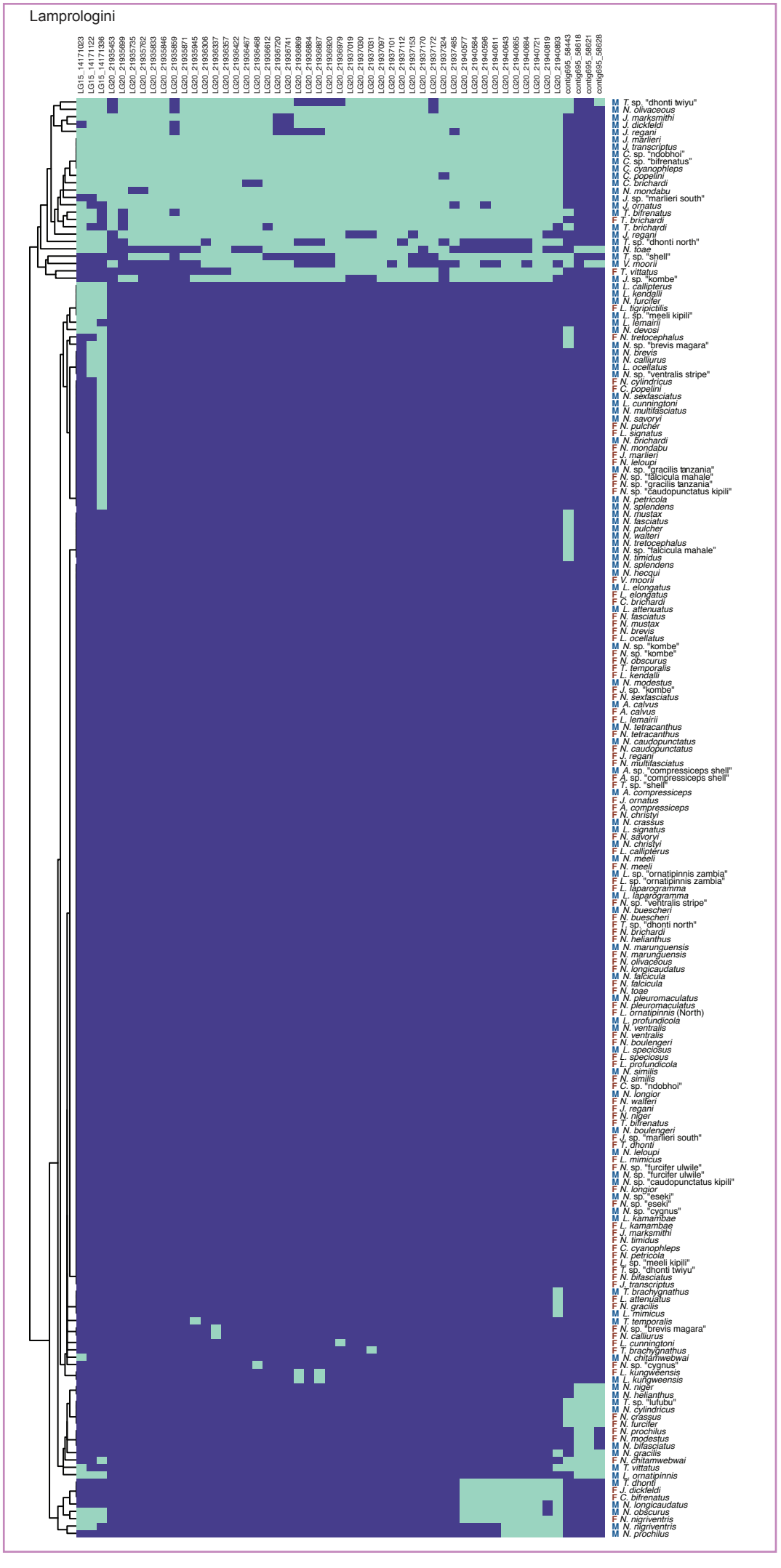


Fig. S2. Genotypes of GWAS outlier regions (approach 1). Heatmaps show individual genotypes for outlier SNPs (F: female individuals, M: male individuals; Haplochromini and Cyprichromini 100 most significant SNPs, Lamprologini three outlier regions on LG15, LG20 and an unplaced contig) of sex-linked regions detected by GWAS shown in fig. S1 for all individuals analysed by GWAS; purple: homozygous; green: heterozygous.

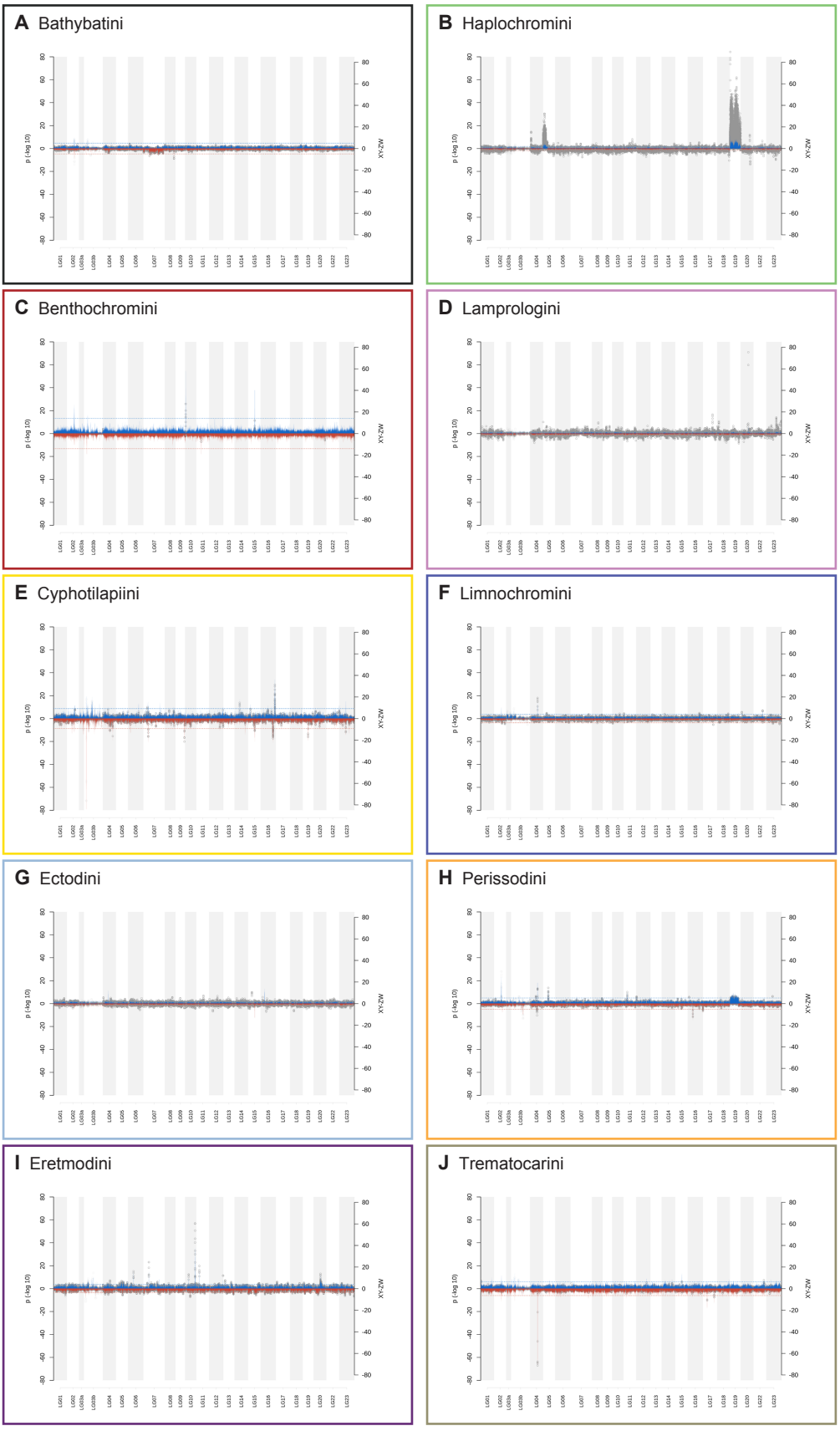


Fig. S3. Sex-specific SNP windows (approach 2) for the tribes Bathybatini (A), Haplochromini (B), Benthochromini (C), Lamprologini (D), Cyphotilapiini (E), Limnochromini (F), Ectodini (G), Perissodini (H), Eretmodini (I), and Trematocarini (J). Gray dots show results for tests of overrepresentation of XY- (left Y-axis, positive scale) and ZW- (left Y-axis, negative scale) SNPs in windows of 10 kb with a slide of 2 kb. Repetitive sequence regions on LGs 02, 03a and 03b were excluded in the final call set of sex chromosomes and are plotted in lighter gray. Background shadings refer to LGs of the reference genome. Vertical lines indicate XY-ZW SNP difference per window normalized by the number of species (right Y-axis), windows with more XY-SNPs are plotted with blue lines, those with more ZW-SNPs with red lines. Dotted horizontal lines indicate the largest absolute XY-ZW difference normalized by species number obtained in each tribe over 100 permutations (right Y-axis).

Fig. S4

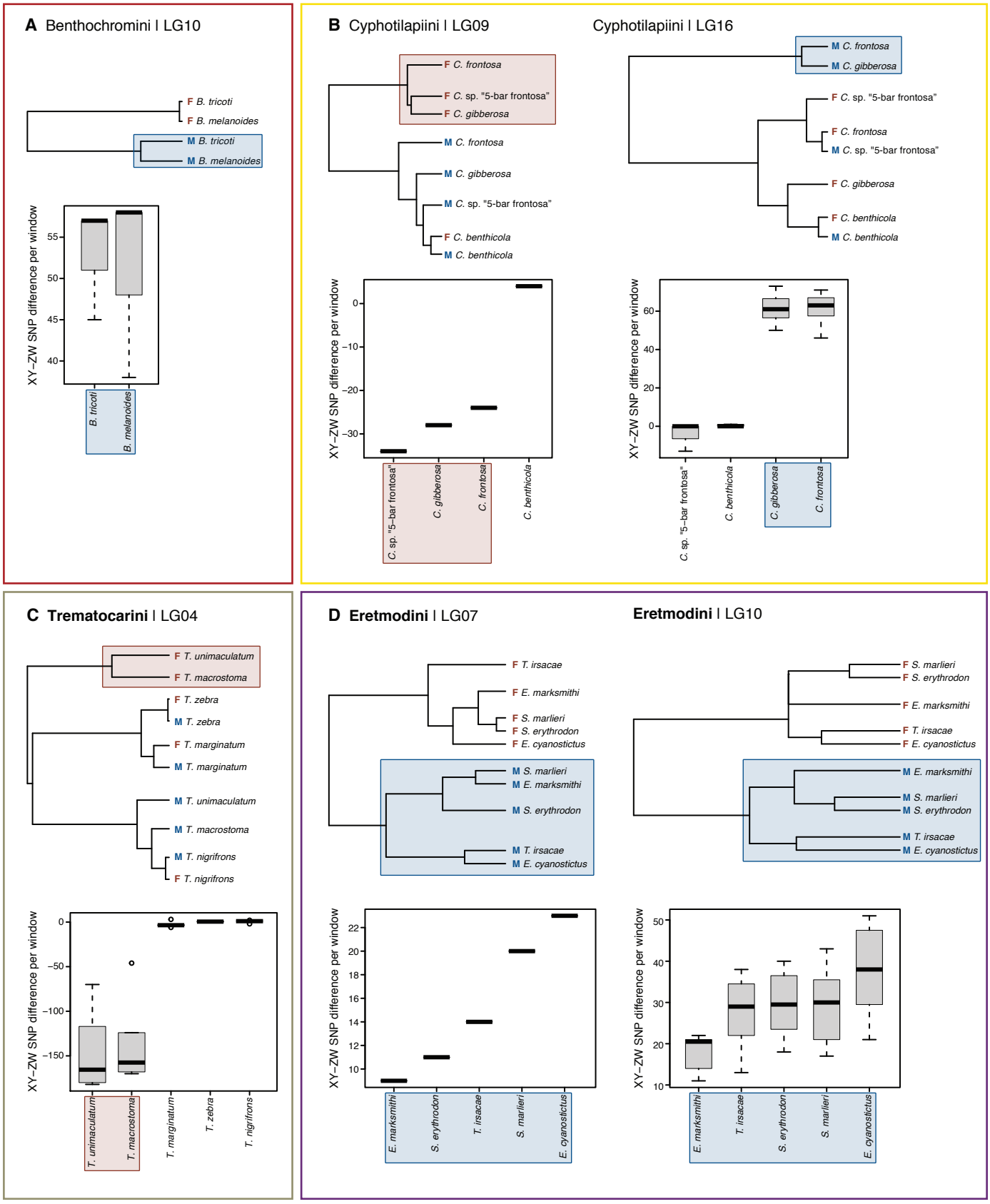

A Benthochromini | LG10
F B. tricoti
F B. melanoides
M B. tricoti
M B. melanoides
XY−ZW SNP difference per window
55
50
45
40
B. tricoti
B. melanoides
B Cyphotilapiini | LG09
F C. frontosa
F C. sp. "5-bar frontosa"
F C. gibberosa
M C. frontosa
M C. gibberosa
M C. sp. "5-bar frontosa"
F C. benthicola
M C. benthicola
XY−ZW SNP difference per window
0
−10
−20
−30
C. sp. "5−bar frontosa"
C. gibberosa
C. frontosa
C. benthicola
Cyphotilapiini | LG16
M C. frontosa
M C. gibberosa
F C. sp. "5-bar frontosa"
F C. frontosa
M C. sp. "5-bar frontosa"
F C. gibberosa
F C. benthicola
M C. benthicola
XY−ZW SNP difference per window
60
40
20
0
C. sp. "5−bar frontosa"
C. benthicola
C. gibberosa
C. frontosa
C Trematocarini | LG04
F T. unimaculatum
F T. macrostoma
F T. zebra
M T. zebra
F T. marginatum
M T. marginatum
M T. unimaculatum
M T. macrostoma
M T. nigrifrons
F T. nigrifrons
XY−ZW SNP difference per window
0
−50
−100
−150
T. unimaculatum
T. macrostoma
T. marginatum
T. zebra
T. nigrifrons
D Eretmodini | LG07
F T. irsacae
F E. marksmithi
F S. marlieri
F S. erythrodon
F E. cyanostictus
M S. marlieri
M E. marksmithi
M S. erythrodon
M T. irsacae
M E. cyanostictus
XY−ZW SNP difference per window
22
20
18
16
14
12
10
E. marksmithi
S. erythrodon
T. irsacae
S. marlieri
E. cyanostictus
Eretmodini | LG10
F S. marlieri
F S. erythrodon
F E. marksmithi
F T. irsacae
F E. cyanostictus
M E. marksmithi
M S. marlieri
M S. erythrodon
M T. irsacae
M E. cyanostictus
XY−ZW SNP difference per window
50
40
30
20
10
E. marksmithi
T. irsacae
S. erythrodon
S. marlieri
E. cyanostictus

Fig. S4 (continued)

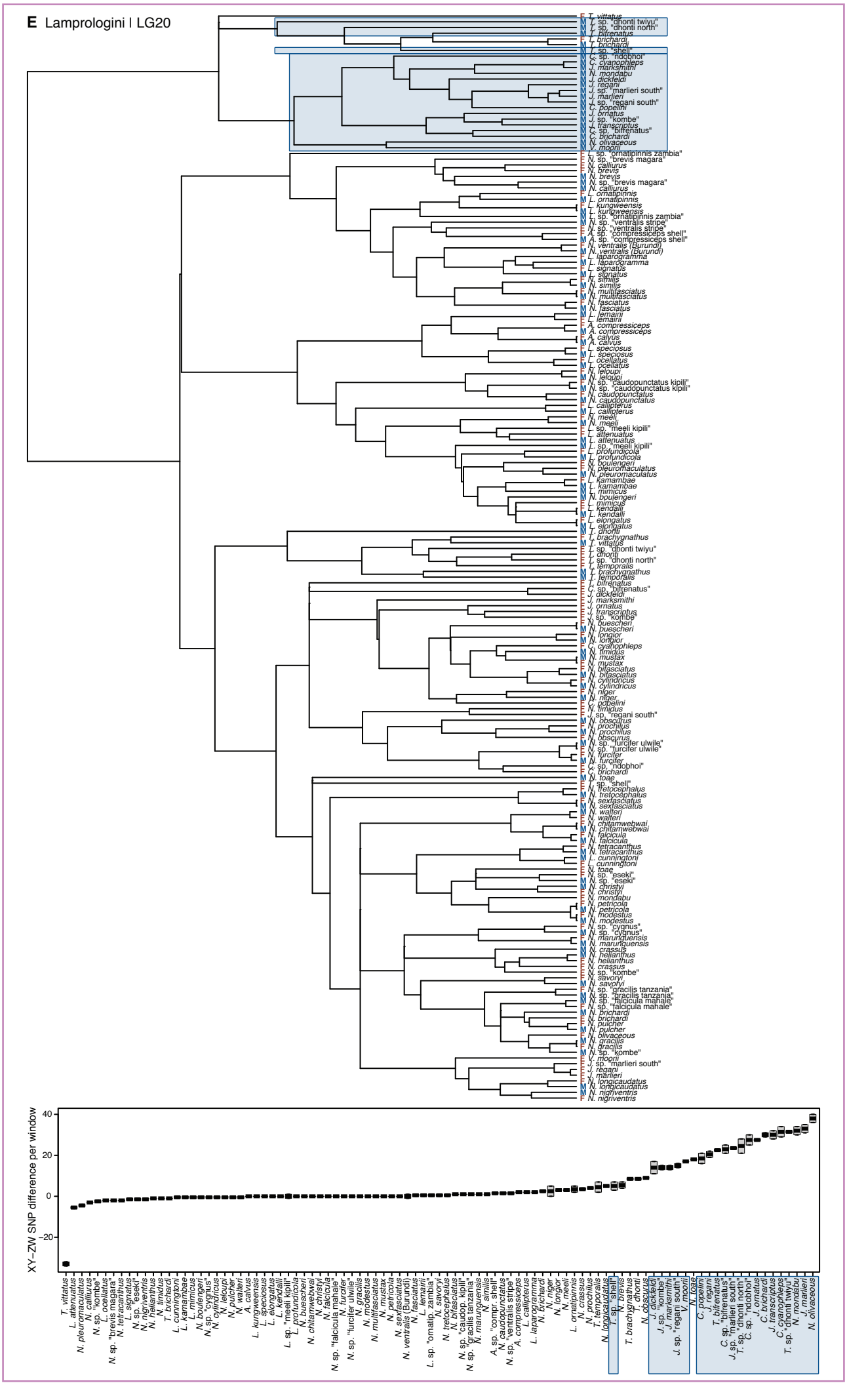


Fig. S4 (continued)

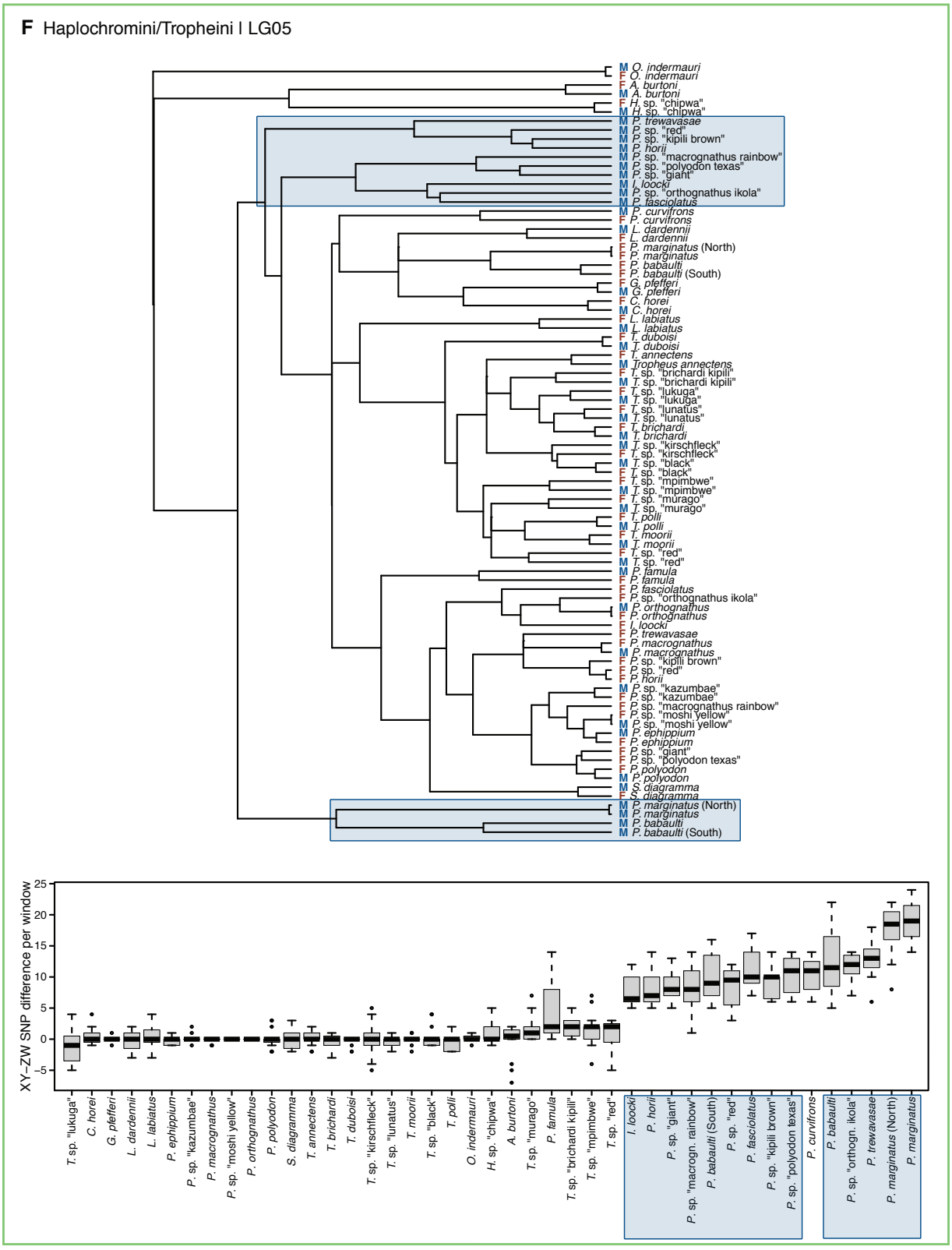


Fig. S4 (continued)

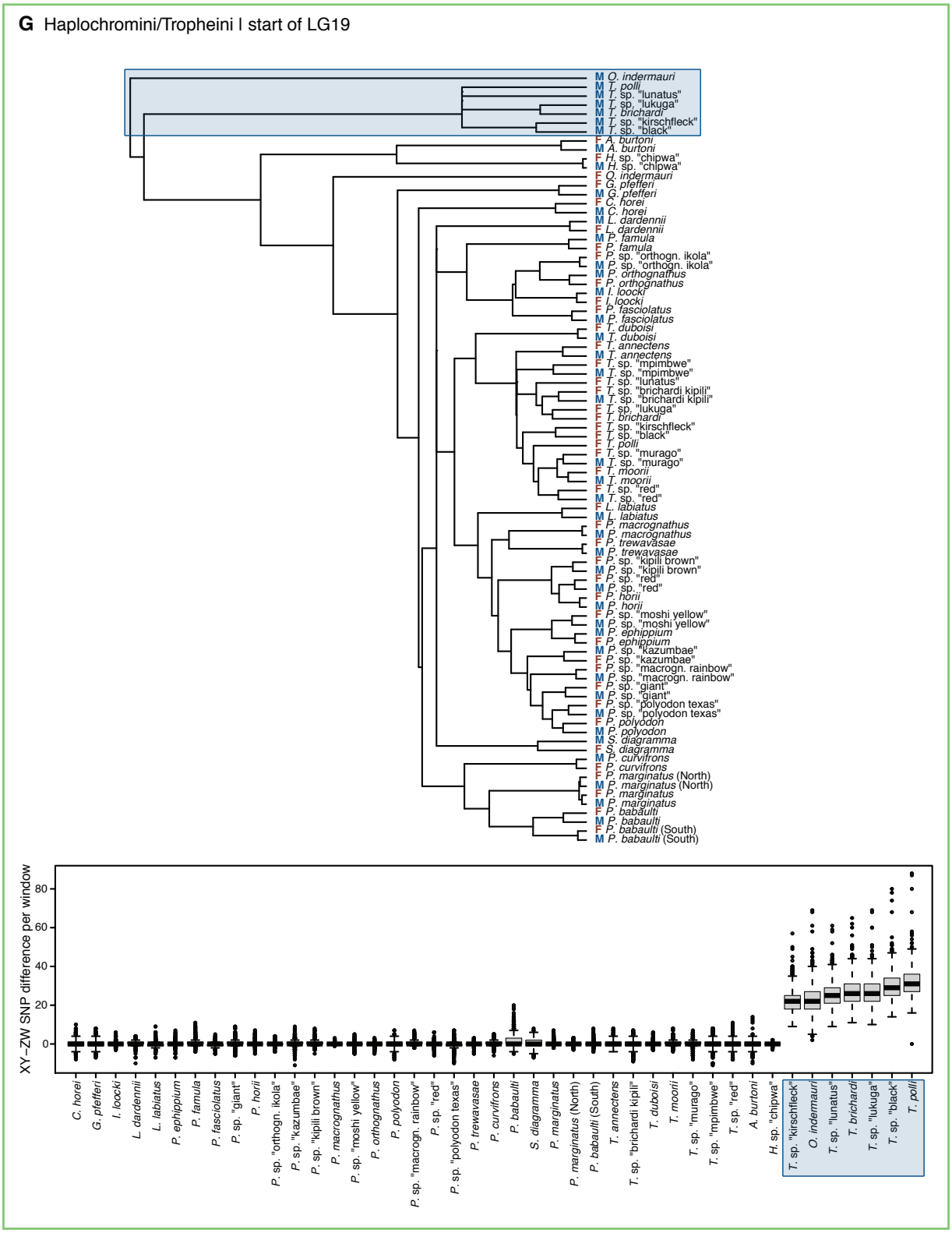
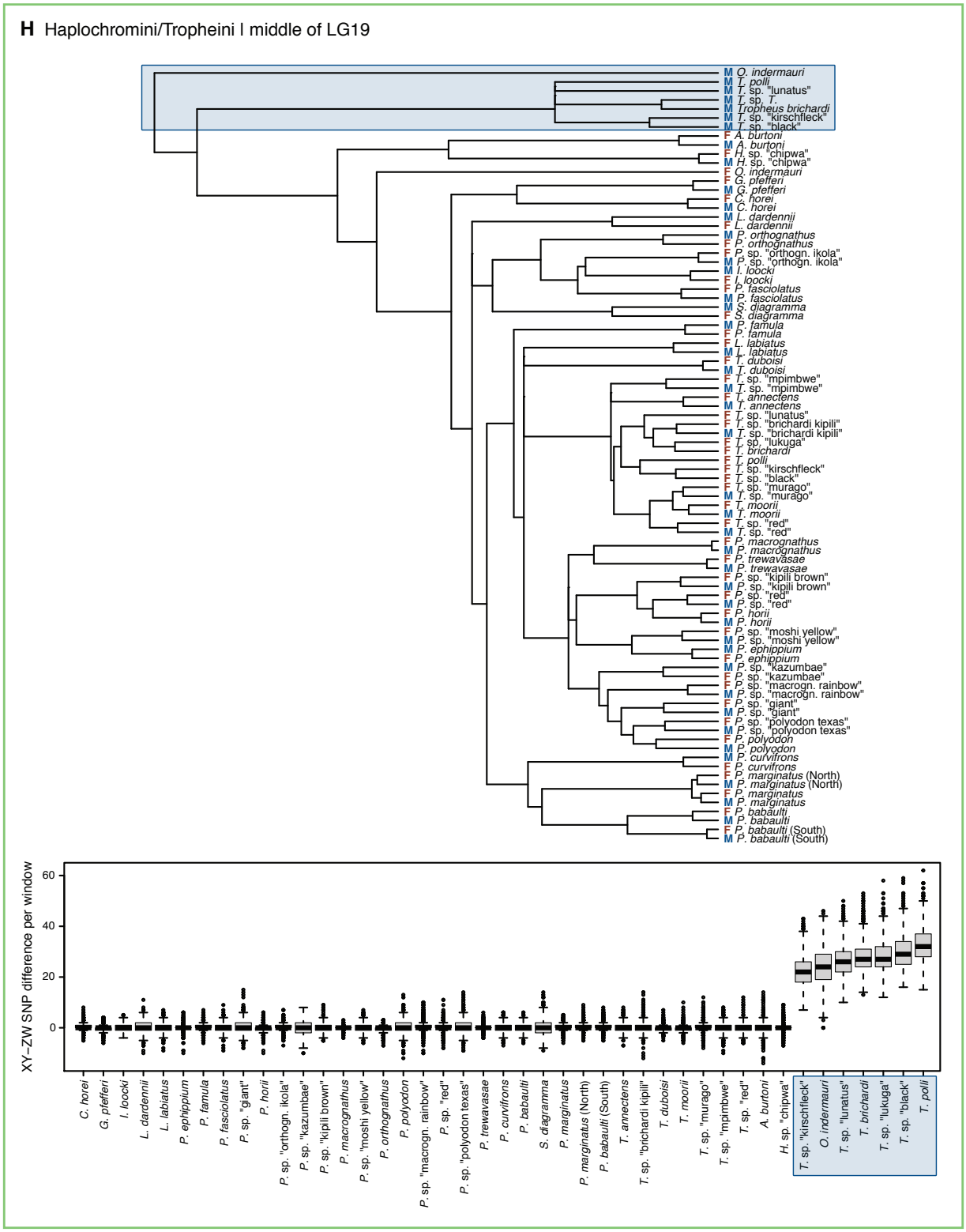


Fig. S4 (continued)



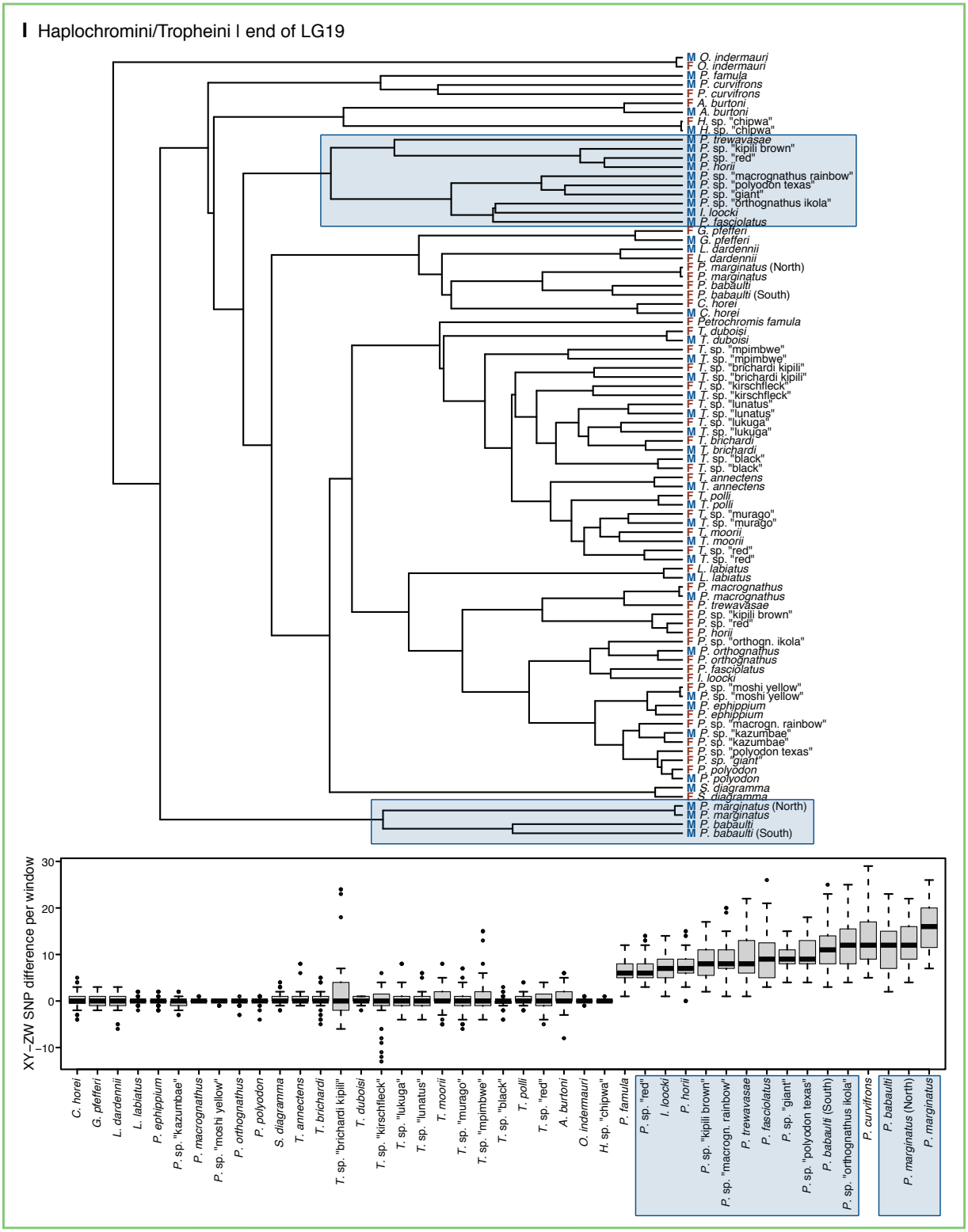


Fig. S4. Species values of sex-specific outlier regions (approach 2) for the tribes Benthochromini (A), Cyphotilapiini (B), Trematocarini (C), Eretmodini (D), Lamprologini (E), and Haplochromini (F-I). Dendograms of samples based on divisive hierarchical clustering of genotype values in outlier windows of approach 2 with individuals clustering by sex colored in blue for XY-systems and in red for ZW-systems and boxplots of per species XY-ZW difference in XY/ZW outlier windows (fig. S3) with color-coding based on the dendograms. Analysis include all individuals analysed by approach 2 (fig. S3). Boxplot center lines represent the median, box limits the upper and lower quartiles, and whiskers the 1.5x interquartile range; points represent outliers. F: female individuals, M: male individuals.

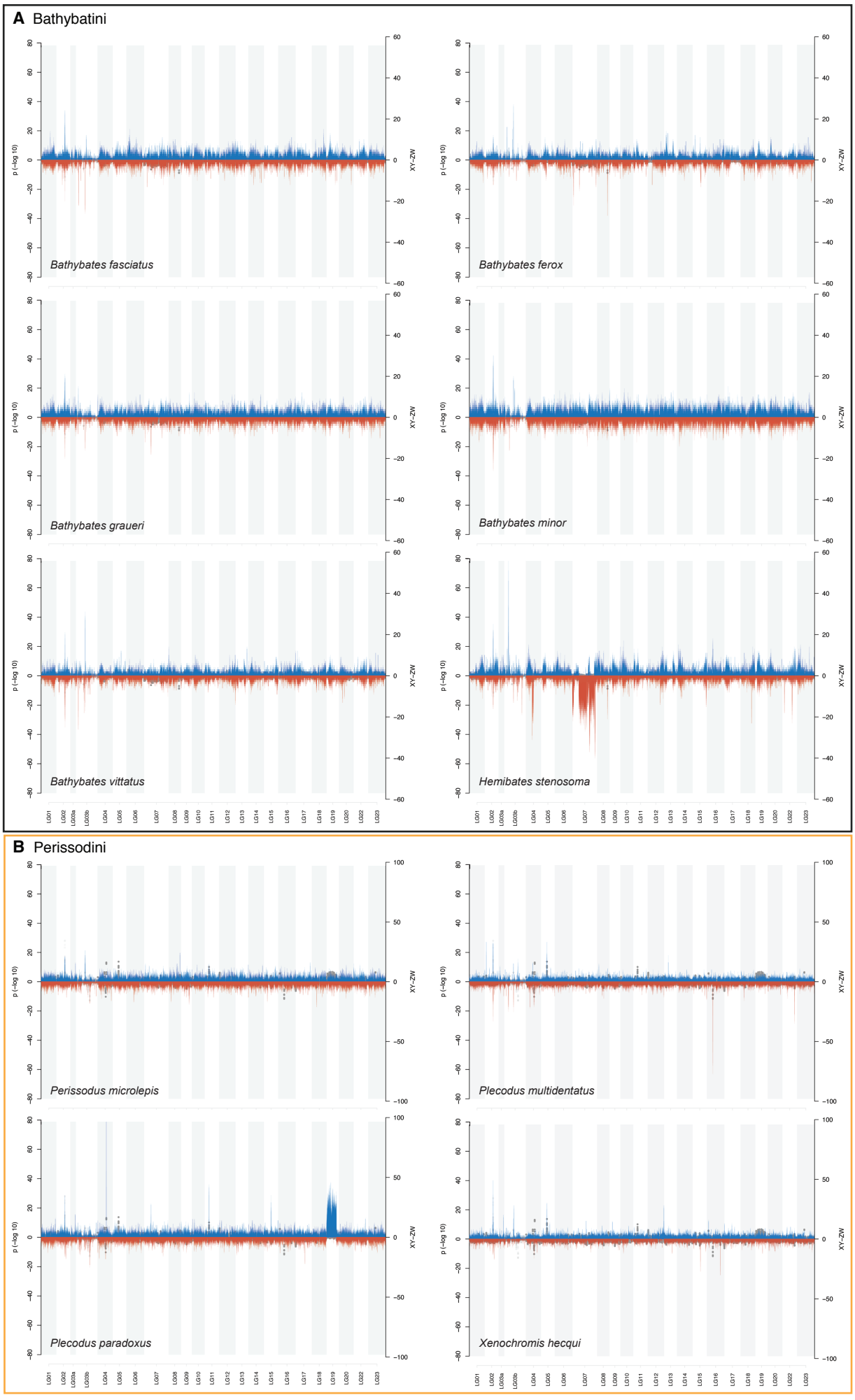


Fig. S5. Chromosome-wide sex-specific alleles in *Hemibates stenosoma* and *Plecodus paradoxus*. Visualization of per species XY-ZW differences in windows of 10 kb with a slide of 2 kb overlaid on the tribewise test results for sex-specific SNPs for Bathybatini (**A**) and Perissodini (**B**) (fig. S3).

Fig. S6

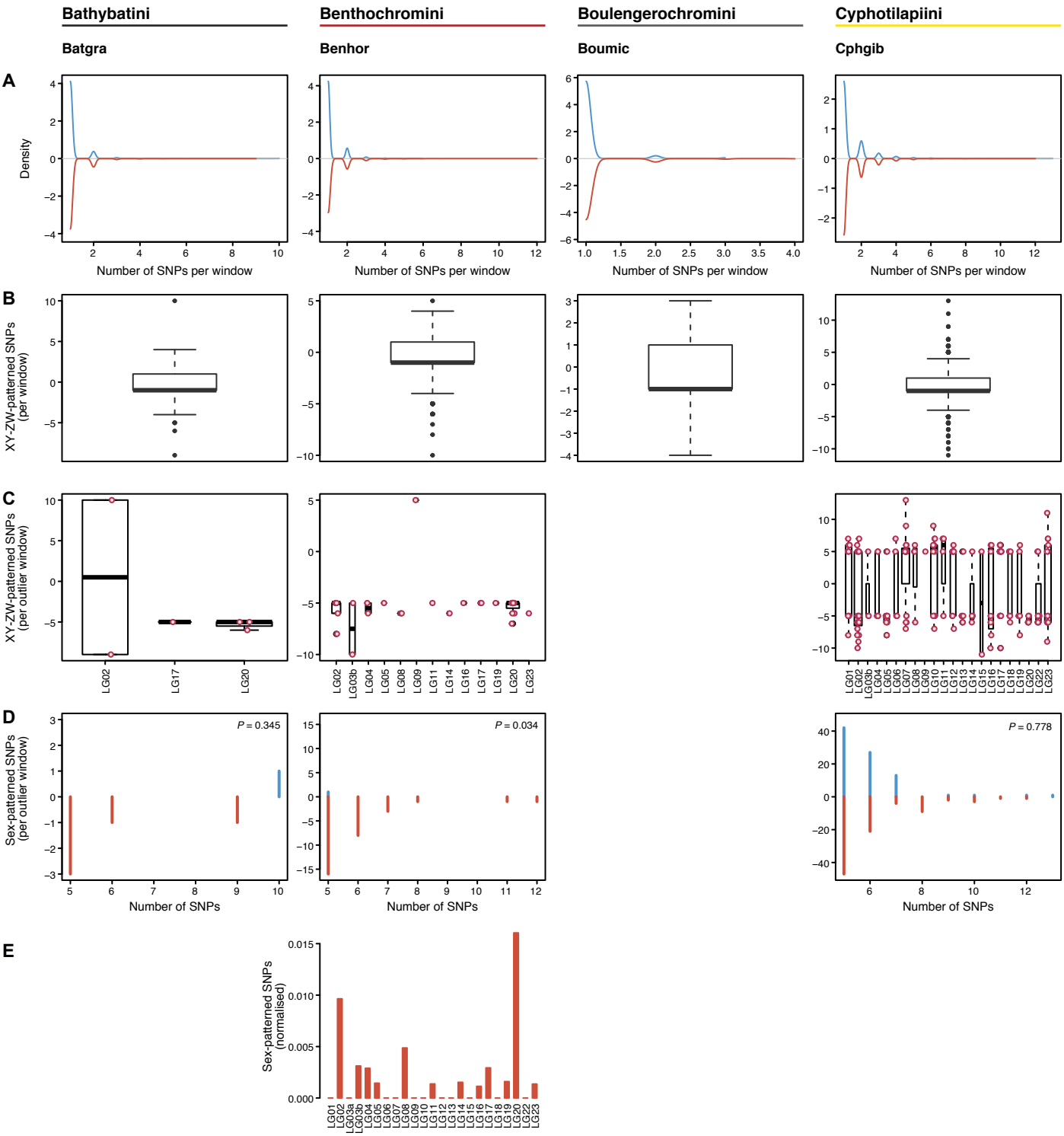

Bathybatini
Benthochromini
Boulengerochromini
Cyphotilapiini
Batgra
Benhor
Boumic
Cphgib
A
Density
Number of SNPs per window
B
XY-ZW-patterned SNPs (per window)
C
XY-ZW-patterned SNPs (per outlier window)
LG02
LG17
LG20
LG03b
LG04
LG05
LG08
LG09
LG11
LG14
LG16
LG17
LG19
LG20
LG23
LG01
LG03a
LG06
LG07
LG10
LG12
LG13
LG15
LG18
LG22
D
Sex-patterned SNPs (per outlier window)
P = 0.345
P = 0.034
P = 0.778
Number of SNPs
E
Sex-patterned SNPs (normalised)

Fig. S6 (continued)

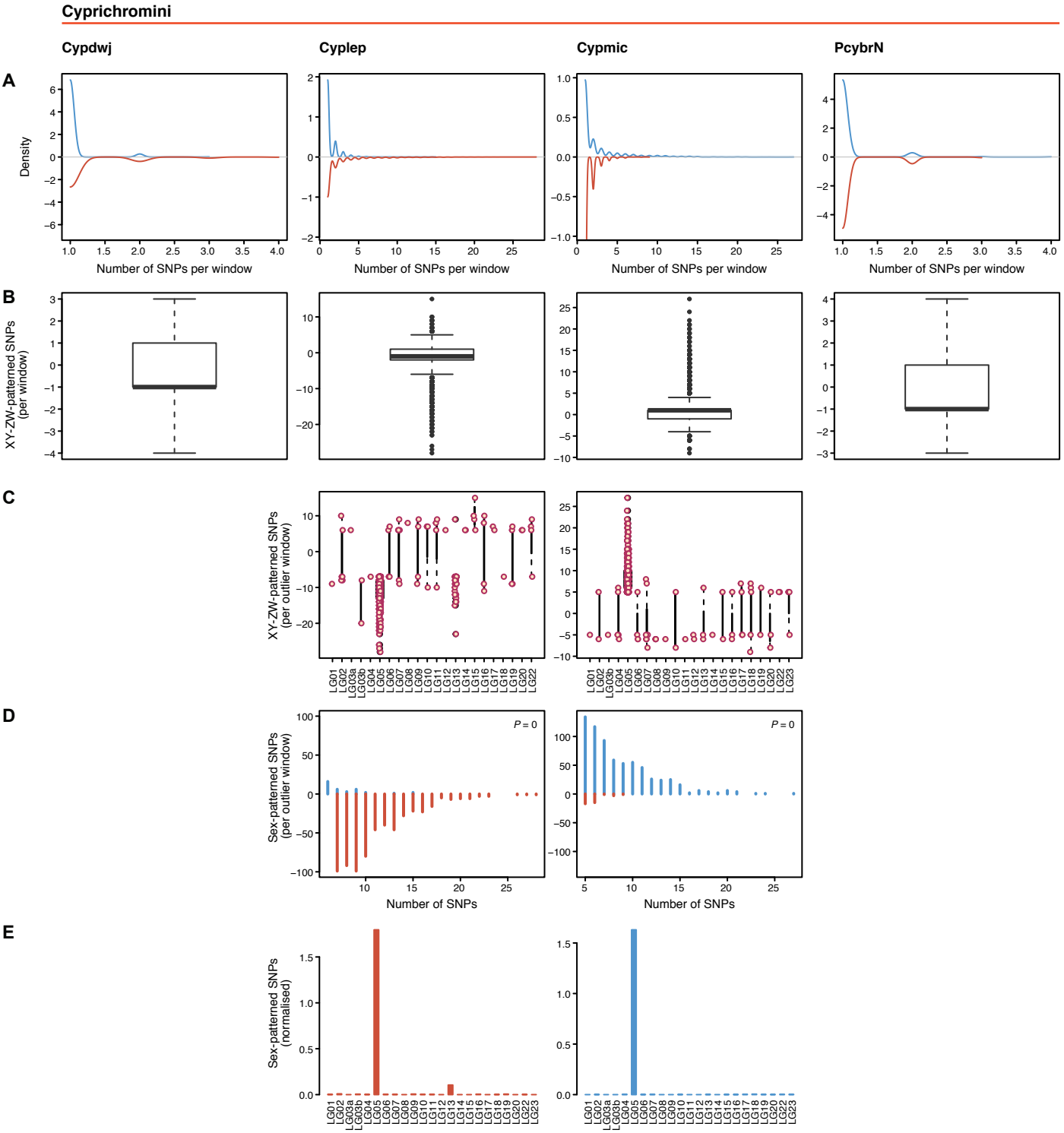


Fig. S6 (continued)

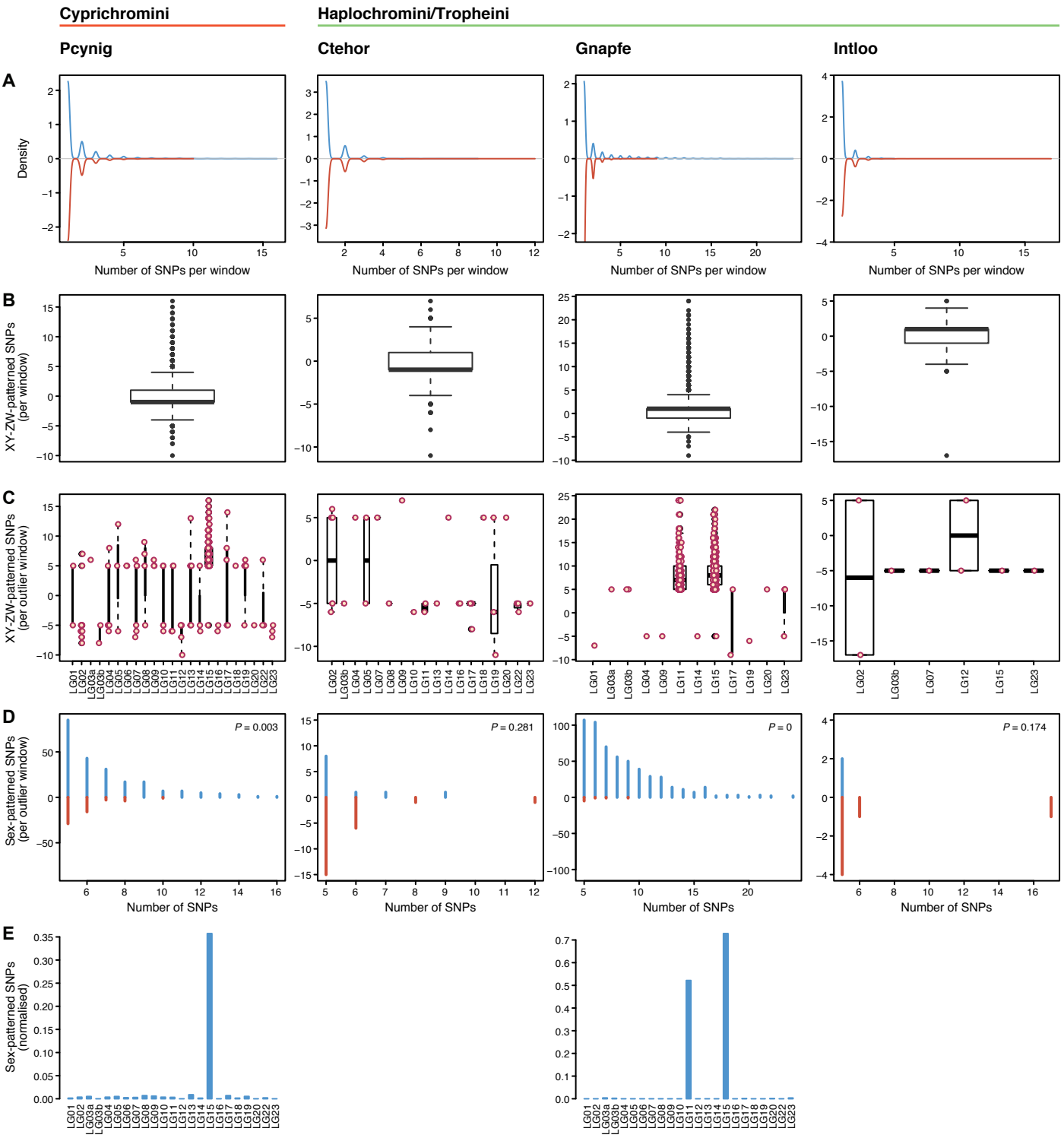


Fig. S6 (continued)

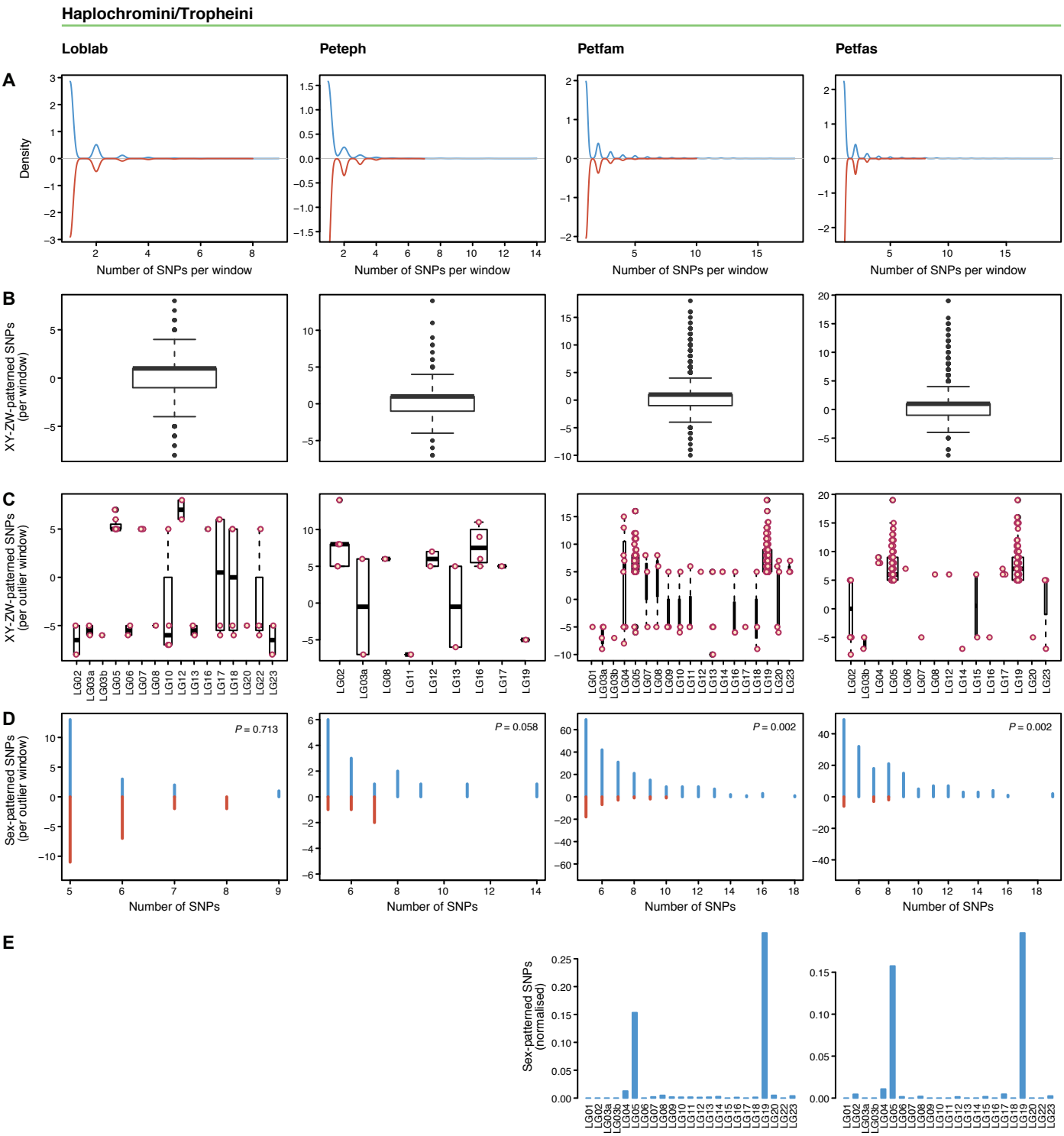


Fig. S6 (continued)

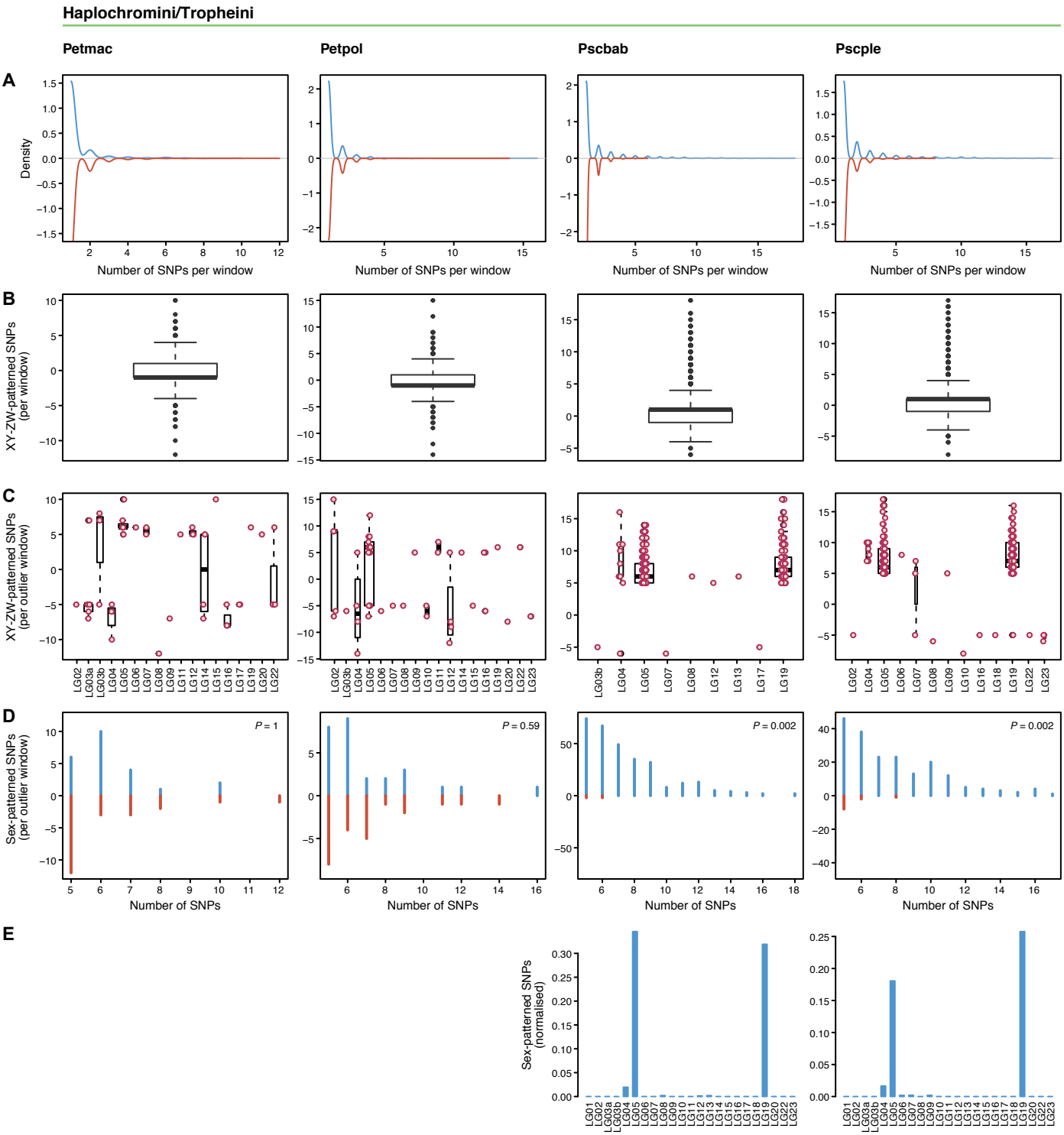


Fig. S6 (continued)

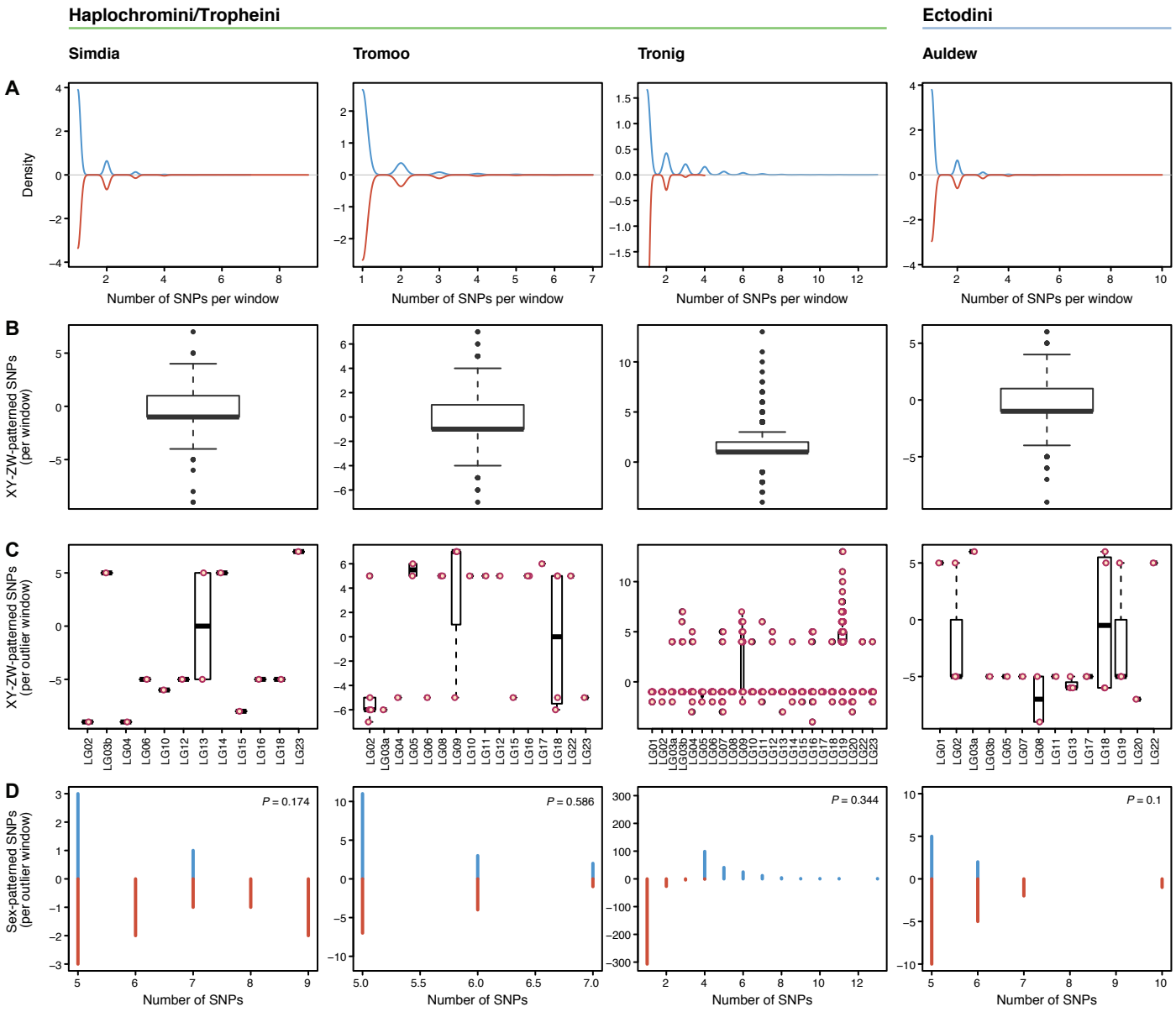


Fig. S6 (continued)

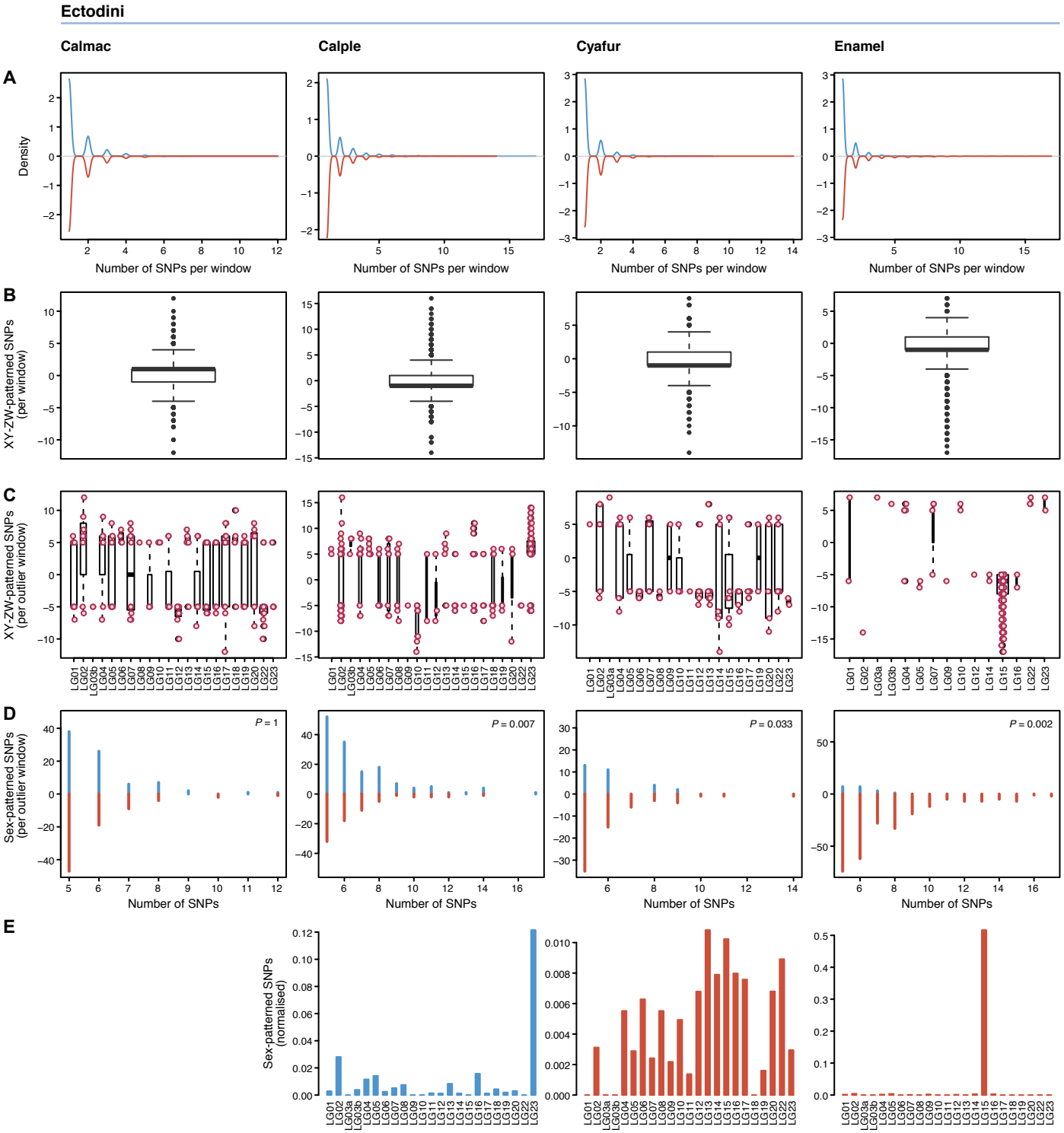


Fig. S6 (continued)

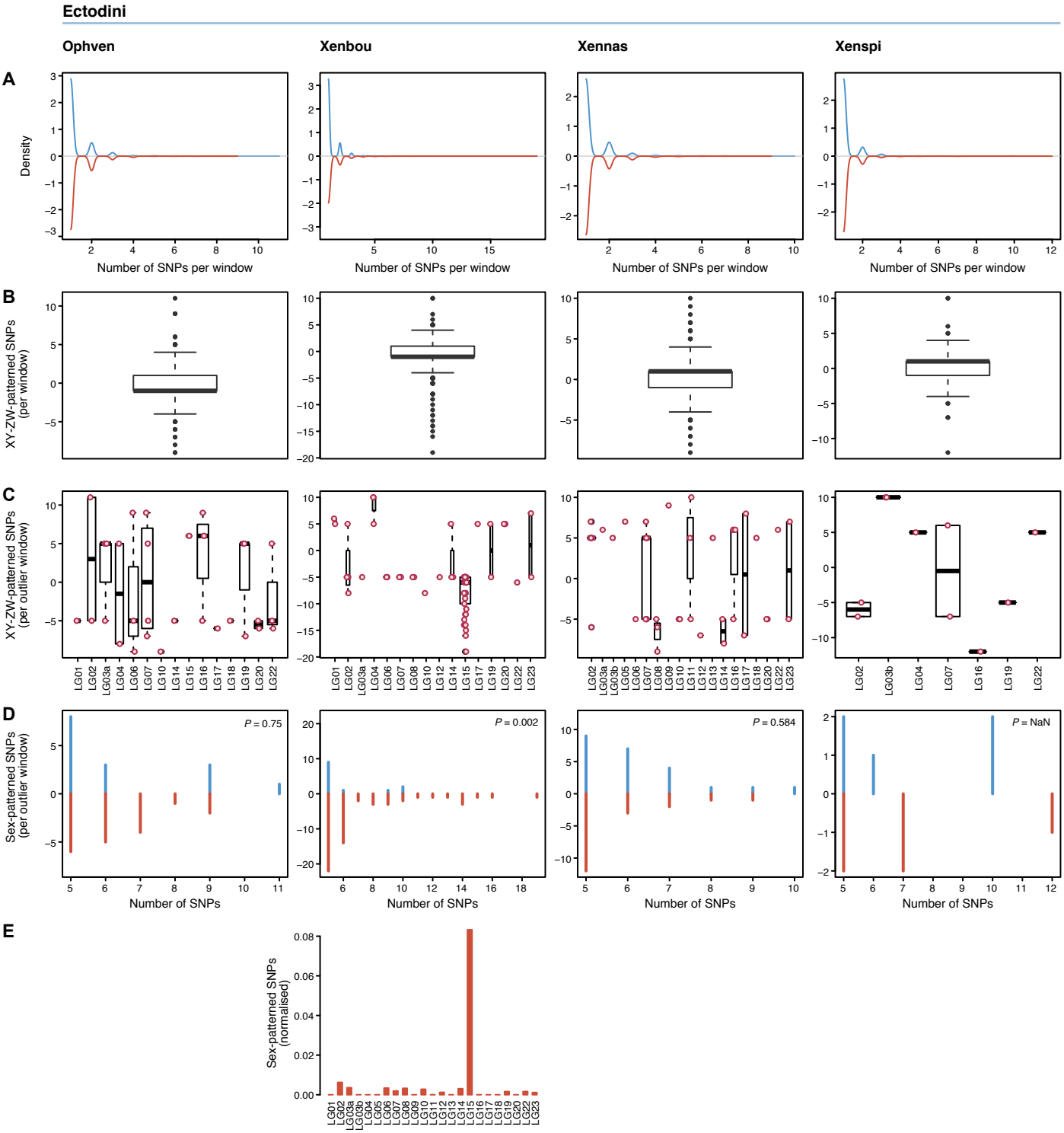


Fig. S6 (continued)

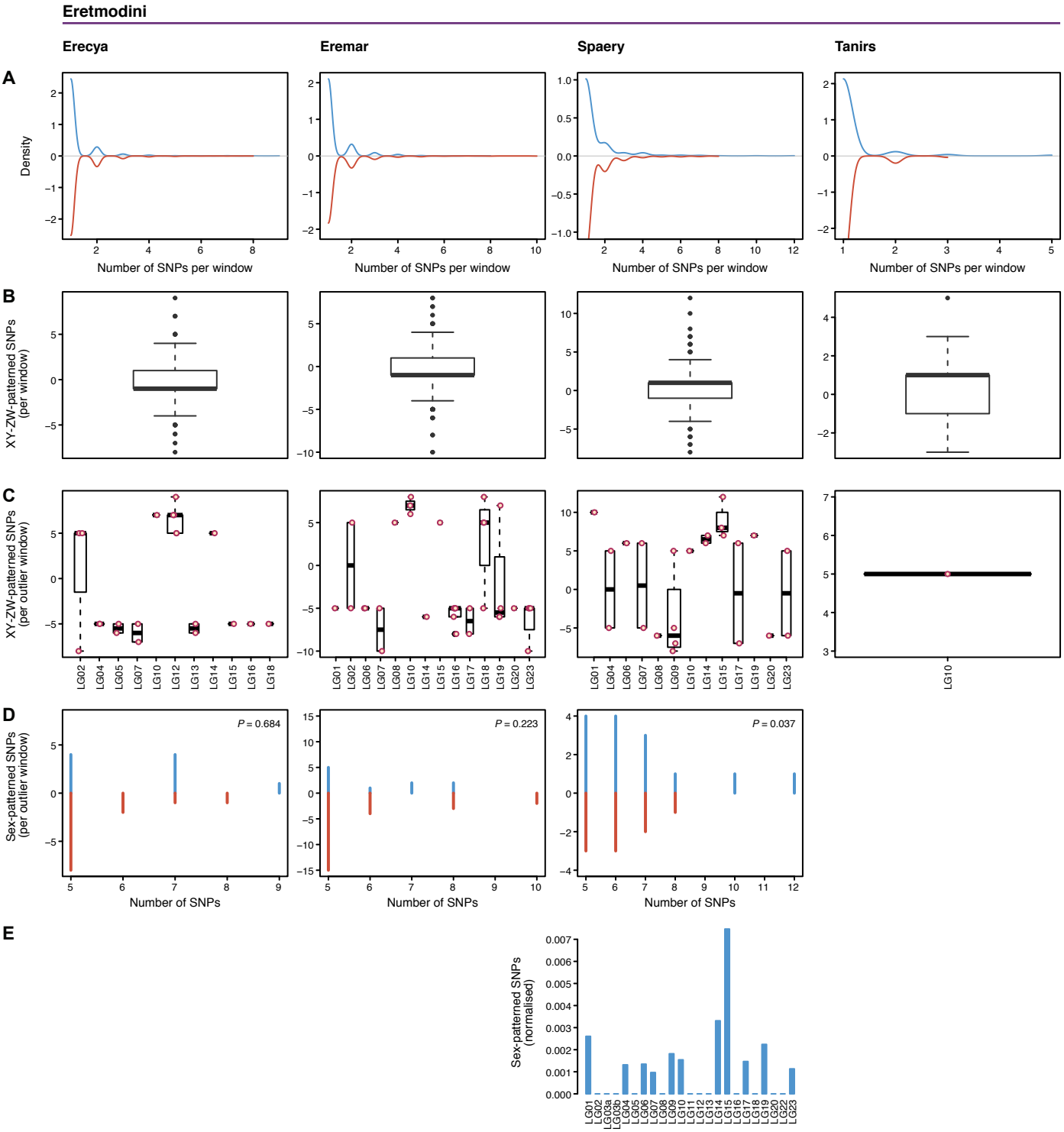


Fig. S6 (continued)

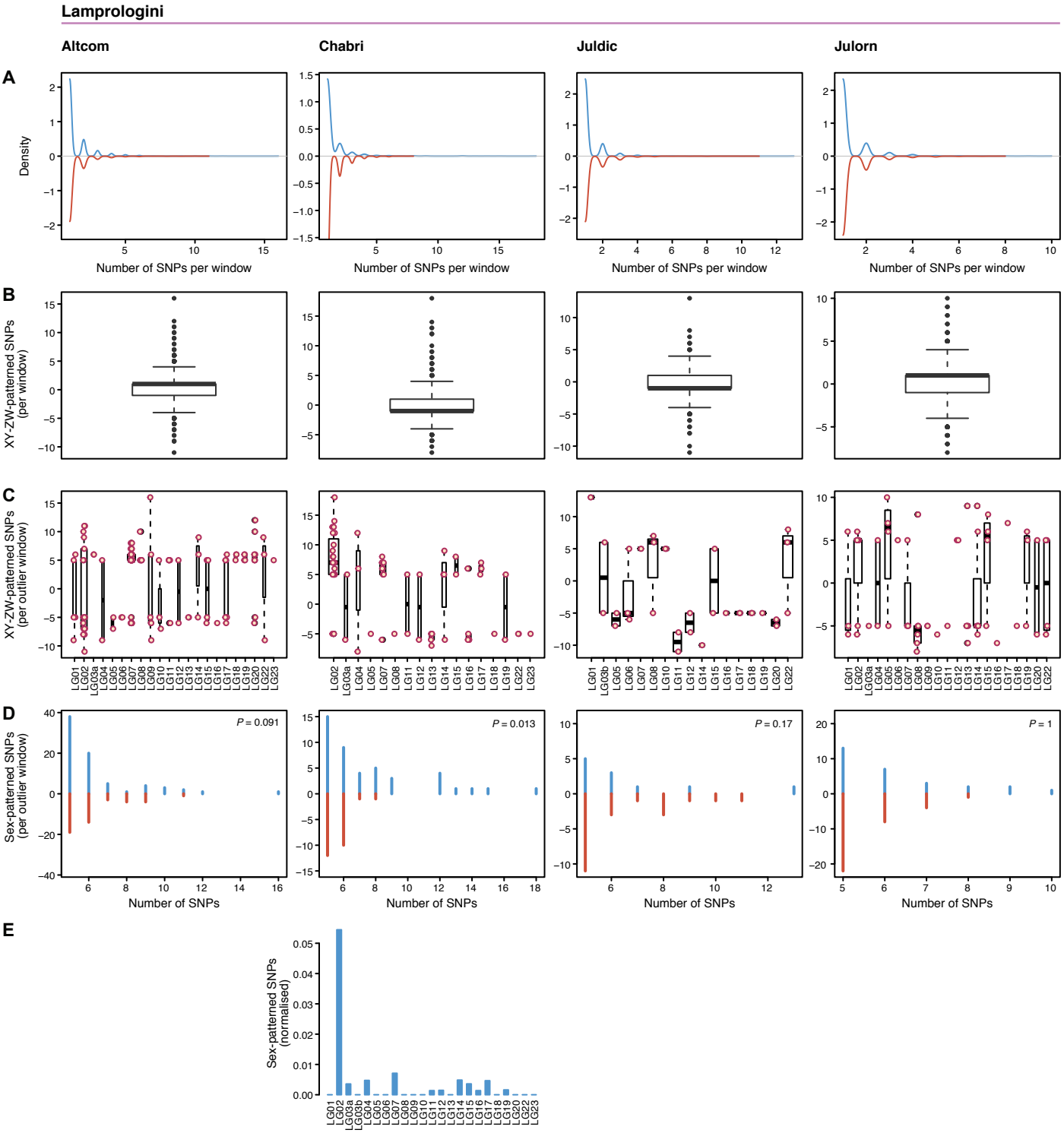


Fig. S6 (continued)

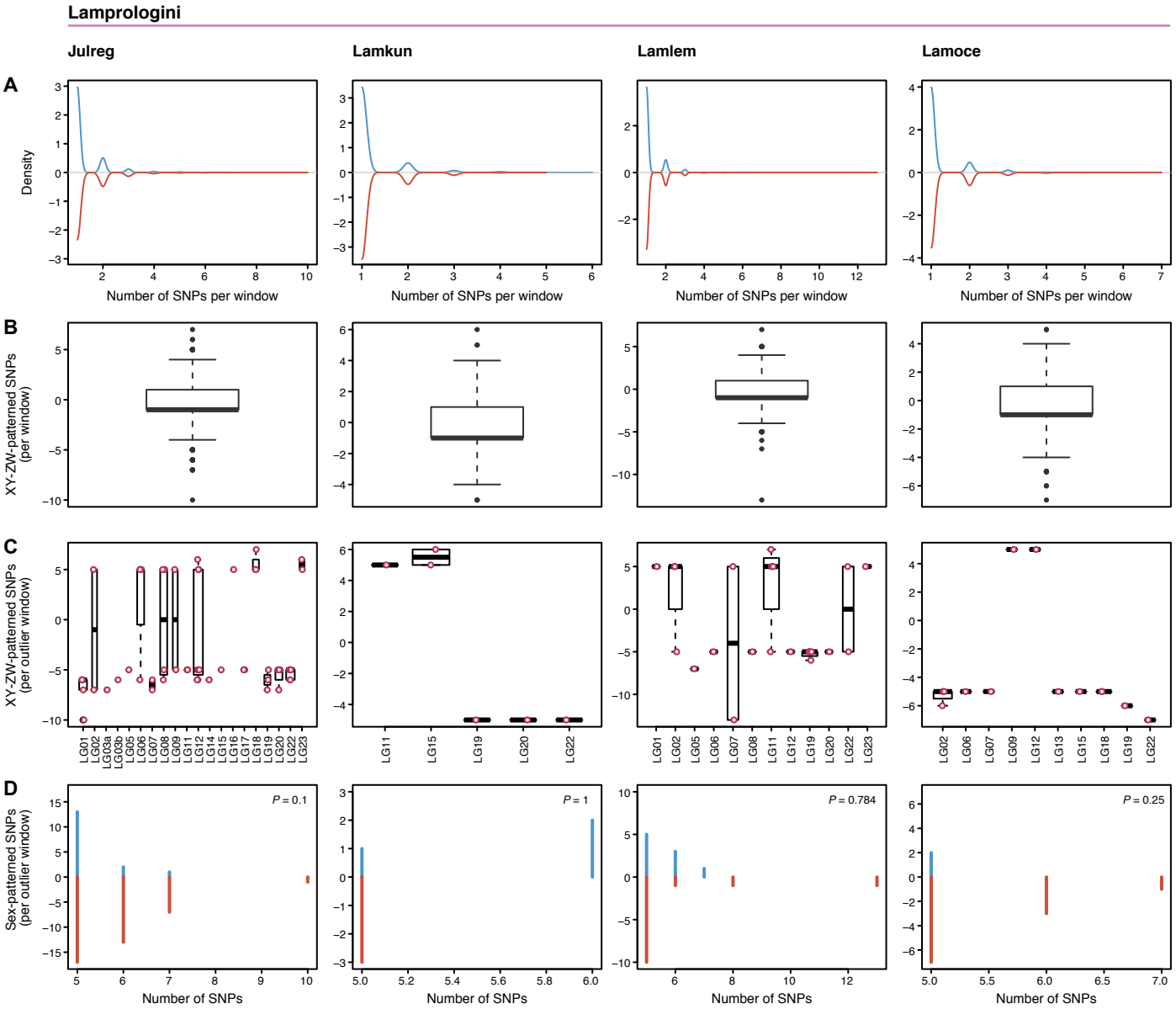


Fig. S6 (continued)

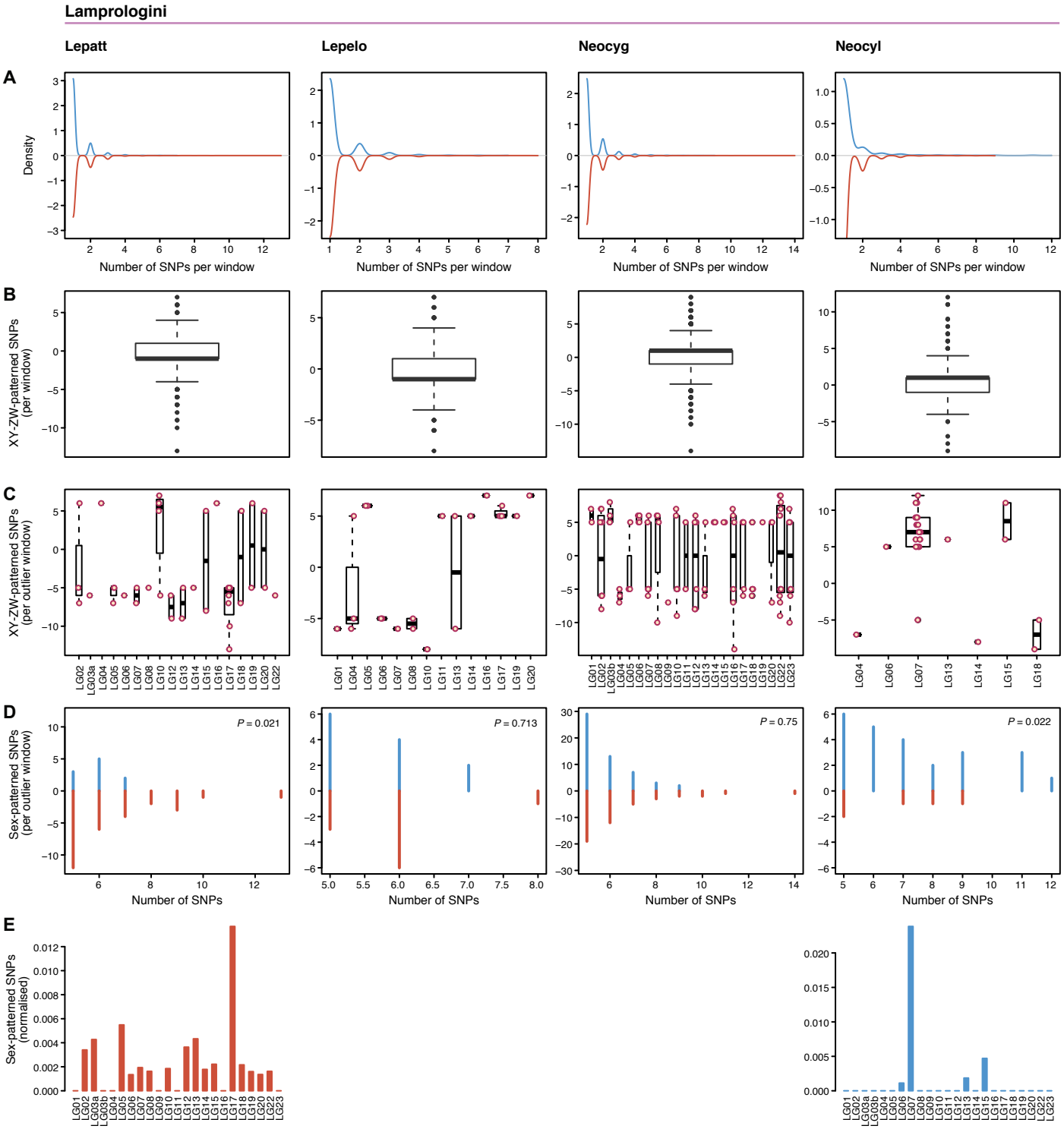


Fig. S6 (continued)

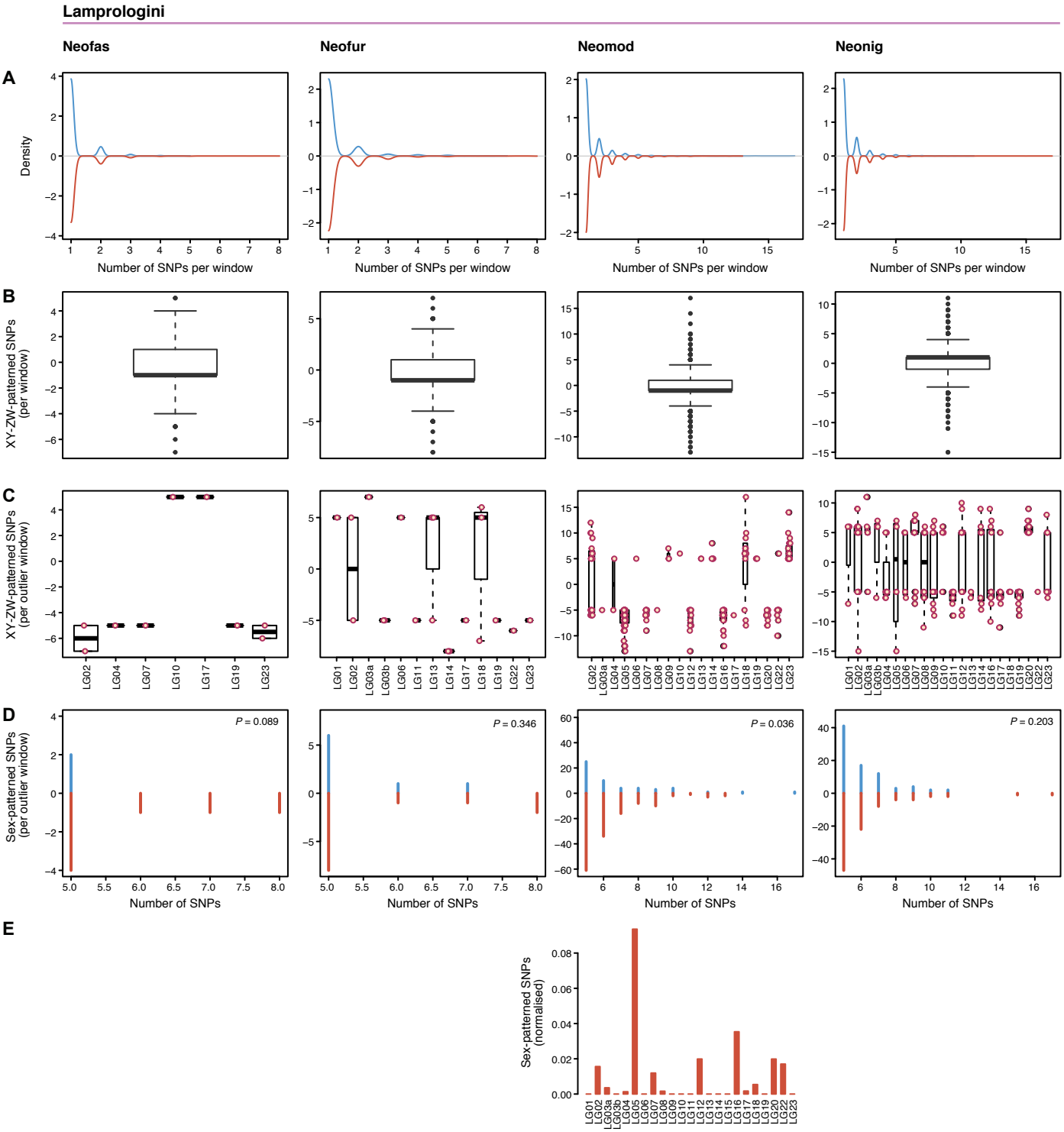


Fig. S6 (continued)

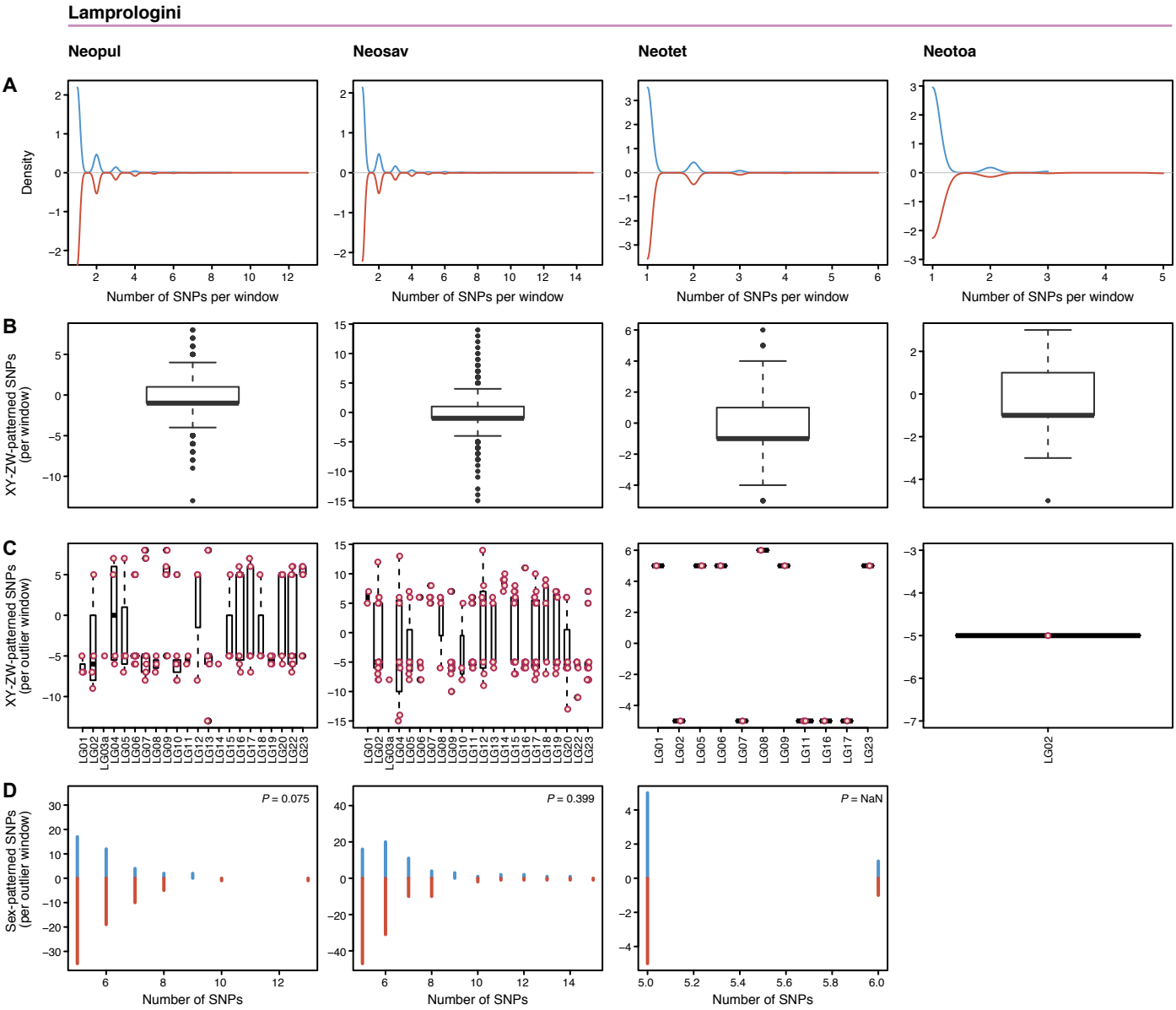


Fig. S6 (continued)

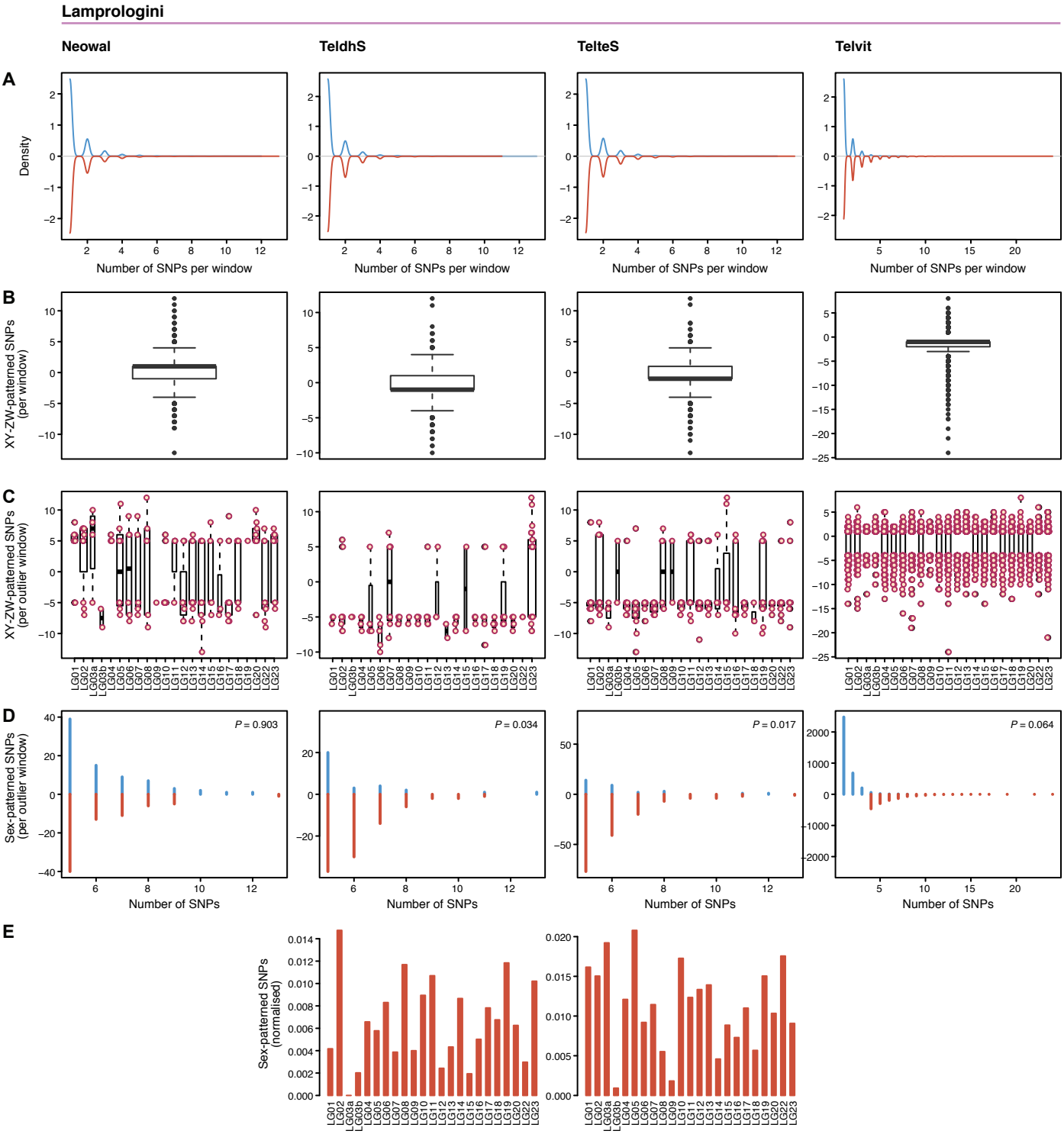


Fig. S6 (continued)

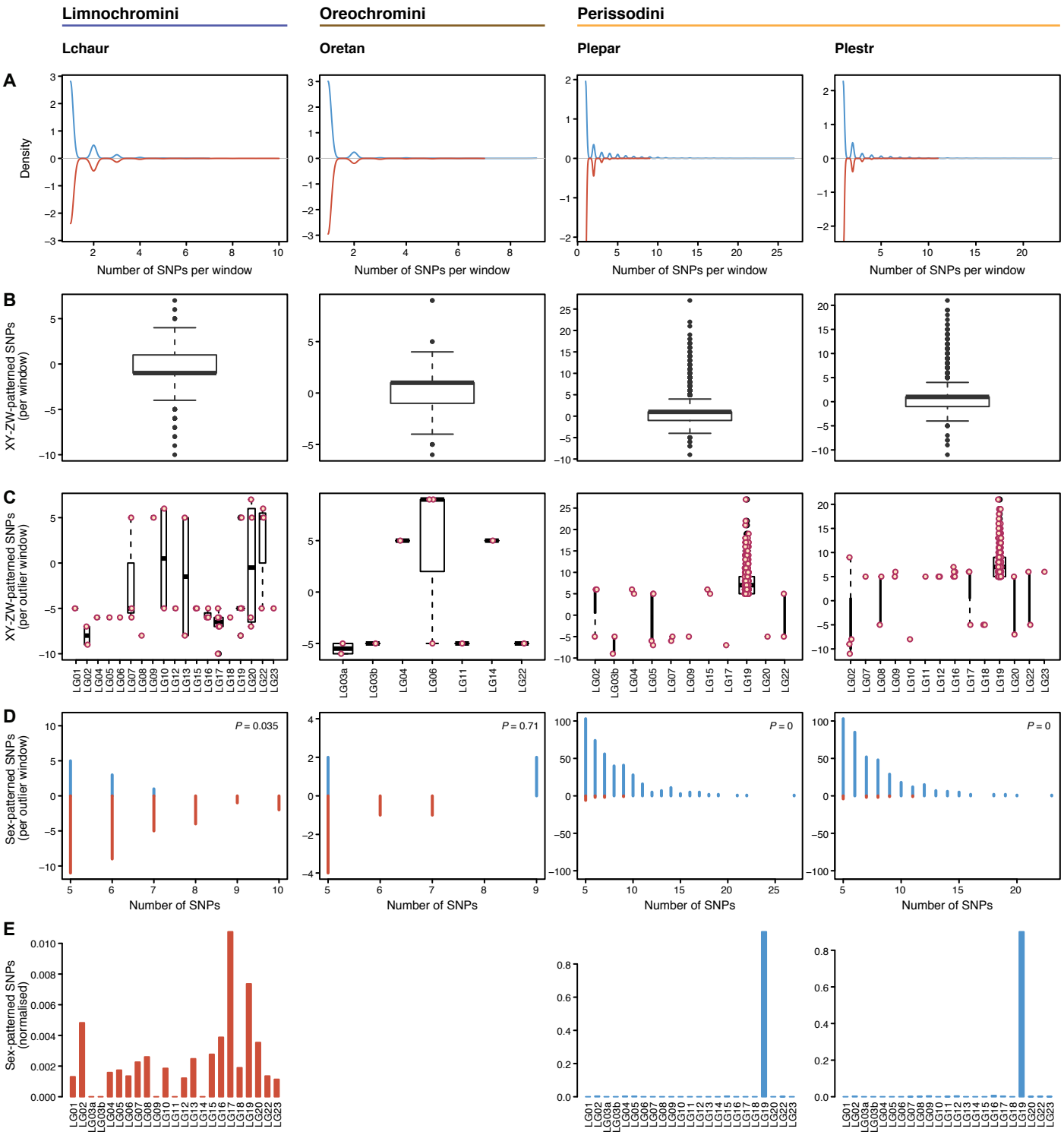


Fig. S6 (continued)

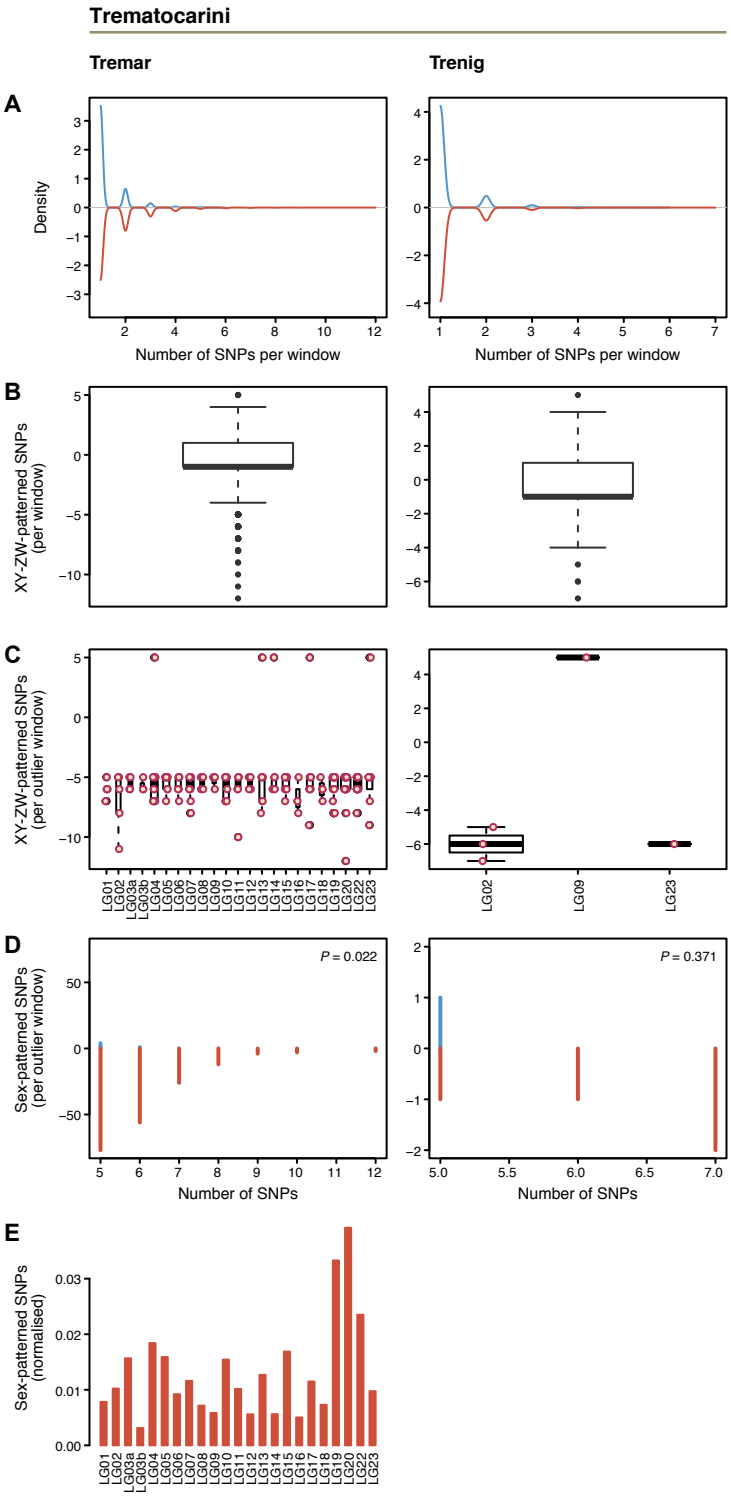


Fig. S6. Identification of sex chromosomes based on transcriptome data (approach 3). Per-species visualization of the results for tests of allele differences based on replicate male and female transcriptome data (approach 3). For each species (columns), four to five plots (rows) are shown, (**A**) Density plot of the number of sex-specific SNPs per 10 kb non-overlapping window; the density of XY- windows is represented in blue (positive y-axis scale) and the density of ZW-windows in red (negative y-axis scale); (**B**) Boxplot of the difference between XY- and ZW-SNPs in 10 kb windows. Windows with a higher number of XY-SNPs are represented with positive values (>0) while windows with a higher number of ZW-SNPs are represented with negative values (<0). Boxplot center lines represent the median; box limits the upper and lower quartiles and whiskers the 1.5x interquartile ranges. The width of the boxes is proportional to the number of observations; (**C**) boxplot of the difference in XY- and ZW-SNPs for outlier windows only and represented per LG. Boxplot center lines represent the median; box limits the upper and lower quartiles and whiskers the 1.5x interquartile ranges. The width of the boxes is proportional to the number of observations. (**D**) Histogram of the number of sex-specific SNPs per 10 kb outlier window, XY-windows are represented in blue (positive y-axis scale) and ZW-windows in red (negative y-axis scale). For species with a detected heterogametic system (**E**) Visualization of the number of sex-specific (XY-system in blue and ZW-system in red) SNPs of outlier windows normalized by chromosome length with species-specific Y-axis scales.

Fig. S7

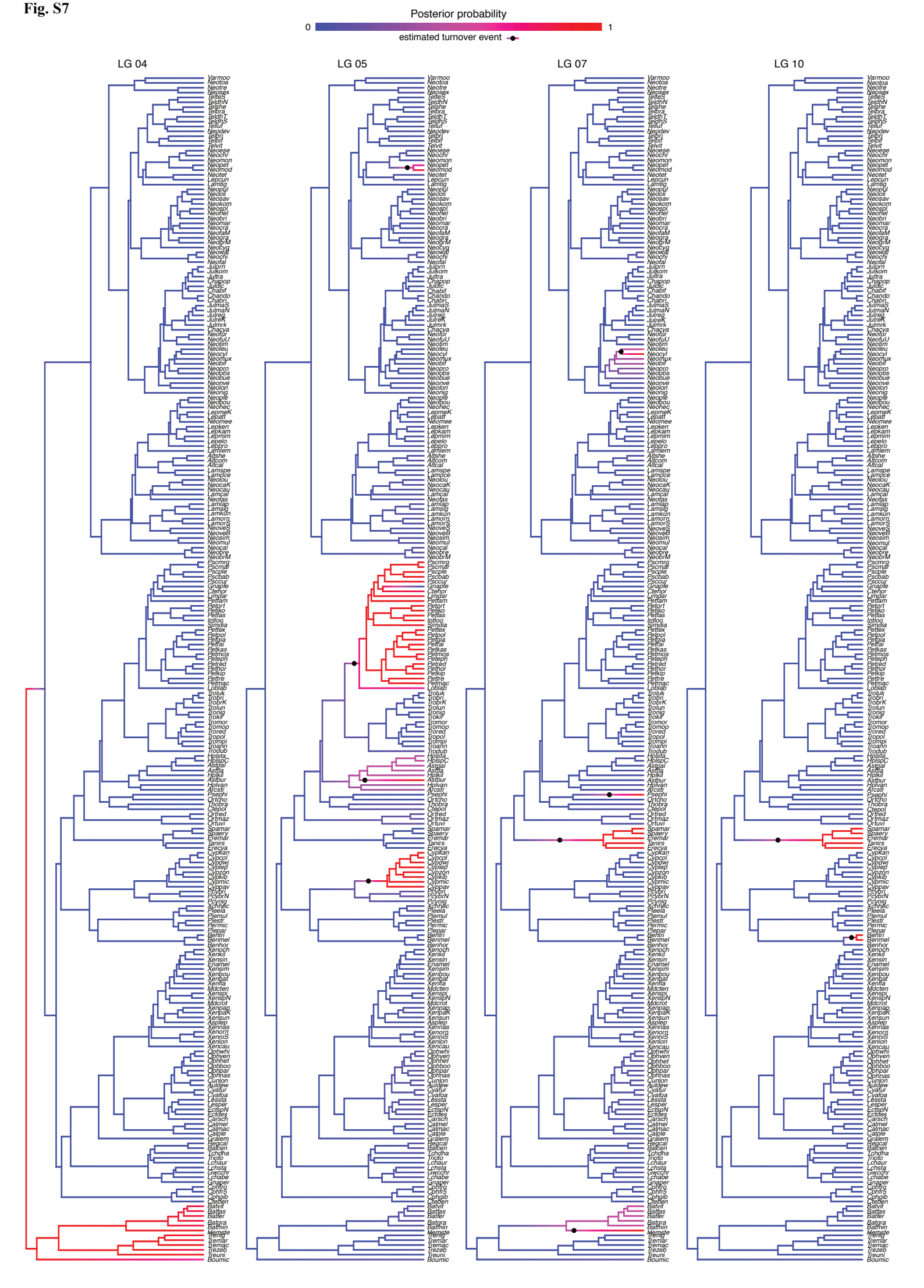

Posterior probability
0
1
estimated turnover event
LG 04
LG 05
LG 07
LG 10

Fig. S7 (continued)

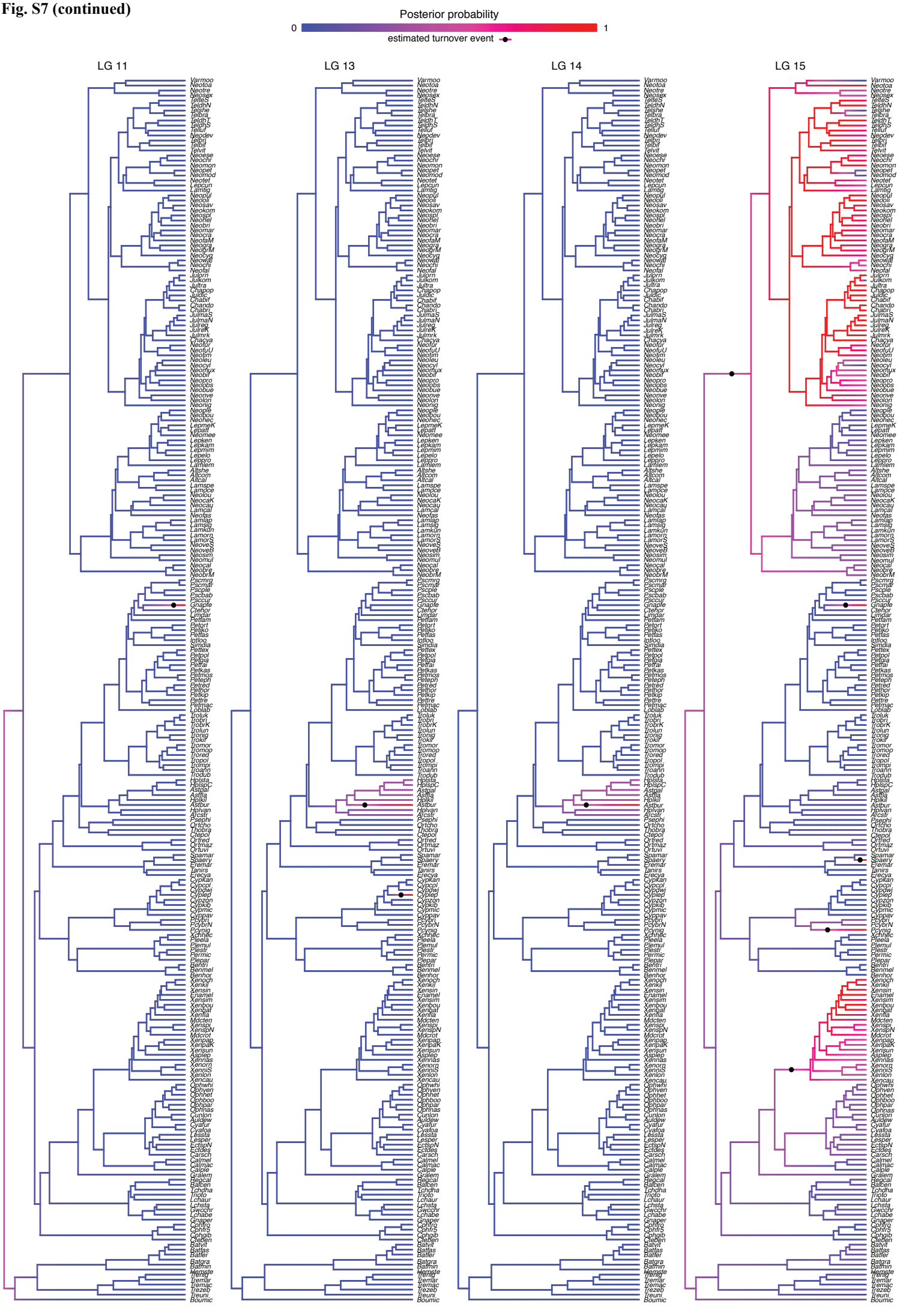
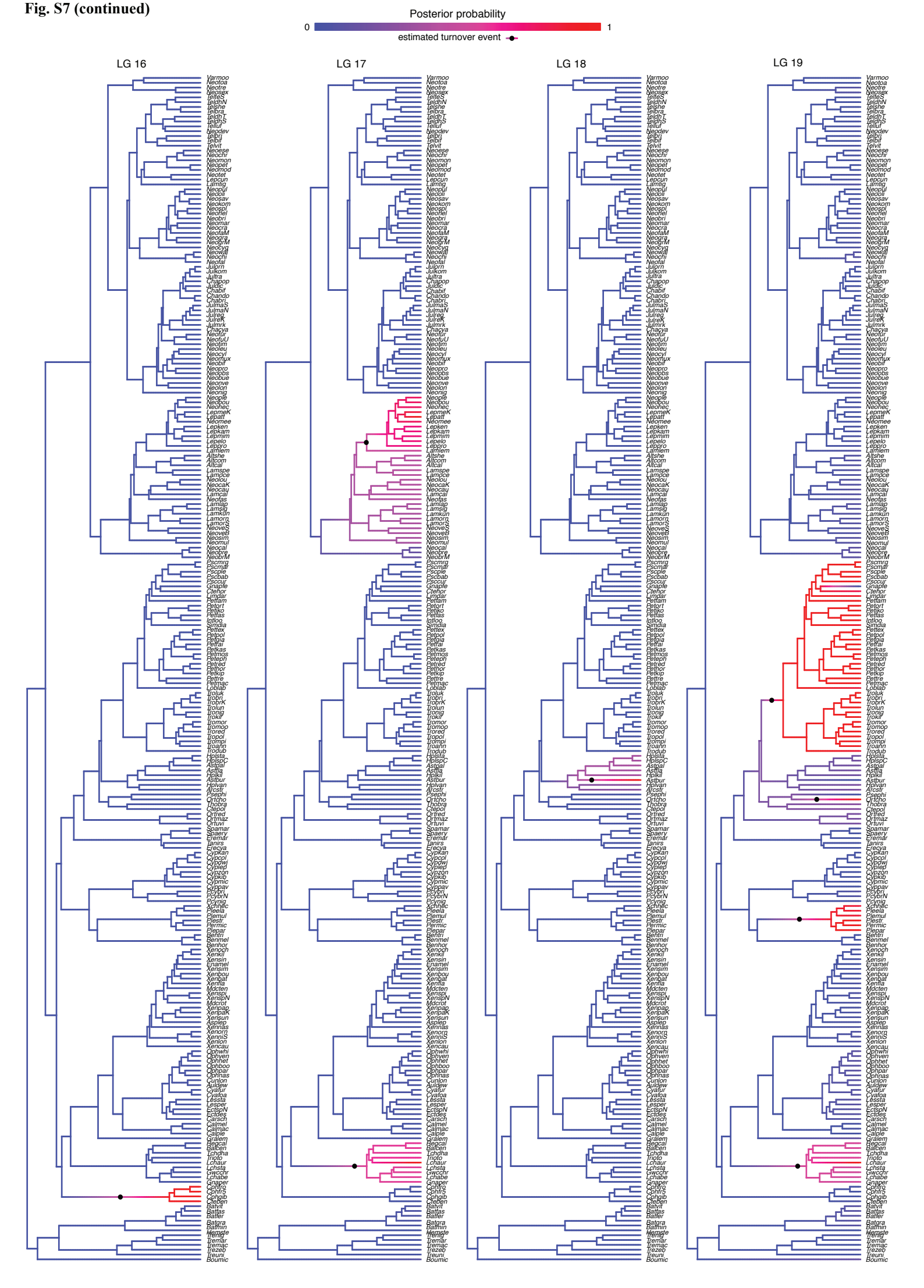


Fig. S7 (continued)



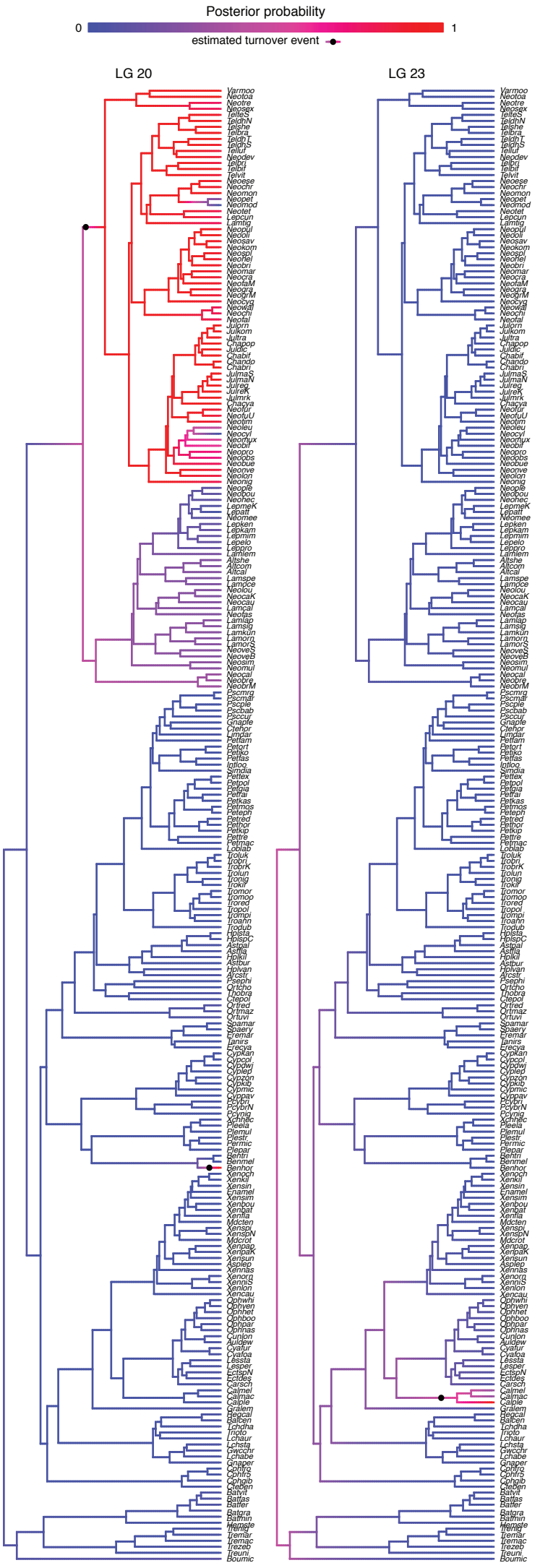


Fig. S7. Sex chromosome turnover points. Sex chromosome turnover points visualized per sex-linked LG derived from ancestral state reconstructions based on stochastic character mappings (using the permissive dataset).

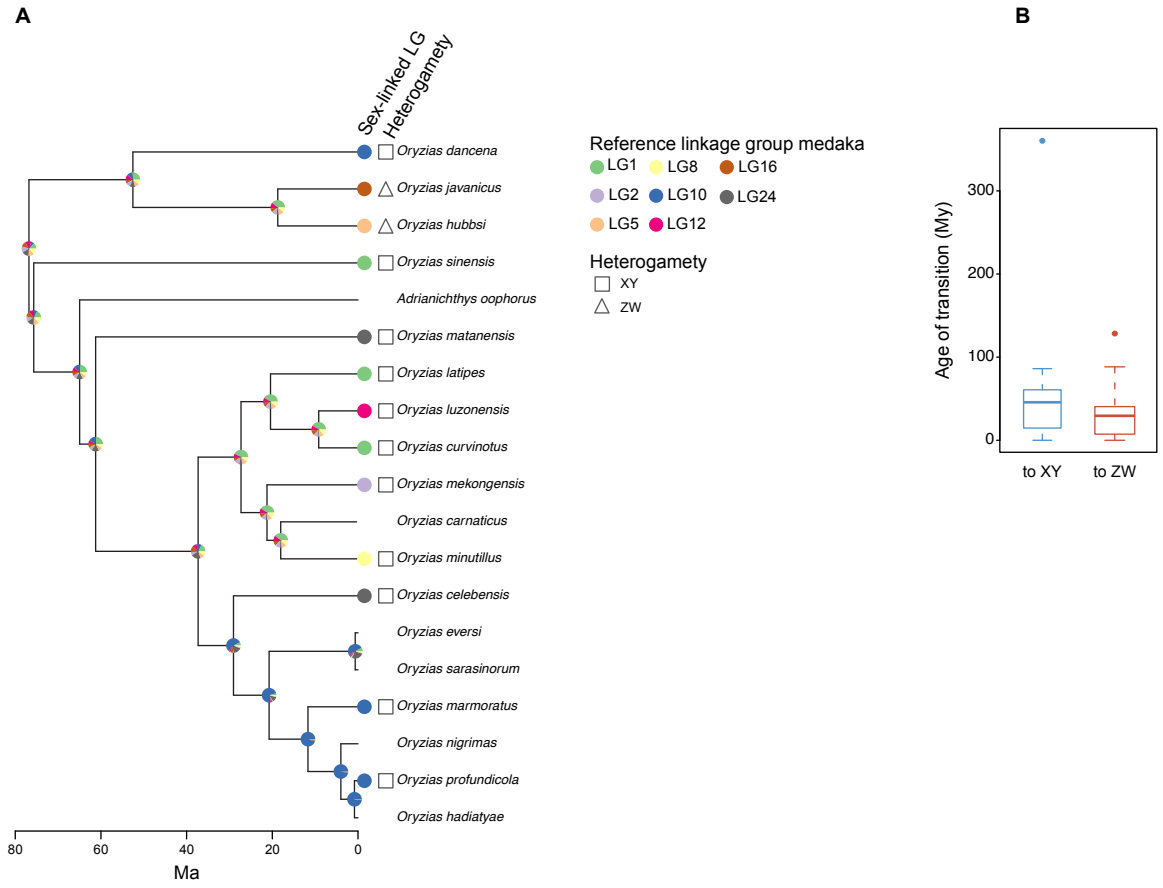


Fig. S8. Sex chromosome evolution in ricefishes. (**A**) Stochastic character mapping for sex chromosomes in ricefishes. Colored circles at tips represent sex-associated LGs with respect to the medaka (*Oryzias latipes*) genome. Pie charts at nodes represent the probability for an LG being a sex chromosome at this time. Heterogametic status is indicated at tips (triangle: ZW, square: XY). (**B**) Age of heterogametic transitions in ray-finned fishes. Transitions to ZW-systems are younger than transitions to XY-systems (Kruskal-Wallis rank sum test P = 0.014). Boxplot center lines represent the median, box limits the upper and lower quartiles, and whiskers the 1.5x interquartile rang, points represent outliers. The size of boxes within a plot is proportional to the number of transition events.

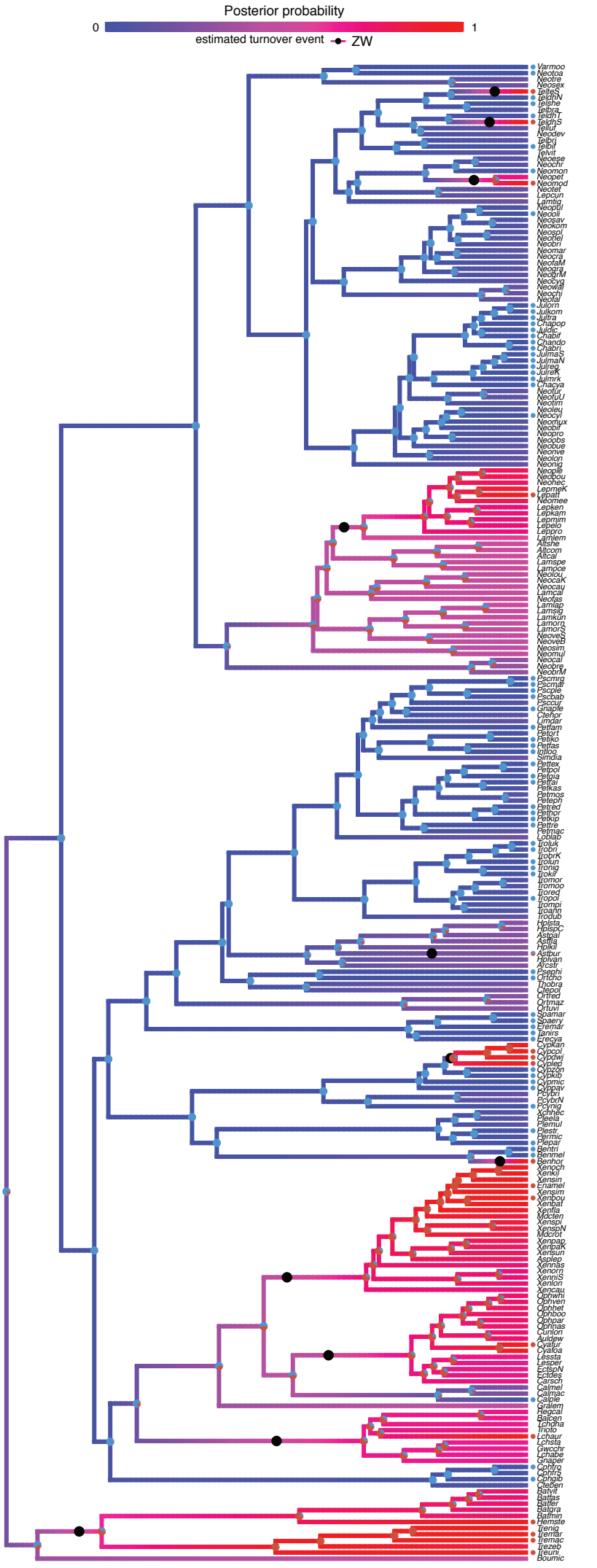


Fig. S9. Stochastic character mapping for heterogametic status and inference of heterogametic changes. Circles at the tips of the tree represent the identified heterogametic status (blue: XY, red: ZW). For species without a symbol, no sex chromosome could be identified. Pie charts at internal nodes represent the reconstructed heterogametic state in 1,000 simulations. Black points represent estimated heterogamety turnover points derived from ancestral state reconstructions based on stochastic character mappings.

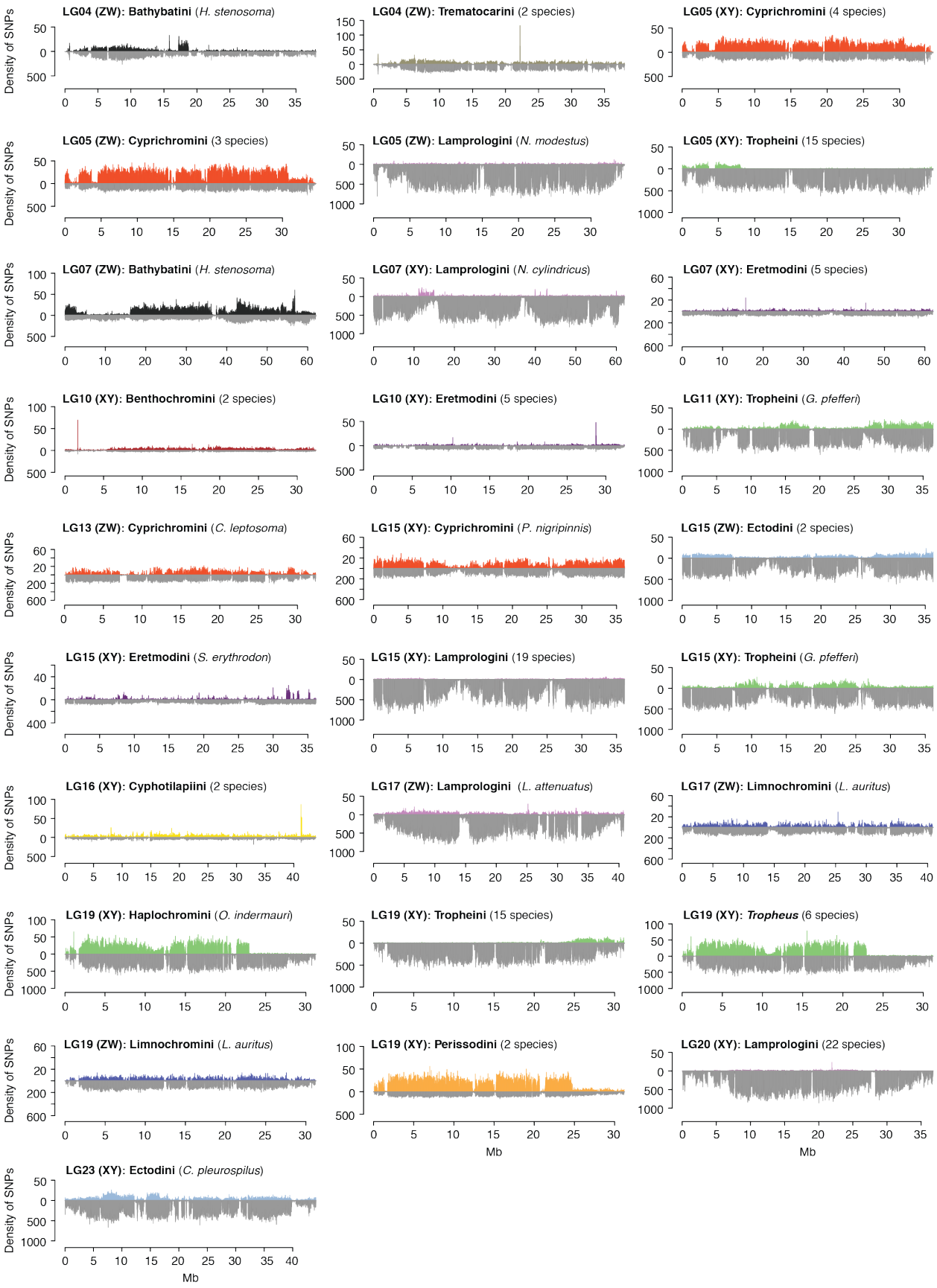


Fig. S10. Distribution of sex-differentiated SNPs along the sex-associated LGs. Each plot represents one identified sex chromosomal system. In each plot, values plotted in tribe colors (see Fig. 1) show the density of XY- or ZW-SNPs, values plotted in gray show the distribution of the total number of SNPs called for each window along the LG.

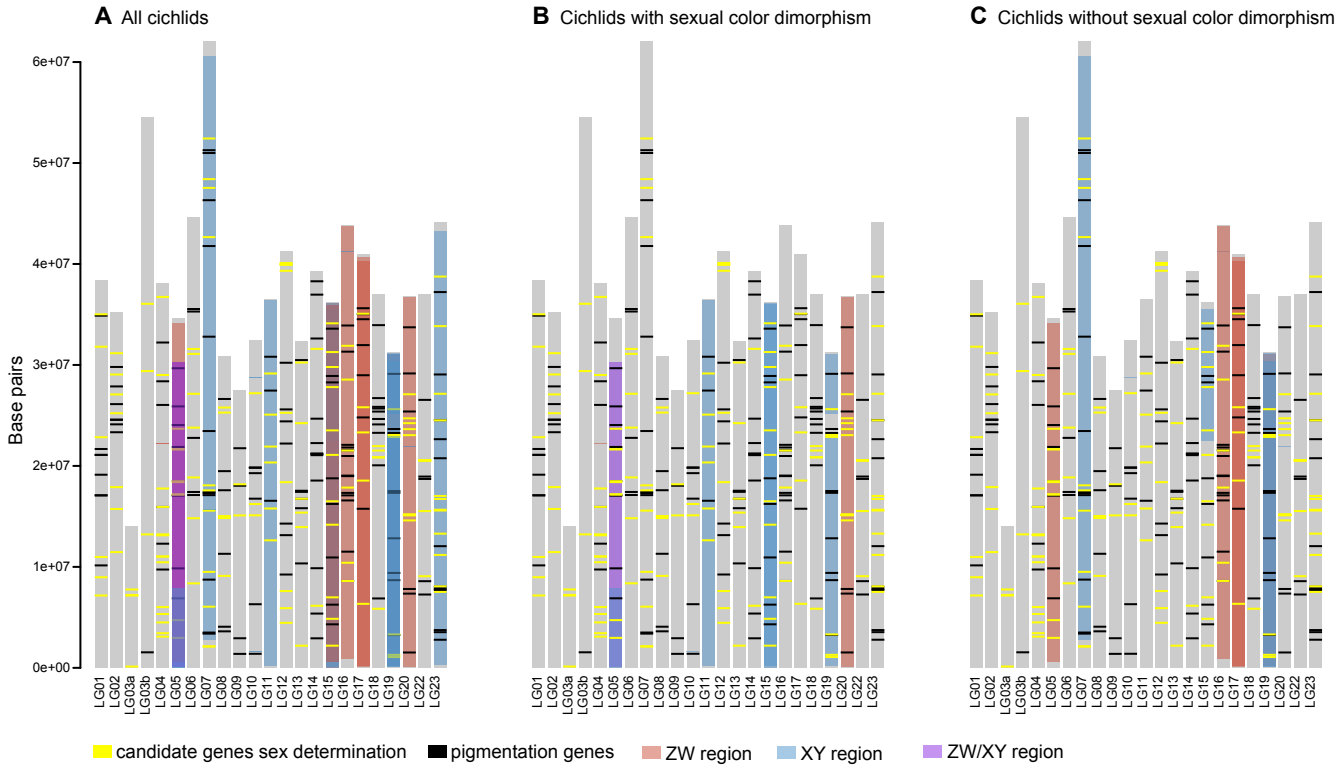


Fig. S11. Distribution of candidate genes for sex determination and pigmentation for (A) all cichlids, (B) cichlids with sexual color dimorphism, and (B) cichlids without sexual color dimorphism. Gray bars represent LGs and lines candidate genes for sex determination and pigmentation. Sex-differentiated regions are indicated with colored shadings.