Supplementary information

Evolutionary dynamics of enlarged sex chromosomes and novel pseudoautosomal regions in Sylvioidea songbirds

Hanna Sigeman, Simon Jacobsen Ellerstrand & Bengt Hansson

This supplementary information includes **Figures S1-S11**.

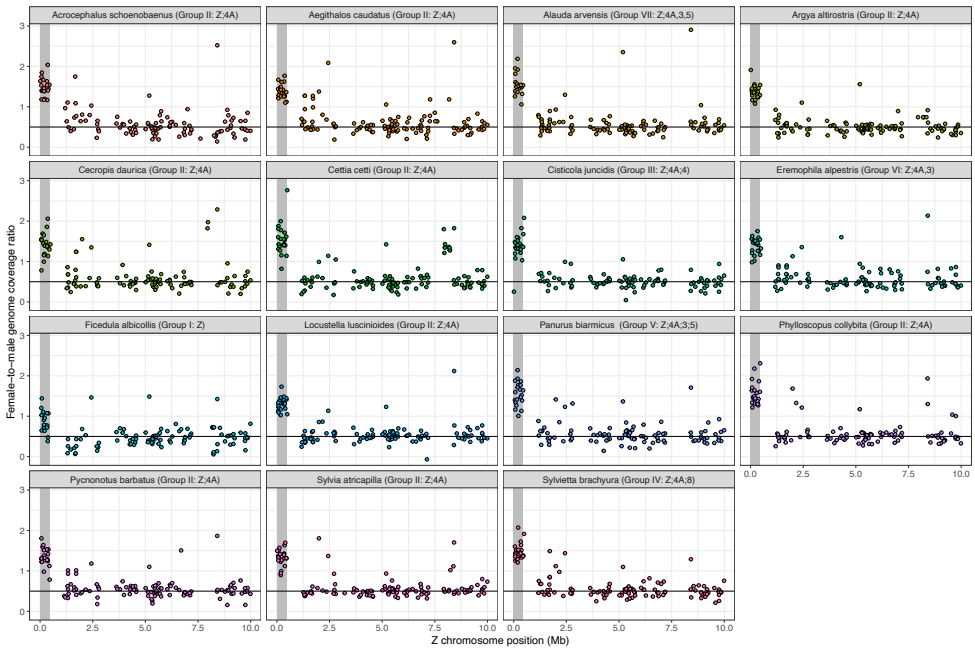

Figure S1. Female-to-male genome coverage values across 0-10 Mb on chromosome Z. The genes in the PAR (0-0.5 Mb; grey) have higher genome coverage values compared to the genes in the fully sex-linked part (0.5-10 Mb). This suggests that the PAR genes have both a Z and W gametolog.

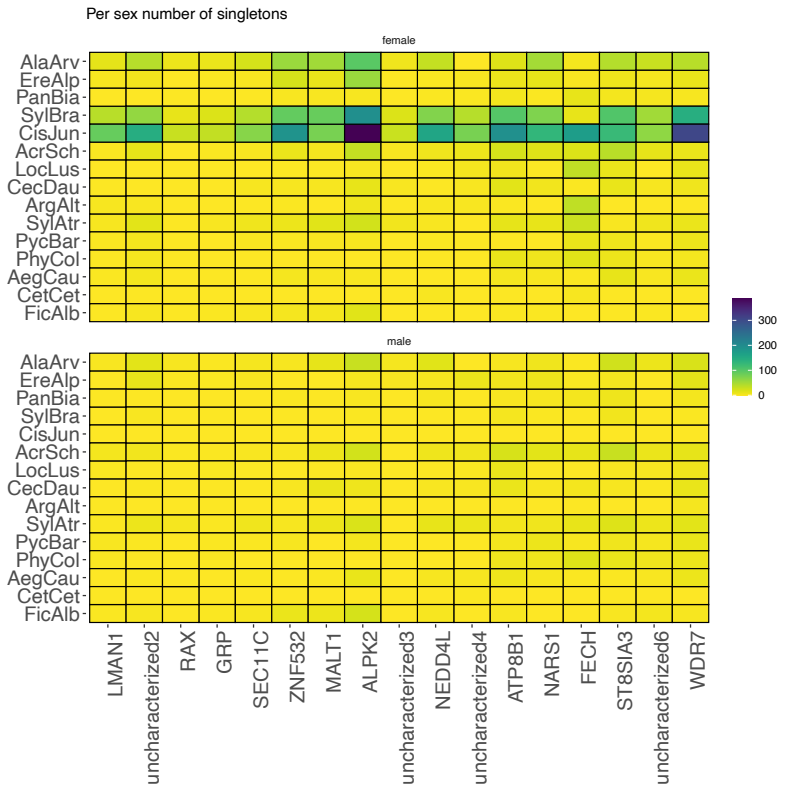

Figure S2. Number of singletons (i.e., genetic variants occuring only in one sample, and in heterozygous form) in females and males separately. Females had more singleton variants than males, indicative of sex-linkage, in three lineages; *Alauda* (AlaArv), *Sylvietta* (SylBra) and *Cisticola* (CisJun).

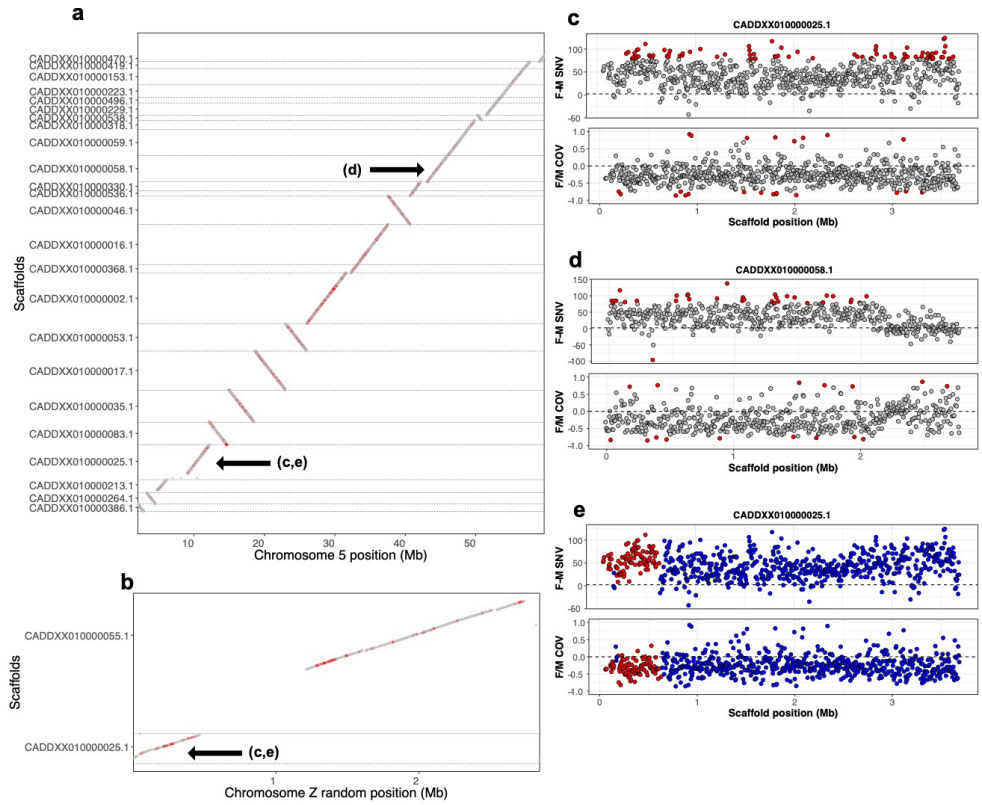

Figure S3. Synteny between *Alauda arvensis* scaffolds and *Taeniopygia guttata* (a) chromosome 5 and the scaffold (b) "Z_random", which contains the *T. guttata* PAR within the first ~0.5 Mb. Genome coverage and single nucleotide variation (SNV) are based on aligned WGS data from one male and one female *A. arvensis* individual. Data points in (a) and (b) are colored red if either the female-to-male genome coverage or SNV exceeds the genome-wide 95% CI, and grey is they are not. Highlighted in (a) and (b) are scaffolds informative of fusion points (c,e) or a novel PAR boundary (d) shown in Figure 3a in the Main text. (c,d) Female-to-male SNV (upper) and genome coverage ratios (lower). Red data points exceed the genome-wide 95% CI, grey data points do not. (e) This panel is based on the same data as (c), but colored according to synteny to zebra finch chromosomes (chromosome 5 in blue; scaffold Z_random in red).

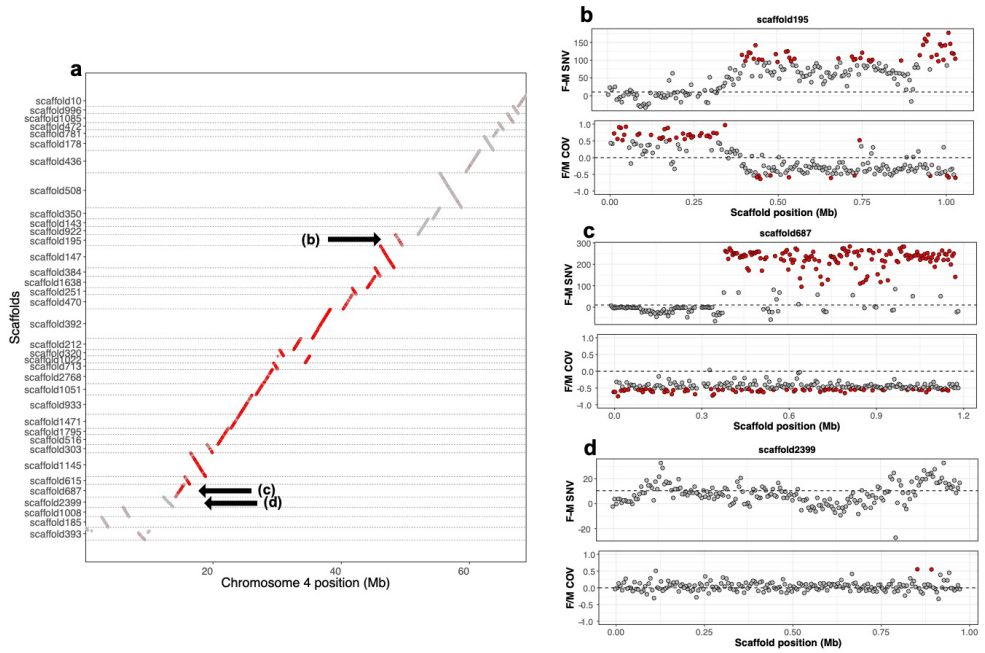

Figure S4. Synteny between *Cisticola juncidis* scaffolds and *Taeniopygia guttata* (a) chromosome 4. Genome coverage and single nucleotide variation (SNV) are based on aligned WGS data from one male and one female *C. juncidis* individual. Data points in (a) are colored red if either the female-to-male genome coverage or SNV exceeds the genome-wide 95% CI, and grey is they are not. Highlighted in (a) are scaffolds informative of (b) a novel PAR boundary or a pair of fission scaffolds (c,d). (b-d) Female-to-male SNV (upper) and genome coverage ratios (lower). Red data points exceed the genome-wide 95% CI, grey data points do not.

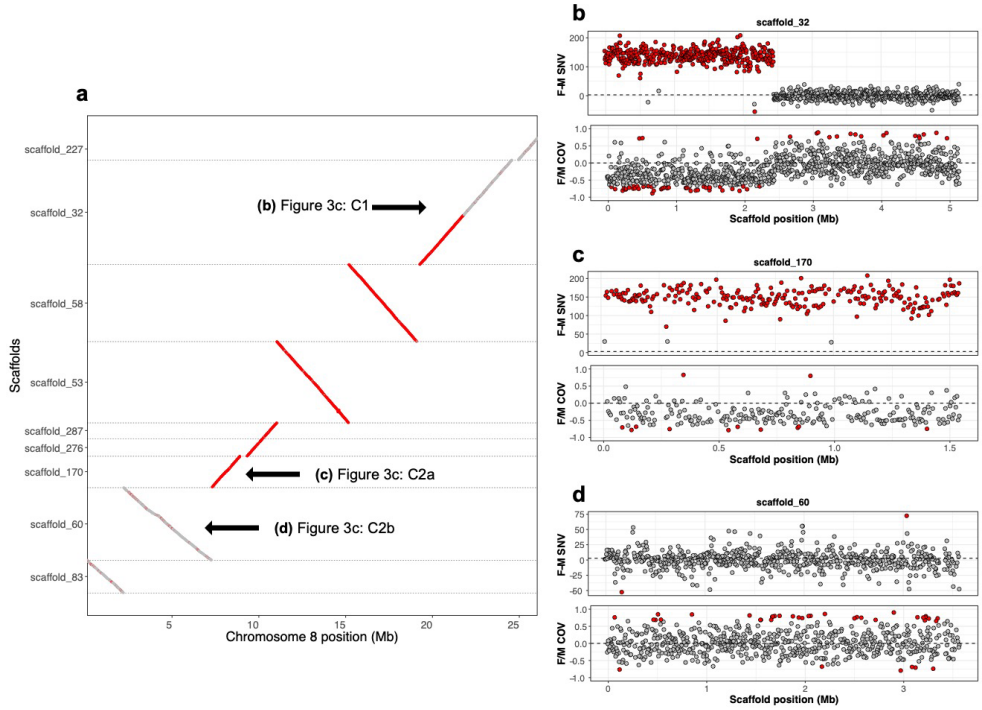

Figure S5. Synteny between *Sylvietta virens* scaffolds and *Taeniopygia guttata* (a) chromosome 8. Genome coverage and single nucleotide variation (SNV) are based on aligned WGS data from one male and one female *S. brachyura* individual. Data points in (a) are colored red if either the female-to-male genome coverage or SNV exceeds the genome-wide 95% CI, and grey is they are not. Highlighted in (a) are scaffolds informative of a novel PAR boundary (b) or fusion points (c,d) shown in Figure 3c in the Main text. (b-d) Female-to-male SNV (upper) and genome coverage ratios (lower). Red data points exceed the genome-wide 95% CI, grey data points do not.

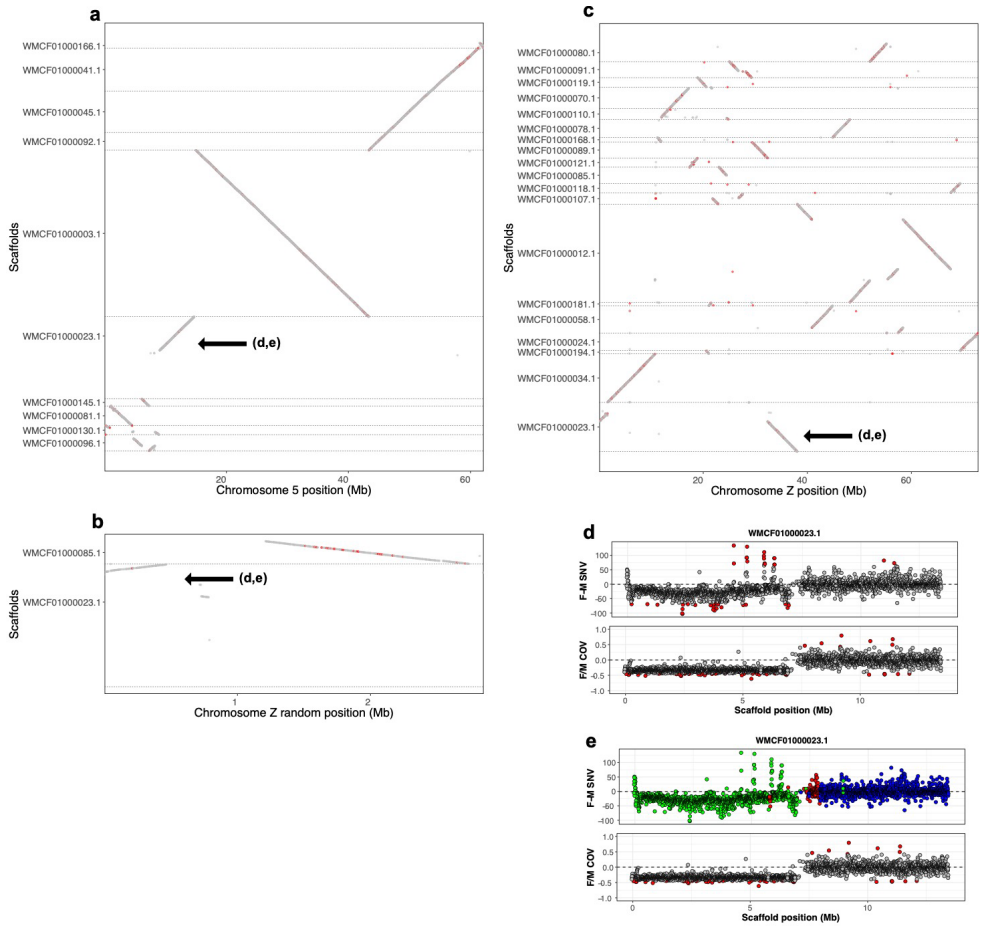

Figure S6. Synteny between *Eremophila alpestris* scaffolds and *Taeniopygia guttata* (a) chromosome 5 and the scaffold (b) "Z_random", which contains the *T. guttata* PAR within the first ~0.5 Mb and the rest of (c) chromosome Z. Genome coverage and single nucleotide variation (SNV) are based on aligned WGS data from one male and one female *E. alpestris* individual. Data points in (a-c) are colored red if either the female-to-male genome coverage or SNV exceeds the genome-wide 95% CI, and grey is they are not. Highlighted in (a-c) is the fusion scaffold shown in Figure 3b (d,e) in the Main text. (d) Female-to-male SNV (upper) and genome coverage ratios (lower). Red data points exceed the genome-wide 95% CI, grey data points do not. (e) This panel is based on the same data as (d), but colored according to synteny to zebra finch chromosomes (chromosome 5 in blue; scaffold Z_random in red, chromosome Z in green).

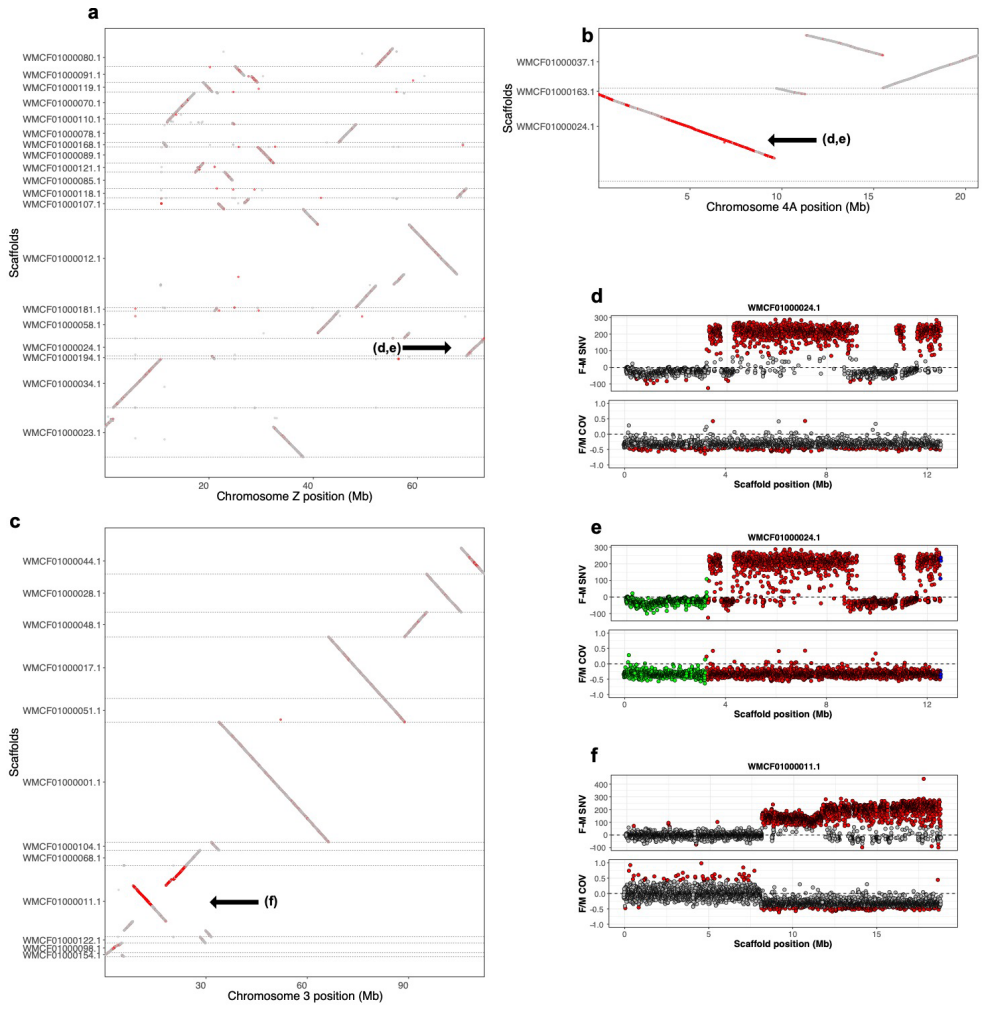

Figure S7. Synteny between *Eremophila alpestris* scaffolds and *Taeniopygia guttata* (a) chromosome Z, (b) chromosome 4A and (c) chromosome 3. Genome coverage and single nucleotide variation (SNV) are based on aligned WGS data from one male and one female *E. alpestris* individual. Data points in (a-c) are colored red if either the female-to-male genome coverage or SNV exceeds the genome-wide 95% CI, and grey is they are not. Highlighted in (a-c) are scaffolds informative of fusion sites (d,e) or (f) a novel PAR boundary. (d,f) Female-to-male SNV (upper) and genome coverage ratios (lower). Red data points exceed the genome-wide 95% CI, grey data points do not. (e) This panel is based on the same data as (d), but colored according to synteny to zebra finch chromosomes (chromosome 4A in red, chromosome Z in green and chromosome 3 in blue).

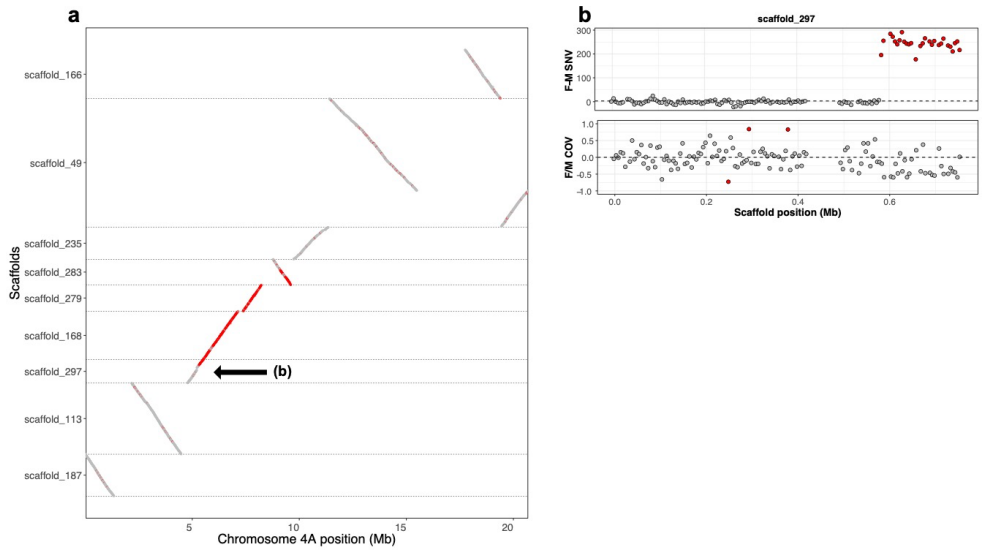

Figure S8. Synteny between *Sylvietta virens* scaffolds and *Taeniopygia guttata* (a) chromosome 4A. Genome coverage and single nucleotide variation (SNV) are based on aligned WGS data from one male and one female *S. brachyura* individual. Data points in (a) are colored red if either the female-to-male genome coverage or SNV exceeds the genome-wide 95% CI, and grey is they are not. Highlighted in (a) is a scaffold informative of (b) a novel PAR boundary. (b) Female-to-male SNV (upper) and genome coverage ratios (lower). Red data points exceed the genome-wide 95% CI, grey data points do not.

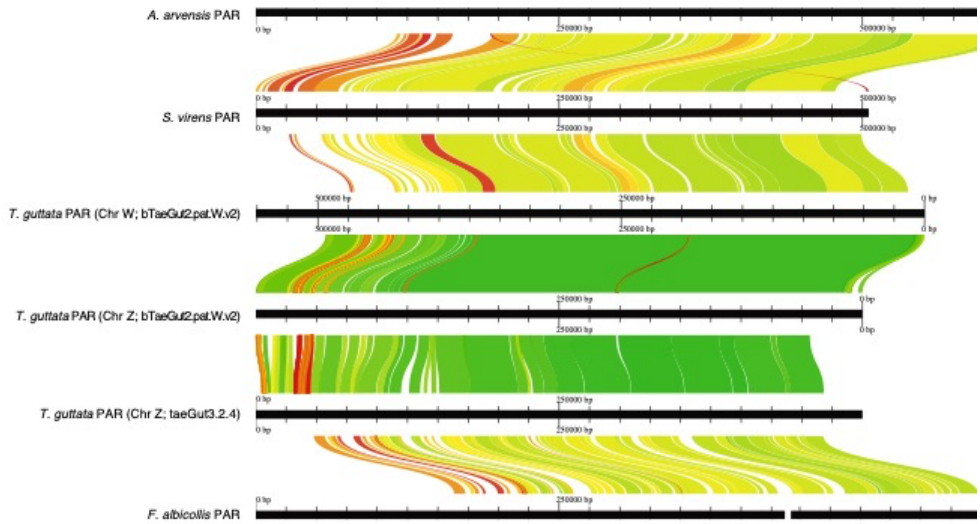

Figure S9. Synteny plot produced from PAR scaffolds of different species (Table S6), using the program AliTV (Ankenbrand et al. 2017)

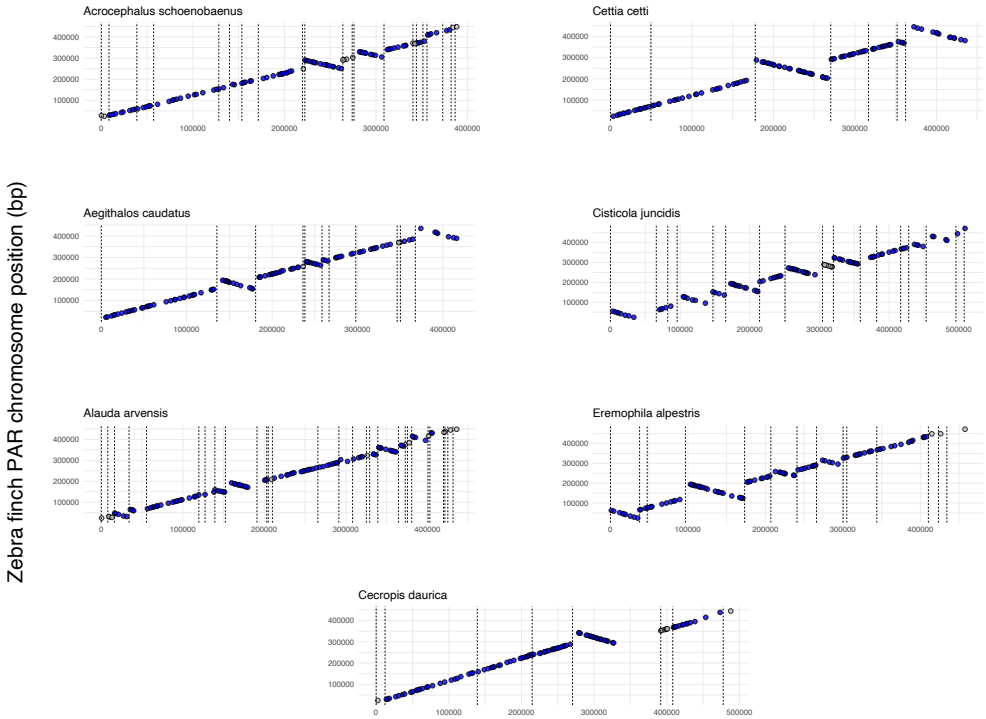

Figure S10a. Synteny plots of scaffolds in 7 of the 14 Sylvioidea reference genomes (assembled from the male samples in Supplementary Table S3) and the zebra finch PAR. Scaffold boundaries are marked with dashed lines. Data points colored in blue are positioned on scaffolds with PAR genes. Grey data points are positioned on scaffolds that contain no genes.

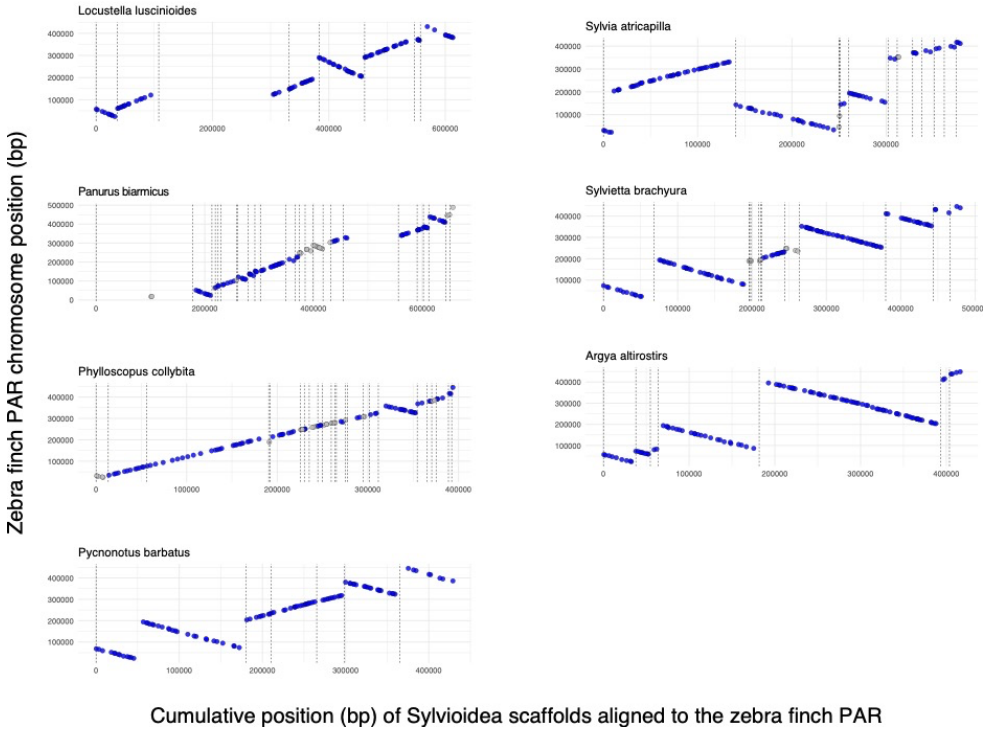

Figure S10b. Synteny plots of scaffolds in 7 of the 14 Sylvioidea reference genomes (assembled from the male samples in Supplementary Table S3) and the zebra finch PAR. Scaffold boundaries are marked with dashed lines. Data points colored in blue are positioned on scaffolds with PAR genes. Grey data points are positioned on scaffolds that contain no genes.

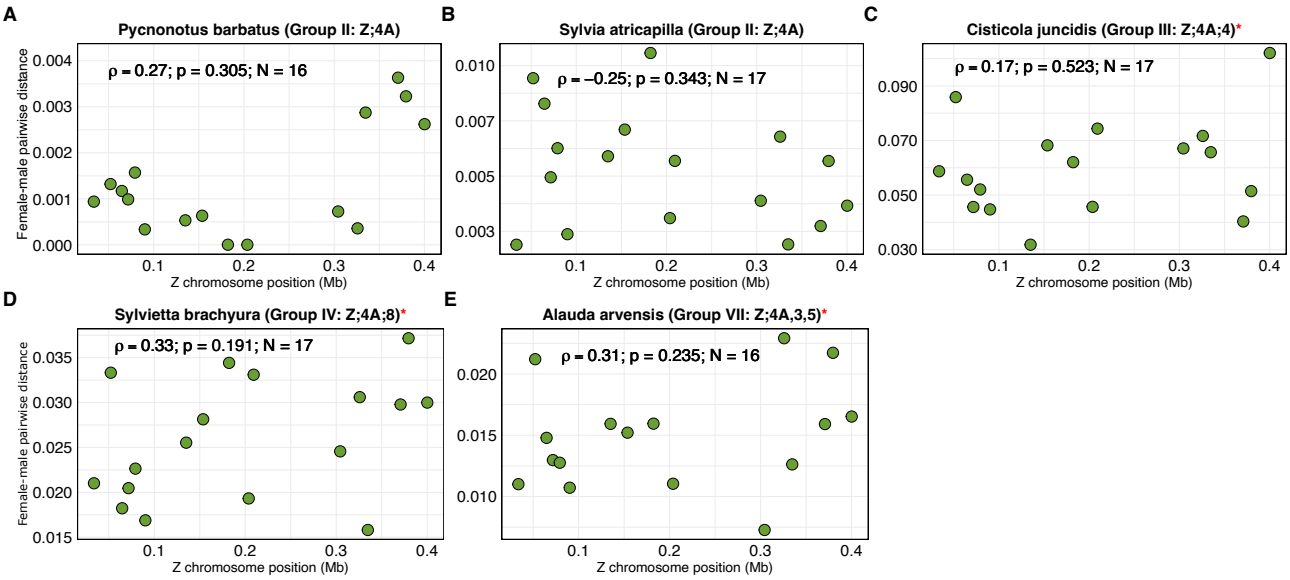

Figure S11. Y-axis: Female-to-male (Z-to-W) PAR gene sequence divergence Y-axis: Chromosome position along the PAR in *Taeniopygia guttata*. The PAB is located at 468 kb. Shown here are species with either established non-collinearity to the gene order in *T. guttata* (a-b) and species where the ancestral PAR is no longer recombining (c-e). See Figure 5 in the Main Text for all other species.

References

Ankenbrand MJ, Hohlfeld S, Hackl T, Förster F. 2017. AliTV—interactive visualization of whole genome comparisons. PeerJ Computer Science 3:e116