

Supplementary information

XYZWfinder: a snakemake pipeline for detecting and visualising sex chromosomes using whole-genome sequencing data

Hanna Sigeman, Bella Sinclair and Bengt Hansson

This supplementary information includes **Supplementary Methods**, **Tables S1-S5**, **Figure S1-S19**.

Supplementary Methods

Expected sex chromosomes/inversion polymorphism scaffold in the nine studied species

Here we give brief information about the current knowledge of the sex chromosome system in each of the studied species, or inversion polymorphism in the case of the ruff. In species where a synteny-species reference genome was used to anchor the scaffolds/chromosomes from the study-species reference genome, we list what chromosomes are expected to show signs of sex-linkage in each synteny-species. These expectations are based either on results from previous studies or based on known syntenies between species (as determined by <https://www.genomicus.biologie.ens.fr/>).

1. *Alauda arvensis* (Eurasian skylark):

Sex chromosome system: Neo-sex chromosome system with shared homology to zebra finch (*Taeniopygia guttata*) chromosomes Z, 3, 4A and 5 (Sigeman et al. 2019; Dierickx et al. 2020).

Details on XYZWfinder analyses: The Eurasian skylark WGS samples (n = 2; Table S1) were analysed using a (highly fragmented) study-species reference genome constructed from the male WGS sample (Table S1, S3; see Sigeman et al. 2019). We ran the XYZWfinder pipeline three times. The first time (a) we used the “consensus reference genome” option, to ensure equal mapping success to the reference genome between sexes (with snakemake-no-synteny). The second and third time, we ran the pipeline using this consensus reference genome as input to the pipeline, while using (b) the zebra finch (*Taeniopygia guttata*) or (c) chicken (*Gallus gallus*) as synteny-species reference genomes (snakemake-synteny).

Expected sex-linked regions:

Synteny-species 1 (zebra finch): chromosomes Z, 3, 4A and 5 (Sigeman et al. 2019; Dierickx et al. 2020).

Synteny-species 2: (chicken): chromosomes Z, 3, 4 and 5 (Based on synteny between zebra finch and chicken genomes: <https://www.genomicus.biologie.ens.fr/>.

2. *Alouatta palliata* (mantled howler monkey):

Sex chromosome system: Multiple sex chromosome system (X_1X_2Y ; Ma et al. 1975). The X_2 chromosome in *A. caraya* (a relative of *A. palliata*) share homology to human chromosomes 3 and 15 (Solari & Rahn 2005).

Details on XYZWfinder analyses: The mantled howler monkey WGS samples (n = 4; Table S1) were analysed using a fragmented study-species reference genome (Table S3). We first ran the pipeline using all samples (n = 4) and with the (a) snakemake-no-synteny option. Then, we used the (b) human (*Homo sapiens*) and (c) meerkat (*Suricata suricatta*) reference genomes (Table S3) as synteny-species reference genomes (snakemake-synteny). We then ran the pipeline (again with human as a synteny-species) using (d) only one sample of each sex (n = 2) and lastly (e) using only 50 % of the sequences from these WGS files (n = 2).

Expected sex-linked regions:

Study-species: Unknown

Synteny-species 1 (human): chromosomes X, 3, 15 (Solari & Rahn 2005).

Synteny-species 2: (meerkat): chromosomes X, 5, 9 (Based on synteny between human and meerkat genomes: <https://www.genomicus.biologie.ens.fr/>).

3. Anolis carolinensis (green anole):

Sex chromosome system: Small XY-system (Alföldi et al. 2011), chromosome LGb.

Details on XYZWfinder analyses: The green anole samples WGS samples (n = 2; Table S1) were analysed using the chromosome-level reference genome of this species (Table S3) as a study-species reference genome (with snakemake-no-synteny).

Expected sex-linked regions:

Study-species: chromosome LGb (Alföldi et al. 2011).

4. Calidris pugnax (ruff)

Inversion polymorphism: Scaffold28 contain the 4.5 Mb large inversion polymorphism controlling male phenotypes in the ruff (Lamichhaney et al. 2016). Scaffold28 share homology to chicken chromosome 11 (Lamichhaney et al. 2016).

Details on XYZWfinder analyses: The ruff samples (n = 3; Table S1) were analysed using the scaffold-level reference genome of the ruff (Table S3) as a study-species reference genome. The pipeline was run twice; first (a) without a synteny-species reference genome (snakemake-no-synteny) and then (b) using the chicken as a synteny-species reference genome (snakemake-synteny).

Expected inversion polymorphic region:

Study-species: scaffold28/NW_015090842.1 (Lamichhaney et al. 2016).

Synteny-species (chicken): chromosome 11 (Lamichhaney et al. 2016).

5. *Drosophila miranda* (fruit fly)

Sex chromosome system: Neo-XY system (Zhou & Bachtrog 2012) involving chromosomes XL, XR (ancestral sex chromosome arms) and Muller-C (fusion).

Details on XYZWfinder analyses: The fruit fly samples ($n = 2$; Table S1) were analysed using the chromosome-level reference genome of this species (Table S3) as a study-species reference genome (snakemake-no-synteny).

Expected sex-linked regions:

Study-species: Chromosomes XL (NC_030302.1), XR (NC_030303.1) and Muller-C (chromosome 3; NC_030305.1) (Zhou & Bachtrog 2012).

6. *Nothobranchius furzeri* (turquoise killifish)

Sex chromosome system: Small relatively undifferentiated XY-system (Reichwald et al. 2015).

Details on XYZWfinder analyses: The turquoise killifish WGS samples ($n = 3$; Table S1) were analysed using a chromosome-level study-species reference genome (Table S3). We first ran the pipeline (snakemake-no-synteny) using all samples ($n = 3$) and then using only one sample of each sex ($n = 2$).

Expected sex-linked regions:

Study-species: Chromosome sgr05 (Reichwald et al. 2015).

7. *Ornithorhynchus anatinus* (platypus)

Sex chromosome system: XY-system consisting of five X chromosomes and five Y chromosomes: X₁₋₅Y₁₋₅ (Grützner et al. 2004).

Details on XYZWfinder analyses: The platypus WGS samples ($n = 2$; Table S1) were analysed (snakemake-no-synteny) using a chromosome-level study-species reference genome (Table S3).

Expected sex-linked regions:

Study-species: Chromosome X_{1-X₅} (Grützner et al. 2004).

8. *Poecilia reticulata* (guppy)

Sex chromosome system: Extremely undifferentiated XY-system (Künstner et al. 2015; Wright et al. 2017), LG12/chr12.

Details on XYZWfinder analyses: The guppy WGS samples (n = 4; Table S1) were analysed (snakemake-no-synteny) using a chromosome-level study-species reference genome (Table S3).

Expected sex-linked regions:

Study-species: Chromosome 12/LG12 (Künstner et al. 2015; Wright et al. 2017).

9. *Pogona vitticeps* (*central bearded dragon*)

Sex chromosome system: Micro-ZW system. Share partial homology to chicken microchromosomes 17 and 23 (Deakin et al. 2016).

Details on XYZWfinder analyses: The central bearded dragon samples (n = 3; Table S1) were analysed using the scaffold-level reference genome of the central bearded dragon (Table S3) as a study-species reference genome. The pipeline was run using the chicken as a synteny-species reference genome (snakemake-synteny).

Expected sex-linked regions:

Synteny-species (chicken): Chromosome 23 and 17 (Deakin et al. 2016).

References:

- Alföldi J, Di Palma F, Grabherr M, et al. 2011. The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature*, **477**, 587–591.
- Deakin JE, Edwards MJ, Patel H, et al. 2016. Anchoring genome sequence to chromosomes of the central bearded dragon (*Pogona vitticeps*) enables reconstruction of ancestral squamate macrochromosomes and identifies sequence content of the Z chromosome. *BMC Genomics*, **17**:447.
- Dierickx EG, Sin SYW, van Veelen HPJ, Brooke ML, Liu Y, Edwards SV, Martin SH. 2020. Genetic diversity, demographic history and neo-sex chromosomes in the Critically Endangered Raso lark. *Proceedings of the Royal Society B*, **287**:20192613.
- Grützner F, Rens W, Tsend-Ayush E, El-Mogharbel N, O'Brien PC, Jones RC, et al. 2004. In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature*, **432**, 913–917.
- Künstner A, Hoffmann M, Fraser BA, Kottler VA, Sharma E, et al. 2016. The Genome of the Trinidadian Guppy, *Poecilia reticulata*, and Variation in the Guanapo Population. *PLOS ONE*, **11**: e0169087
- Lamichhaney S, Fan G, Widemo F, et al. 2016. Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nature Genetics*, **48**, 84–88.

- Ma N, S, F, Jones T, C, Thorington R, W, Miller A, Morgan L. 1975. Y-Autosome Translocation in the Howler Monkey. *Journal of Medical Primatology*, **4**:299-307.
- Reichwald K, Petzold A, Koch P, et al. 2015. Insights into Sex Chromosome Evolution and Aging from the Genome of a Short-Lived Fish. *Cell*, **163**:1527-1538.
- Sigeman H, Ponnikas S, Chauhan P, Dierickx E, Brooke MdeL, Hansson B. 2019. Repeated sex chromosome evolution in vertebrates supported by expanded avian sex chromosomes. *Proceedings of the Royal Society B*, **286**, 20192051.
- Solari A, J, Rahn M, 2005. I: Fine structure and meiotic behaviour of the male multiple sex chromosomes in the genus Alouatta. *Cytogenetic and Genome Research*, **108**:262-267.
- Wright AE, Darolti I, Bloch NI, Oostra V, Sandkam B, Buechel SD, et al. 2017. Convergent recombination suppression suggests role of sexual selection in guppy sex chromosome formation. *Nature Communications*, **8**, 14251.
- Zhou Q, & Bachtrog D. 2012. Sex-specific adaptation drives early sex chromosome evolution in Drosophila. *Science*, **337**, 341–345.

Phylopic/Wikimedia commons credits:

The guppy silhouette (Figure 5g) was created by Josefine Bohr Brask and shared via phylopic.org under a CC By 3.0 license (<https://creativecommons.org/licenses/by/3.0/>). Not modified from its original state.

The ruff silhouette (Figure 5i) was created by Alexandre Vong and shared via phylopic.org under a CC By 3.0 license (<https://creativecommons.org/licenses/by/3.0/>). Not modified from its original state.

Silhouettes in Figure 5a-c,e,f,h were downloaded from phylopic.org and were without copyright restrictions.

The silhouette in Figure 5d is from Wikimedia commons.

Supplementary Tables

Table S1. WGS paired-end data used for analyses. SRA files from the same BioSample were merged prior to analyses.

Species	Common name	Sex	Heterogamety	Alias	SRA code	BioSample
<i>Alauda arvensis</i>	Eurasian skylark	female	Heterogametic	A_arvensis_F	SRR10340221	SAMN13107494
<i>Alauda arvensis</i>	Eurasian skylark	male	Homogametic	A_arvensis_M	SRR10340220	SAMN13107495
<i>Alouatta palliata</i>	mantled howler monkey	female	Homogametic	A_palliata_F1	SRR9655168	SAMN10407156
<i>Alouatta palliata</i>	mantled howler monkey	female	Homogametic	A_palliata_F2	SRR9655169	SAMN10407140
<i>Alouatta palliata</i>	mantled howler monkey	male	Heterogametic	A_palliata_M1	SRR9655170	SAMN10407285
<i>Alouatta palliata</i>	mantled howler monkey	male	Heterogametic	A_palliata_M2	SRR9655171	SAMN10407294
<i>Anolis carolinensis</i>	green anole	female	Homogametic	A_carolinensis_F	SRR5508303	SAMN06891217
<i>Anolis carolinensis</i>	green anole	male	Heterogametic	A_carolinensis_M	SRR5508304	SAMN06891216
<i>Calidris pugnax</i>	ruff	faeder/male	Heterozygotic inversion	C_pugnax_F1	ERR1001621	SAMEA3522198
<i>Calidris pugnax</i>	ruff	resident/male	Homozygotic inversion	C_pugnax_R	ERR1001618	SAMEA3522199
<i>Calidris pugnax</i>	ruff	faeder/male	Heterozygotic inversion	C_pugnax_F2	ERR1001620	SAMEA3522198
<i>Drosophila miranda</i>	fruit fly	female	Homogametic	D_miranda_F	SRR1738162	SAMN03220544
<i>Drosophila miranda</i>	fruit fly	male	Heterogametic	D_miranda_M	SRR1738163	SAMN03220545
<i>Nothobranchius furzeri</i>	turquoise killifish	female	Homogametic	N_furzeri_F	ERR583470	SAMEA2698544
<i>Nothobranchius furzeri</i>	turquoise killifish	male	Heterogametic	N_furzeri_M1	ERR583467	SAMEA2698541
<i>Nothobranchius furzeri</i>	turquoise killifish	male	Heterogametic	N_furzeri_M2	ERR583468	SAMEA2698543
<i>Ornithorhynchus anatinus</i>	platypus	male	Heterogametic	O_anatinus_M	ERR2298705	SAMEA104585244
<i>Ornithorhynchus anatinus</i>	platypus	female	Homogametic	O_anatinus_F	ERR2298704	SAMEA104585243
<i>Poecilia reticulata</i>	guppy	female	Homogametic	P_reticulata_F1	SRR5034743	SRR5034743
<i>Poecilia reticulata</i>	guppy	female	Homogametic	P_reticulata_F2	SRR5034730	SAMN06036146
<i>Poecilia reticulata</i>	guppy	male	Heterogametic	P_reticulata_M1	SRR5034745	SAMN06036147

<i>Poecilia reticulata</i>	guppy	male	Heterogametic	P_reticulata_M2	SRR5034738	SAMN06036148
<i>Pogona vitticeps</i>	central bearded dragon	ZW female	Heterogametic	P_vitticeps_F_ZW*	ERR409919	SAMEA2300449
<i>Pogona vitticeps</i>	central bearded dragon	ZW female	Heterogametic	P_vitticeps_F_ZW*	ERR409918	SAMEA2300449
<i>Pogona vitticeps</i>	central bearded dragon	ZZ male	Homogametic	P_vitticeps_M_ZZ*	ERR409943	SAMEA2300447
<i>Pogona vitticeps</i>	central bearded dragon	ZZ male	Homogametic	P_vitticeps_M_ZZ*	ERR409944	SAMEA2300447
<i>Pogona vitticeps</i>	central bearded dragon	ZZ male	Homogametic	P_vitticeps_M_ZZ*	ERR409945	SAMEA2300447
<i>Pogona vitticeps</i>	central bearded dragon	ZZ female	Homogametic	P_vitticeps_F_ZZ*	ERR409939	SAMEA2300448
<i>Pogona vitticeps</i>	central bearded dragon	ZZ female	Homogametic	P_vitticeps_F_ZZ*	ERR409938	SAMEA2300448

* WGS files originating from the same individual (BioSample) were merged prior to alignment.

Table S2. Settings used to trim each WGS data file using trimmomatic (see Methods).

Species	Alias	Trimming settings
<i>Alauda arvensis</i>	A_arvensis_F	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Alauda arvensis</i>	A_arvensis_M	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Alouatta palliata</i>	A_palliata_F1	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Alouatta palliata</i>	A_palliata_F2	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Alouatta palliata</i>	A_palliata_M1	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Alouatta palliata</i>	A_palliata_M2	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Anolis carolinensis</i>	A_carolinensis_F	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Anolis carolinensis</i>	A_carolinensis_M	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Calidris pugnax</i>	C_pugnax_F1	TruSeq3-PE.fa:2:30:10 LEADING:15 SLIDINGWINDOW:4:25 MINLEN:90
<i>Calidris pugnax</i>	C_pugnax_R	TruSeq3-PE.fa:2:30:10 LEADING:15 SLIDINGWINDOW:4:25 MINLEN:90
<i>Calidris pugnax</i>	C_pugnax_F2	TruSeq3-PE.fa:2:30:10 LEADING:15 SLIDINGWINDOW:4:25 MINLEN:90

<i>Drosophila miranda</i>	D_miranda_F	TruSeq3-PE.fa:2:30:10 LEADING:20 SLIDINGWINDOW:4:30 MINLEN:50
<i>Drosophila miranda</i>	D_miranda_M	TruSeq3-PE.fa:2:30:10 LEADING:20 SLIDINGWINDOW:4:30 MINLEN:50
<i>Nothobranchius furzeri</i>	N_furzeri_F	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:60
<i>Nothobranchius furzeri</i>	N_furzeri_M1	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:60
<i>Nothobranchius furzeri</i>	N_furzeri_M2	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:60
<i>Ornithorhynchus anatinus</i>	O_anatinus_M	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Ornithorhynchus anatinus</i>	O_anatinus_F	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Poecilia reticulata</i>	P_reticulata_F1	TruSeq3-PE.fa:2:30:10 LEADING:15 SLIDINGWINDOW:4:25 MINLEN:50
<i>Poecilia reticulata</i>	P_reticulata_F2	TruSeq3-PE.fa:2:30:10 LEADING:15 SLIDINGWINDOW:4:25 MINLEN:50
<i>Poecilia reticulata</i>	P_reticulata_M1	TruSeq3-PE.fa:2:30:10 LEADING:15 SLIDINGWINDOW:4:25 MINLEN:50
<i>Poecilia reticulata</i>	P_reticulata_M2	TruSeq3-PE.fa:2:30:10 LEADING:15 SLIDINGWINDOW:4:25 MINLEN:50
<i>Pogona vitticeps</i>	P_vitticeps_F_ZW	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Pogona vitticeps</i>	P_vitticeps_F_ZW	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Pogona vitticeps</i>	P_vitticeps_M_ZZ	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Pogona vitticeps</i>	P_vitticeps_M_ZZ	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Pogona vitticeps</i>	P_vitticeps_M_ZZ	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Pogona vitticeps</i>	P_vitticeps_F_ZZ	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90
<i>Pogona vitticeps</i>	P_vitticeps_F_ZZ	TruSeq3-PE.fa:2:30:10 LEADING:15 TRAILING:30 SLIDINGWINDOW:4:20 MINLEN:90

Table S3. Reference genomes used as either study-species or synteny-species reference genome. Only reference genomes from homogametic samples were used as study-species reference genomes.

Species	Common name	Assembly ID	GenBank accession	Sex	Heterogamety
<i>Alauda arvensis</i>	Eurasian skylark	skylark_min1kb	NA (https://doi.org/10.5061/dryad.95x69p8f9)	male	Homogametic
<i>Alouatta palliata</i>	mantled howler monkey	AloPal_v1_BIUU	GCA_004027835.1	female	Homogametic

<i>Anolis carolinensis</i>	green anole	AnoCar2.0	GCA_000090745.2	female	Homogametic
<i>Calidris pugnax</i>	ruff	ASM143184v1	GCA_001431845.1	not collected	NA
<i>Drosophila miranda</i>	fruit fly	DroMir_2.2	GCA_000269505.2	female	Homogametic
<i>Gallus gallus</i>	chicken	GRCg6a	GCA_000002315.5	female	Heterogametic
<i>Homo sapiens</i>	human	GRCh38.p13	GCA_000001405.28	male	Heterogametic
<i>Nothobranchius furzeri</i>	turquoise killifish	Nfu_20140520	GCA_001465895.2	female	Homogametic
<i>Ornithorhynchus anatinus</i>	platypus	ASM227v2/ornAna2	GCA_000002275.2	female	Homogametic
<i>Poecilia reticulata</i>	guppy	Guppy_female_1.0+MT	GCA_000633615.2	female	Homogametic
<i>Pogona vitticeps</i>	central bearded dragon	pvi1.1	GCA_900067755.1	male	Homogametic
<i>Suricata suricatta</i>	meerkat	meerkat_22Aug2017_guvM2_HiC	GCA_006229205.1	female	Homogametic
<i>Taeniopygia guttata</i>	zebra finch	taeGut3.2.4	GCA_000151805.2	male	Homogametic

Table S4. Information on what sex chromosome(s)/scaffold(s) were detected with the pipeline

Species	Homogametic samples	Heterogametic samples	Snakefile	Study-species reference genome	Synteny-species reference genome
<i>Alauda arvensis</i>	A_arvensis_M	A_arvensis_F	snakefile-le-synten_y	Alauda arvensis	Taeniopygia guttata (Z;3;4A;5*)
<i>Alauda arvensis</i>	A_arvensis_M	A_arvensis_F	snakefile-le-synten_y	Alauda arvensis (consensus)	Taeniopygia guttata (Z;3;4A;5*)
<i>Alauda arvensis</i>	A_arvensis_M (downsampled by 50 %)	A_arvensis_F (downsampled by 50 %)	snakefile-le-synten_y	Alauda arvensis (consensus)	Taeniopygia guttata (Z;3;4A;5*)
<i>Alauda arvensis</i>	A_arvensis_M	A_arvensis_F	snakefile-le-synten_y	Alauda arvensis	Gallus gallus (Z;3;4;5*)
<i>Alauda arvensis</i>	A_arvensis_M	A_arvensis_F	snakefile-le-synten_y	Alauda arvensis	Anolis carolinensis (1;2;GL34319 3.1*)
<i>Alouatta palliata</i>	A_palliata_F1; A_palliata_F2	A_palliata_M1; A_palliata_M2	snakefile-le-synten_y	Alouatta palliata	Homo sapiens (X;3*)
<i>Alouatta palliata</i>	A_palliata_F1; A_palliata_F2	A_palliata_M1; A_palliata_M2	snakefile-le-synten_y	Alouatta palliata	Suricata suricatta (X;5*)
<i>Alouatta palliata</i>	A_palliata_F1	A_palliata_M1	snakefile-le-synten_y	Alouatta palliata	Homo sapiens (X;3*)
<i>Alouatta palliata</i>	A_palliata_F1 (downsampled by 50 %)	A_palliata_M1 (downsampled by 50 %)	snakefile-le-no-synten_y	Alouatta palliata	Homo sapiens (X;3*)
<i>Anolis carolinensis</i>	A_carolinensis_F	A_carolinensis_M	snakefile-no-synten_y	Anolis carolinensis (LGb*)	NA
<i>Calidris pugnax</i>	C_pugnax_R (resident)	C_pugnax_F1; C_pugnax_F2 (faeder)	snakefile-no-synten_y	Calidris pugnax (scaffold28/KQ48216 4.1**)	NA
<i>Calidris pugnax</i>	C_pugnax_R (resident)	C_pugnax_F1; C_pugnax_F2 (faeder)	snakefile-le-no-synten_y	Calidris pugnax	Gallus gallus (11*)
<i>Drosophila miranda</i>	D_miranda_F	D_miranda_M	snakefile-no-synten_y	Drosophila miranda (XL;XR;3*)	NA
<i>Nothobranchius furzeri</i>	N_furzeri_F	N_furzeri_M1; N_furzeri_M2	snakefile-no-	Nothobranchius furzeri (sgr05*)	NA

			synten y		
<i>Ornithorhynchus anatinus</i>	O_anatinus_F	O_anatinus_M	snakefile-no-synten y	Ornithorhynchus anatinus (X1;X2;X3;X4;X5*)	NA
<i>Poecilia reticulata</i>	P_reticulata_F1;P_reticulata_F2	P_reticulata_M1;P_reticulata_M2	snakefile-no-synten y	Poecilia reticulata (LG12***)	NA
<i>Pogona vitticeps</i>	P_vitticeps_M_ZZ;P_vitticeps_F_ZZ	P_vitticeps_F_ZW	snakefile-no-synten y	Pogona vitticeps	Gallus gallus (17;23***)
* Sex chromosomes/scaffolds that were identified with the pipeline					
** Chromosome/scaffolds with inversion polymorphism regions identified with the pipeline					
*** Sex chromosomes/scaffolds that were not identified with the pipeline					

Table S5. Average coverage values for all aligned samples (Table S1), calculated as: Total aligned base pairs (all reads)/Length of study-species reference genome

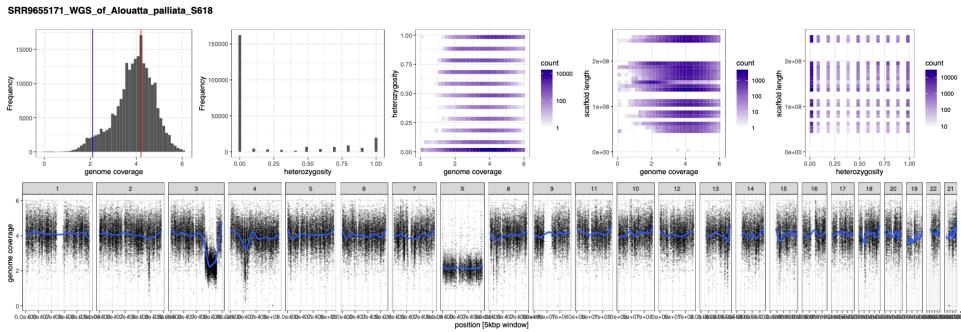
Alias (WGS data)	Study-species reference genome	Average coverage		
		all mismatches	<=2 mismatches	0 mismatches
A_arvensis_F	<i>Alauda arvensis</i>	27.31	15.09	4.54
A_arvensis_M	<i>Alauda arvensis</i>	27.28	23.07	16.73
A_arvensis_F	<i>Alauda arvensis</i> (consensus)	27.30	16.89	6.42
A_arvensis_M	<i>Alauda arvensis</i> (consensus)	27.29	22.10	13.46
A_arvensis_F (downsampled by 50 %)	<i>Alauda arvensis</i> (consensus)	15.20	9.41	3.57
A_arvensis_M (downsampled by 50 %)	<i>Alauda arvensis</i> (consensus)	15.00	12.16	7.40
A_palliata_F1	<i>Alouatta palliata</i>	4.45	4.33	3.74
A_palliata_F2	<i>Alouatta palliata</i>	6.01	5.85	5.06
A_palliata_M1	<i>Alouatta palliata</i>	2.65	2.56	2.19
A_palliata_M2	<i>Alouatta palliata</i>	4.47	4.32	3.72
A_palliata_F1 (downsampled by 50%)	<i>Alouatta palliata</i>	2.61	2.54	2.19
A_palliata_M1 (downsampled by 50%)	<i>Alouatta palliata</i>	1.56	1.50	1.29
A_carolinensis_F	<i>Anolis carolinensis</i>	12.54	11.69	8.74
A_carolinensis_M	<i>Anolis carolinensis</i>	10.32	9.58	6.78
C_pugnax_F1	<i>Calidris pugnax</i>	6.16	5.66	4.28
C_pugnax_F2	<i>Calidris pugnax</i>	5.66	5.21	3.92
C_pugnax_R	<i>Calidris pugnax</i>	6.50	5.98	4.55
D_miranda_F	<i>Drosophila miranda</i>	4.02	3.38	2.88
D_miranda_M	<i>Drosophila miranda</i>	9.87	8.10	5.99
N_furzeri_M1	<i>Nothobranchius furzeri</i>	17.27	15.53	12.44
N_furzeri_M2	<i>Nothobranchius furzeri</i>	25.15	23.01	18.66
N_furzeri_F	<i>Nothobranchius furzeri</i>	27.84	25.44	21.72
O_anatinus_M	<i>Ornithorhynchus anatinus</i>	13.93	13.25	9.85
O_anatinus_F	<i>Ornithorhynchus anatinus</i>	14.33	13.55	10.04
P_reticulata_F1	<i>Poecilia reticulata</i>	40.01	34.14	19.73
P_reticulata_F2	<i>Poecilia reticulata</i>	36.66	31.49	18.39
P_reticulata_M1	<i>Poecilia reticulata</i>	34.71	29.65	17.15

P_reticulata_M2	<i>Poecilia reticulata</i>	42.25	35.88	20.66
P_vitticeps_F_ZW	<i>Pogona vitticeps</i>	30.29	16.31	7.42
P_vitticeps_M_ZZ	<i>Pogona vitticeps</i>	28.29	21.12	13.89
P_vitticeps_F_ZZ	<i>Pogona vitticeps</i>	46.34	28.22	12.71

Supplementary Figures

Here, we present figures supporting results that are specifically mentioned in the Main text. Information on how each result plot was generated is found in the Figure legend. Unless otherwise stated, all samples from each species were used to analyse the data (Table S1). The reference genomes used, either as a study-species reference genome or as a synteny-species reference genome, correspond to Table S3.

a



b

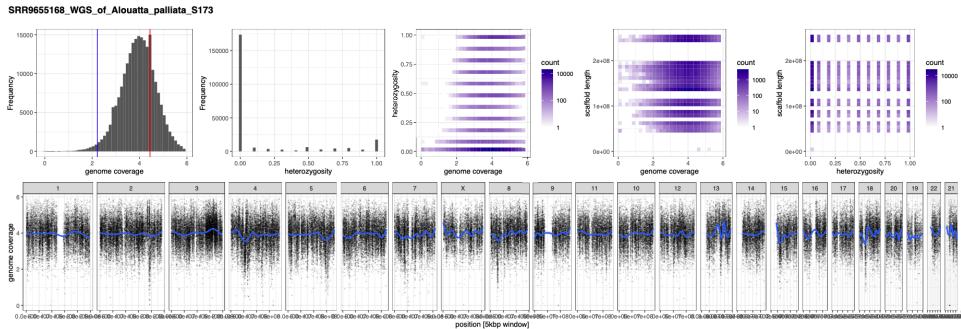


Figure S1. Example of a “Confirm-sexing” plot (see Figure 2), from one (a) male (sample SRR9655171) and one (b) female (sample SRR9655168) mantled howler monkey (*A. palliata*).

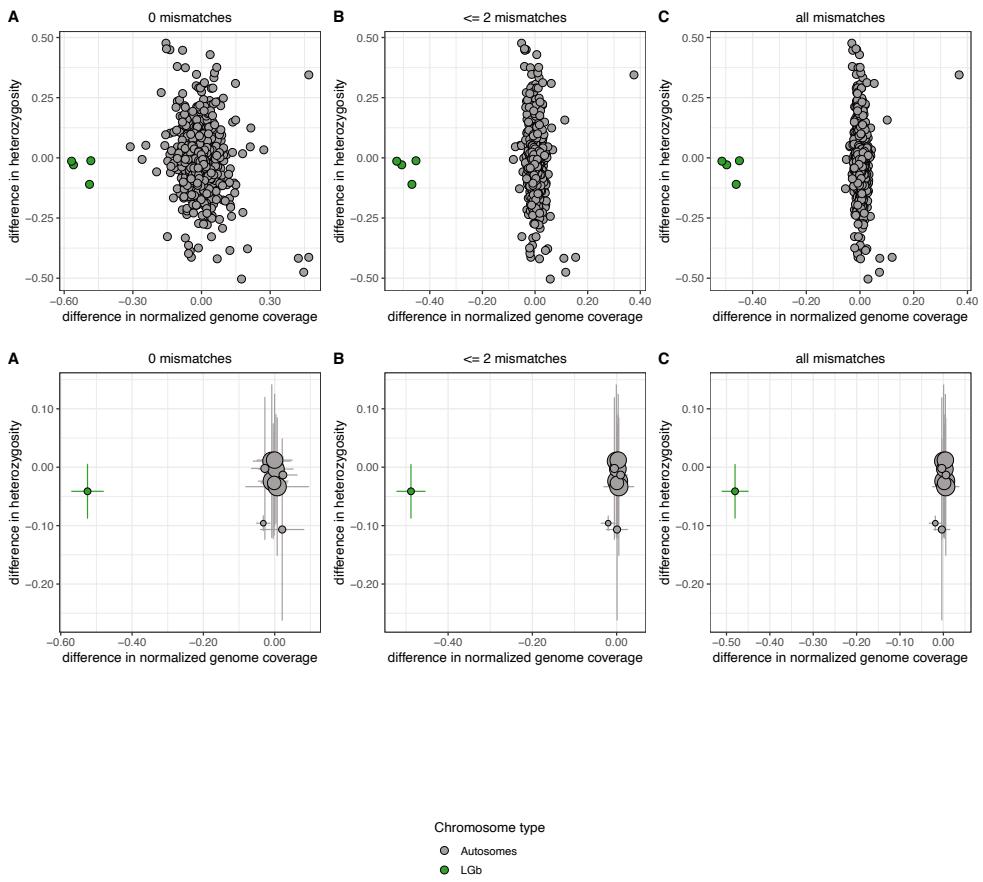


Figure S2. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male and 1 female green anole (*A. carolinensis*), analysed using snakefile-no-synteny (i.e. without the use of a synteny-species reference genome). The previously identified sex chromosome in this species (LGb) is a clear outlier.

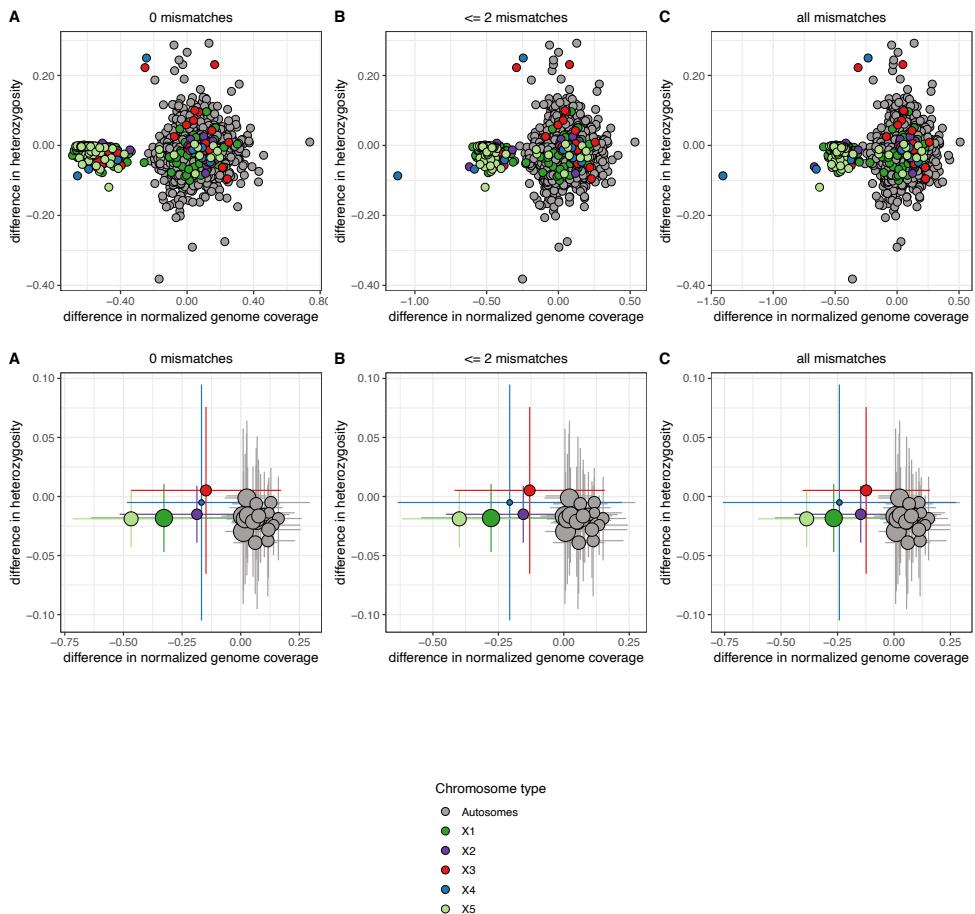
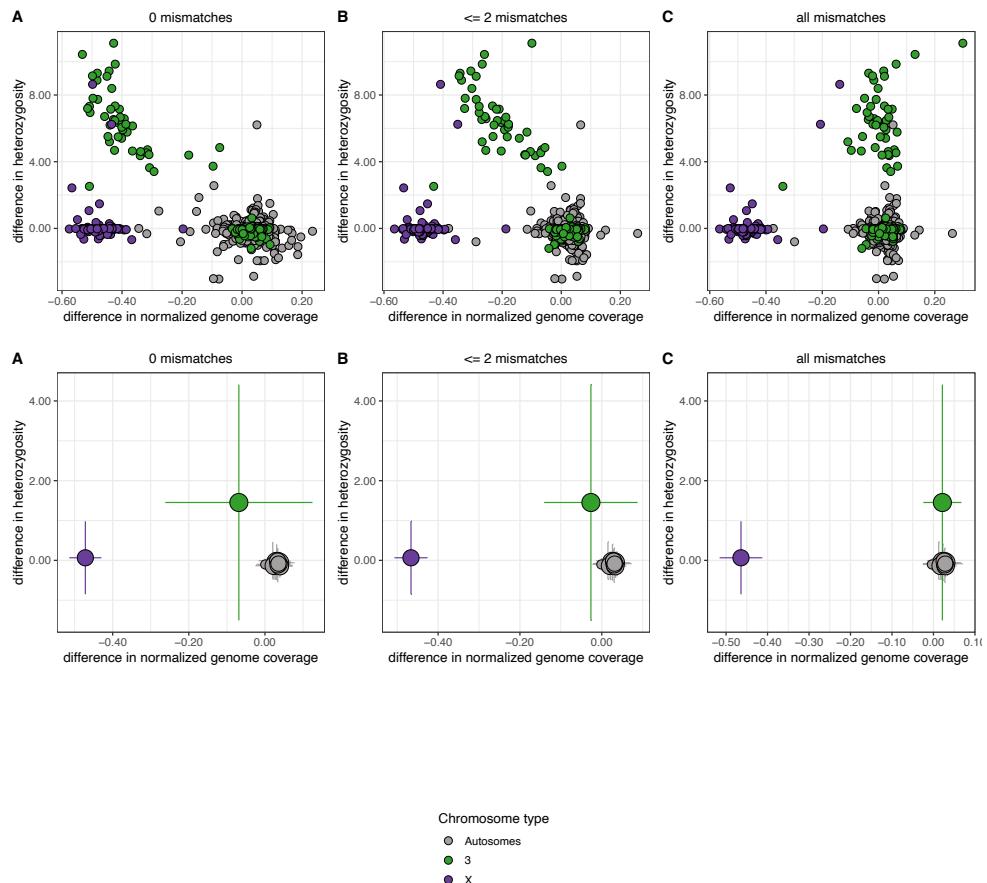
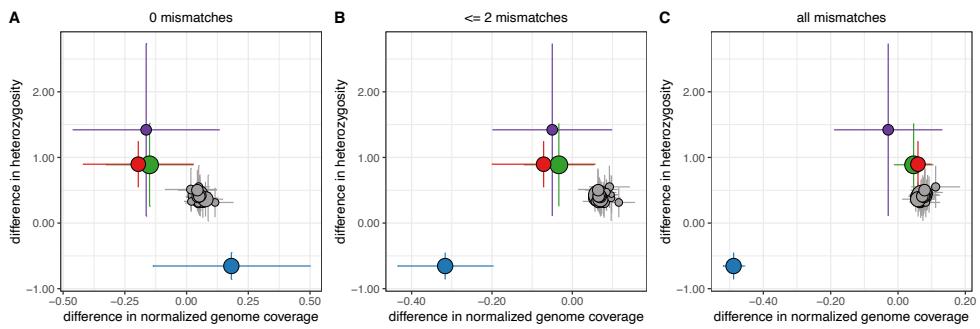
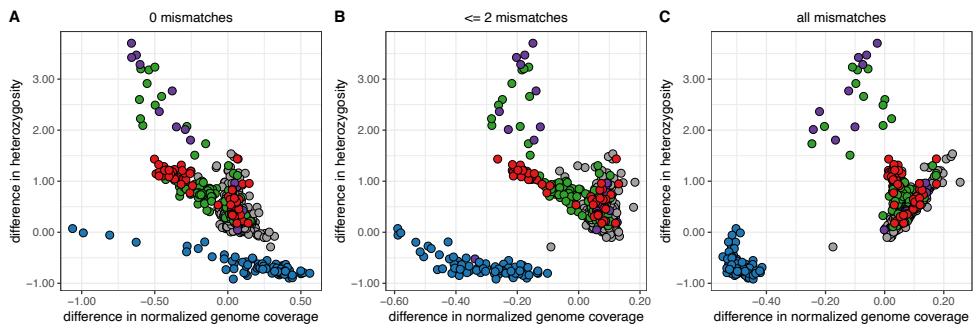


Figure S3. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male and 1 female platypus (*O. anatinus*), analysed using snakefile-no-synteny (without the use of a synteny-species reference genome). The previously identified sex chromosomes in this species (X1, X2, X3, X4 and X5) are clear outliers.





Chromosome type

- Autosomes
- 3
- 4A
- 5
- Z

Figure S5. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male and 1 female Eurasian skylark (*A. arvensis*; consensus-reference genome), analysed using snakefile-synteny and zebra finch (*T. guttata*) as a synteny-species reference genome. The previously identified sex chromosomes in this species (Z, 3, 4A and 5) are clear outliers.

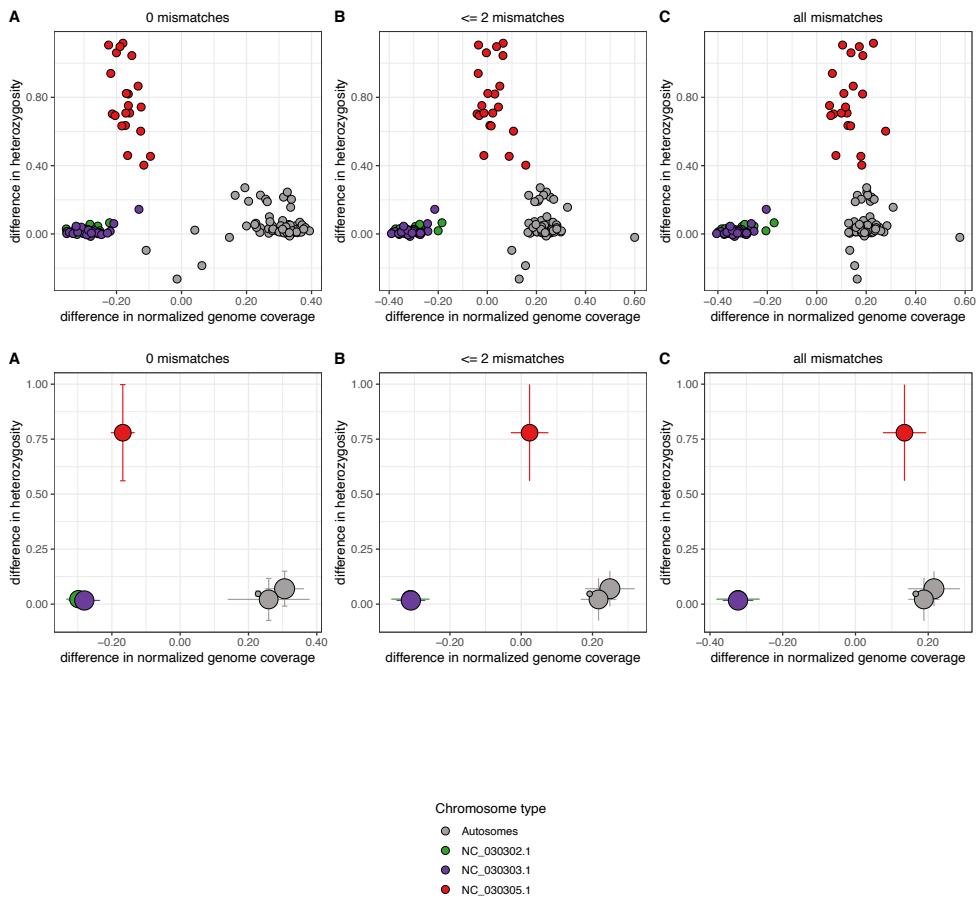


Figure S6. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male and 1 female fruit flies (*D. miranda*), analysed using snakefile-synteny (without a synteny-species reference genome). The previously identified sex chromosomes in this species (XL; NC_030302.1, XR; NC_030303.1 and 3; NC_030305.1) are clear outliers.

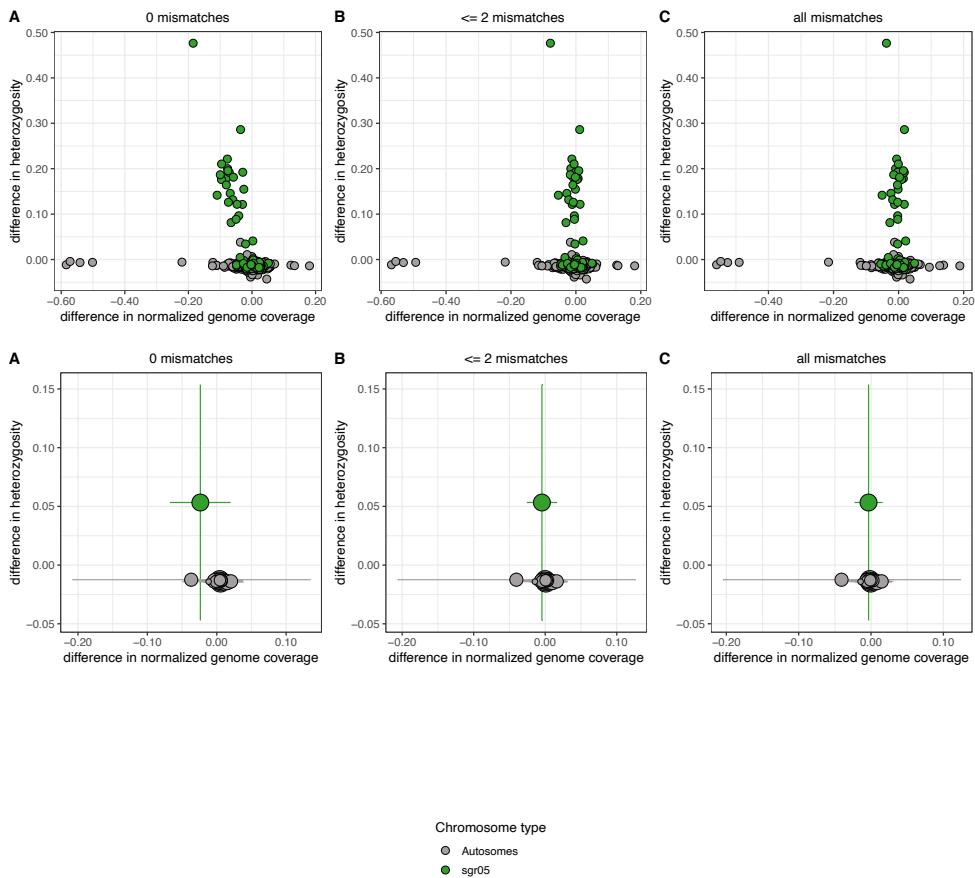


Figure S7. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 2 male and 1 female turquoise killifish (*N. furzeri*), analysed using snakefile-no-synteny (without a synteny-species reference genome). The previously identified sex chromosomes in this species (sgr05) is a clear outlier.

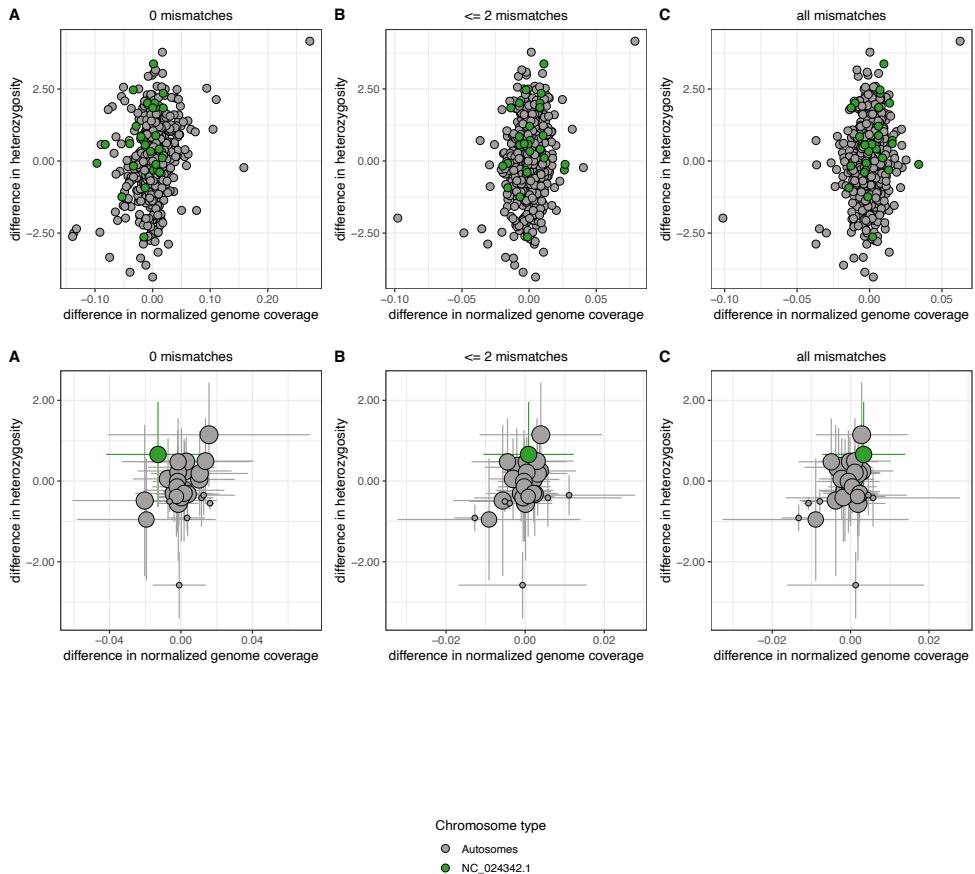


Figure S8. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 2 male and 2 female guppies (*P. reticulata*), analysed using snakefile-no-synteny (without a synteny-species reference genome). The previously identified sex chromosomes in this species (Chromosome 12/NC_024342.1) is not a clear outlier.

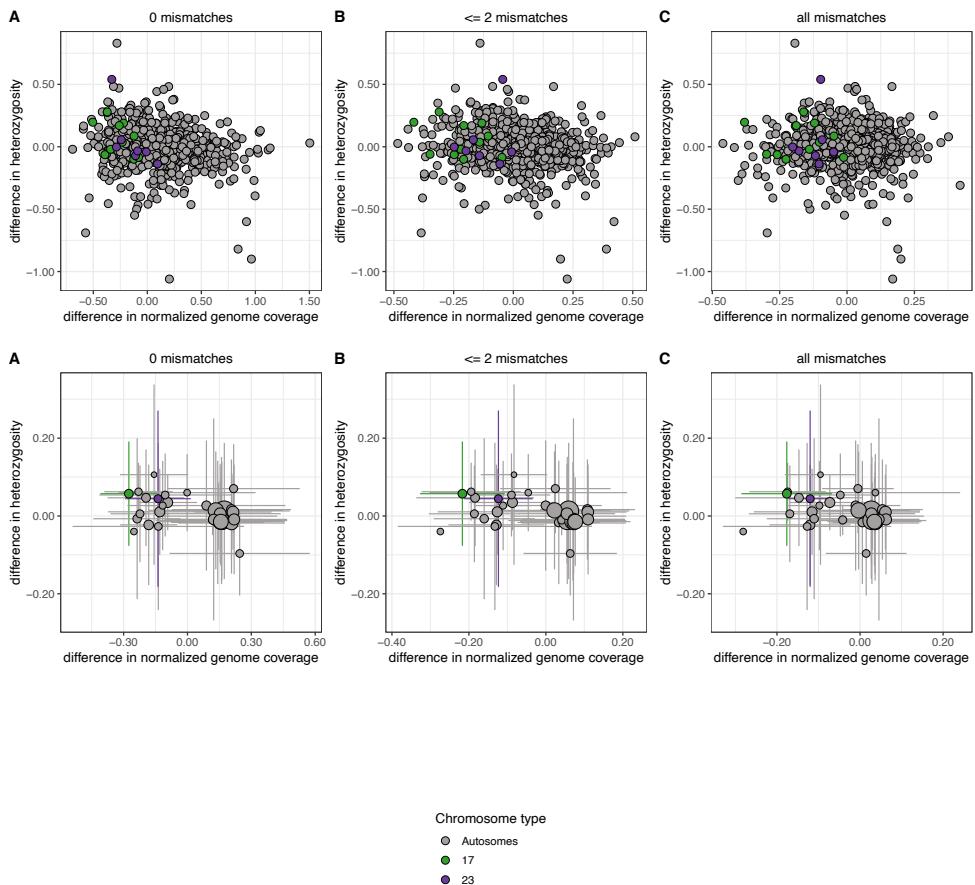


Figure S9. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male homozygotic (ZZ), 1 female homozygotic (ZZ) and 1 female heterozygotic (ZW) central bearded dragon (*P. vitticeps*), analysed using snakefile-no-synteny and with chicken as a synteny-species reference genome. The previously identified sex chromosomes in this species (17 and 23) were not outliers compared to autosomes.

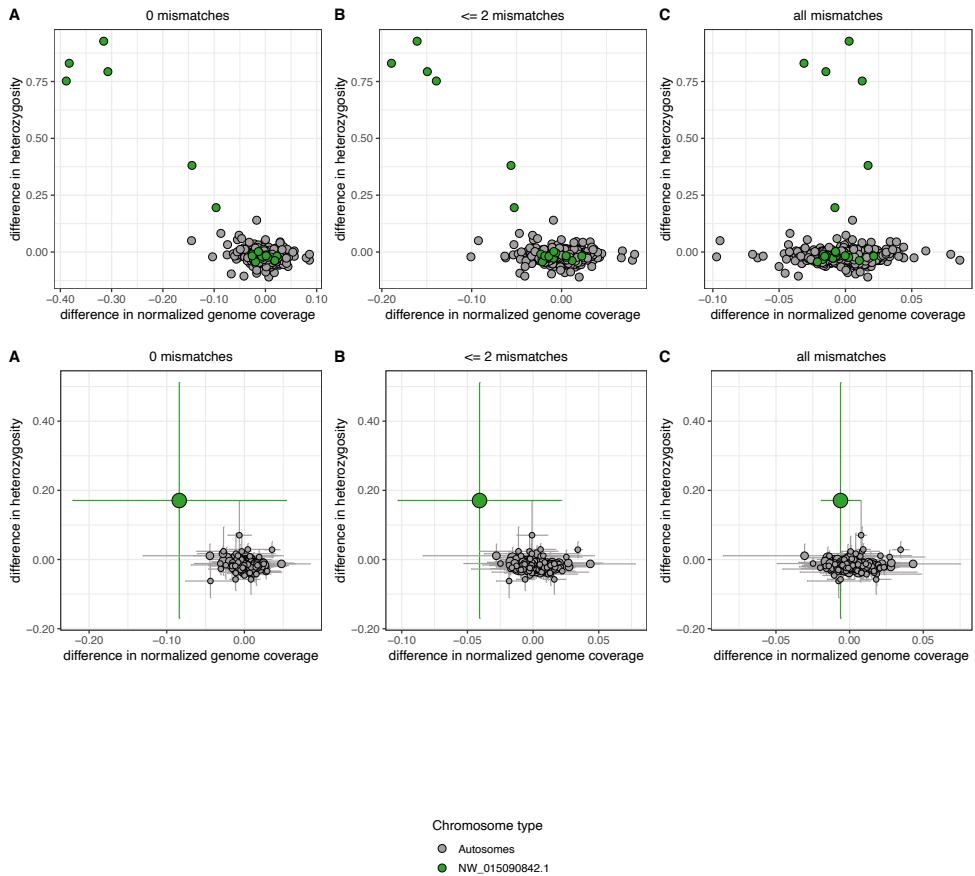


Figure S10. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 3 male ruffs (*C. pugnax*); 2 faeder male (heterozygotic inversion) and 1 resident male (homozygotic inversion). The data was analysed using snakefile-no-synteny (without a synteny-species reference genome). The previously identified scaffold containing the inversion polymorphism (scaffold28/NW_015090842.1) was a clear outlier.

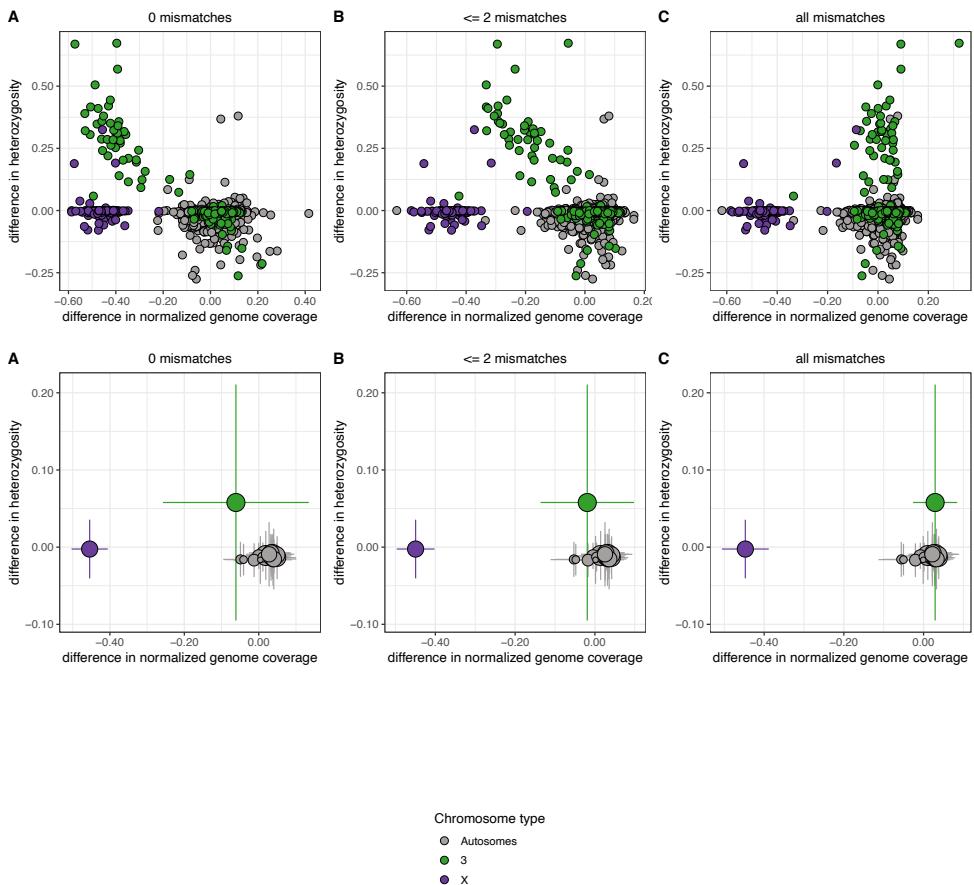


Figure S11. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male (*A_palliata_M1*) and 1 female (*A_palliata_F1*) mantled howler monkeys (*A. palliata*), analysed using snakefile-synteny and human (*H. sapiens*) as a synteny-species reference genome. The previously identified sex chromosomes in this species (X and 3) are clear outliers.

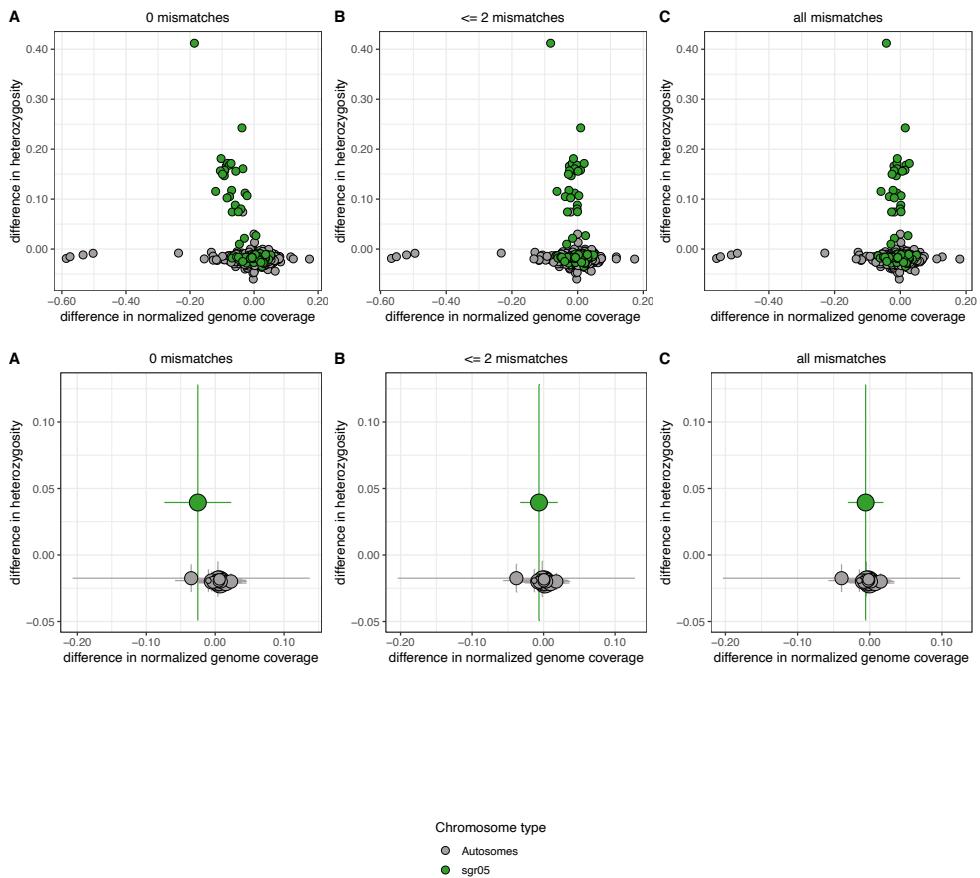


Figure S12. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male (*N. furzeri*_M1) and 1 female (*N. furzeri*_F) turquoise killifish (*N. furzeri*), analysed using snakefile-no-synteny (without a synteny-species reference genome). The previously identified sex chromosomes in this species (sgr05) is a clear outlier.

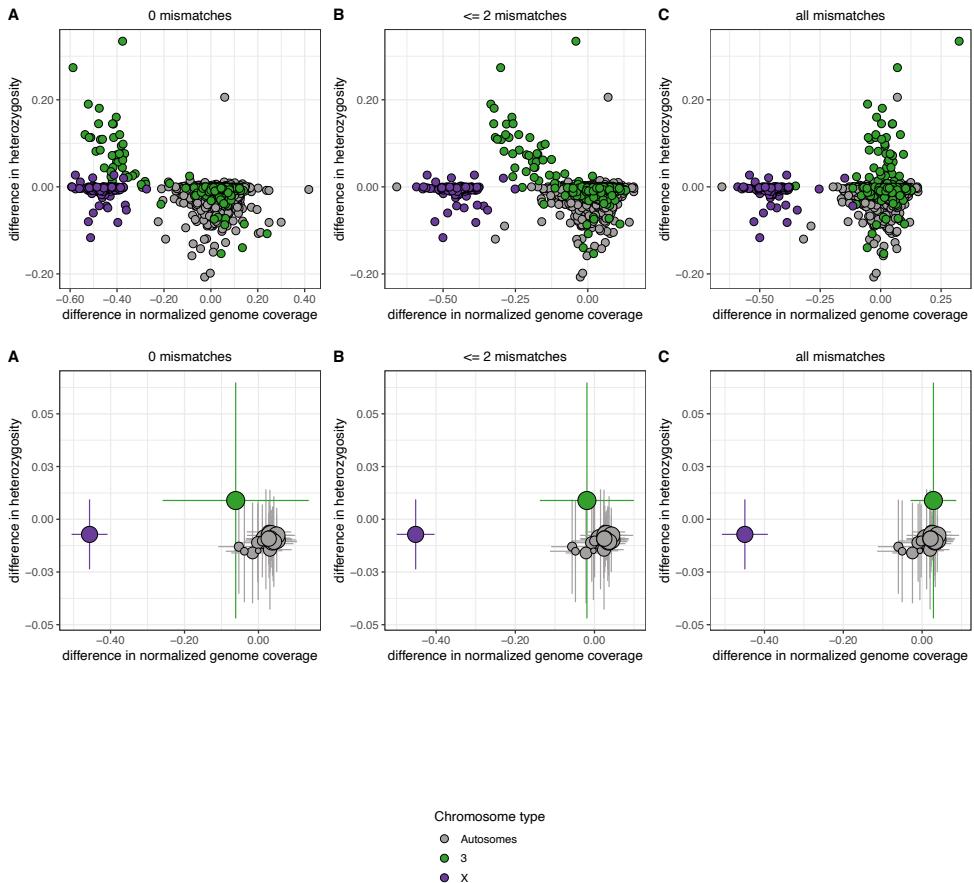


Figure S13. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male (*A. palliata*_M1) and 1 female *A. palliata*_F1 mantled howler monkeys (*A. palliata*), subsampled to contain 50 % of the original number of reads. The data was analysed using snakefile-synteny and human (*H. sapiens*) as a synteny-species reference genome. The previously identified sex chromosomes in this species (X and 3) are clear outliers.

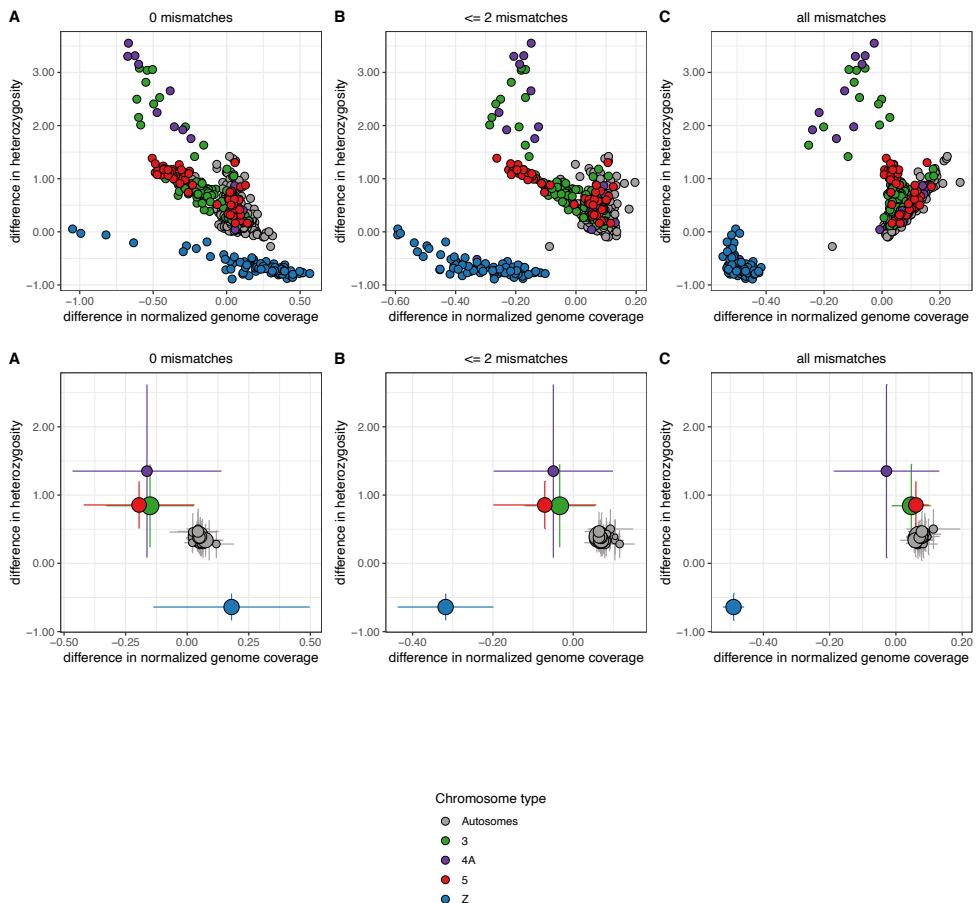


Figure S14. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male and 1 female Eurasian skylark (*A. arvensis*), subsampled to contain 50 % of the original number of reads. The data was analysed using snakefile-synteny and zebra finch (*T. guttata*) as a synteny-species reference genome. The previously identified sex chromosomes in this species (Z, 3, 4A and 5) are clear outliers.

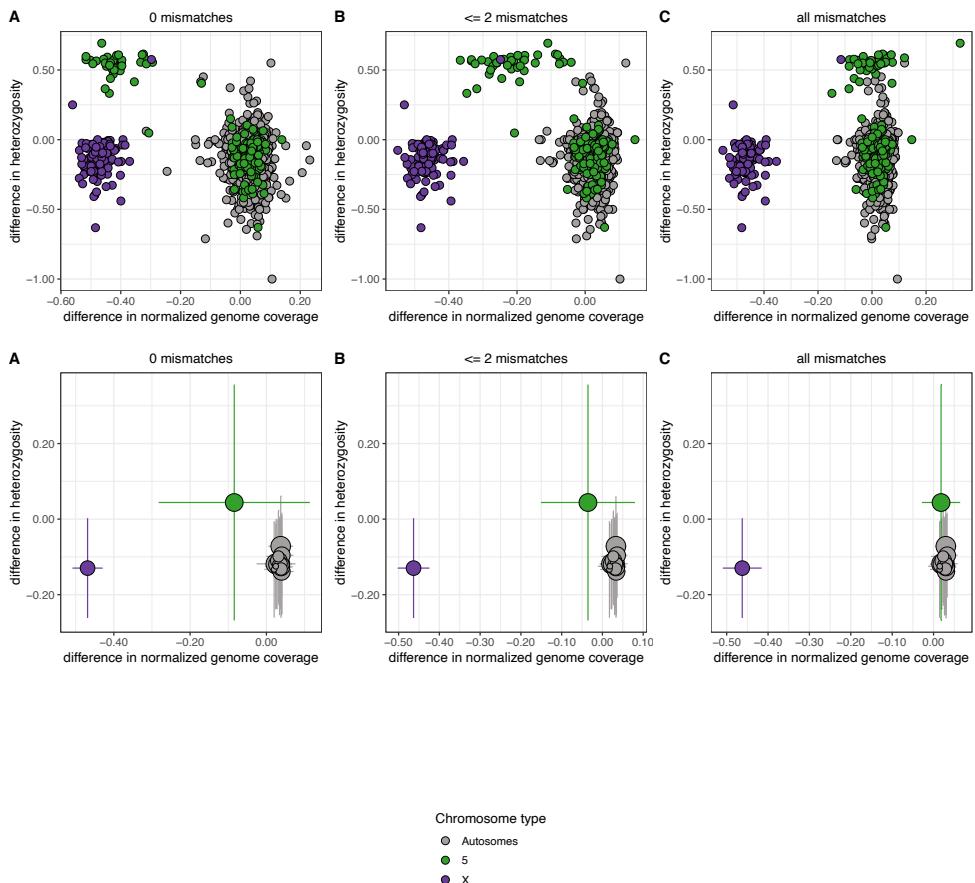


Figure S15. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 2 male and 2 female mantled howler monkeys (*A. palliata*), analysed using snakefile-synteny and meerkat (*S. suricatta*) as a synteny-species reference genome. The previously identified sex chromosomes in this species (X and 5) are clear outliers.

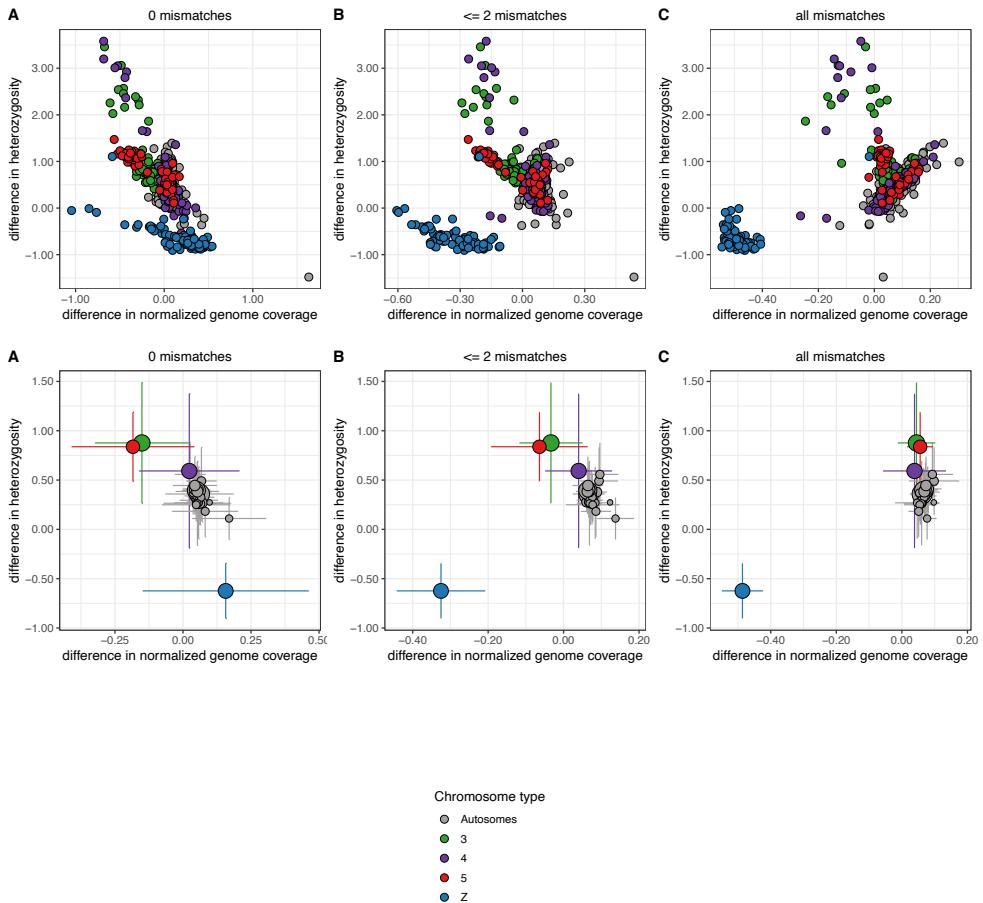


Figure S16. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 1 male and 1 female Eurasian skylark (*A. arvensis*; consensus-reference genome), analysed using snakefile-synteny and chicken (*G. gallus*) as a synteny-species reference genome. The previously identified sex chromosomes in this species (Z, 3 and 5) are clear outliers. Chromosome 4 is not a clear outlier, perhaps as a result of the sex-linked part of this chromosome only constituting 10 Mb of this chromosome (91 Mb in total). Figure S17 show the same data as in this plot, but across chromosome positions in the chicken genome. In this plot (see below), the beginning of chromosome 4A (which is homologous to chromosome 4A in the zebra finch) show a sex-linked pattern.

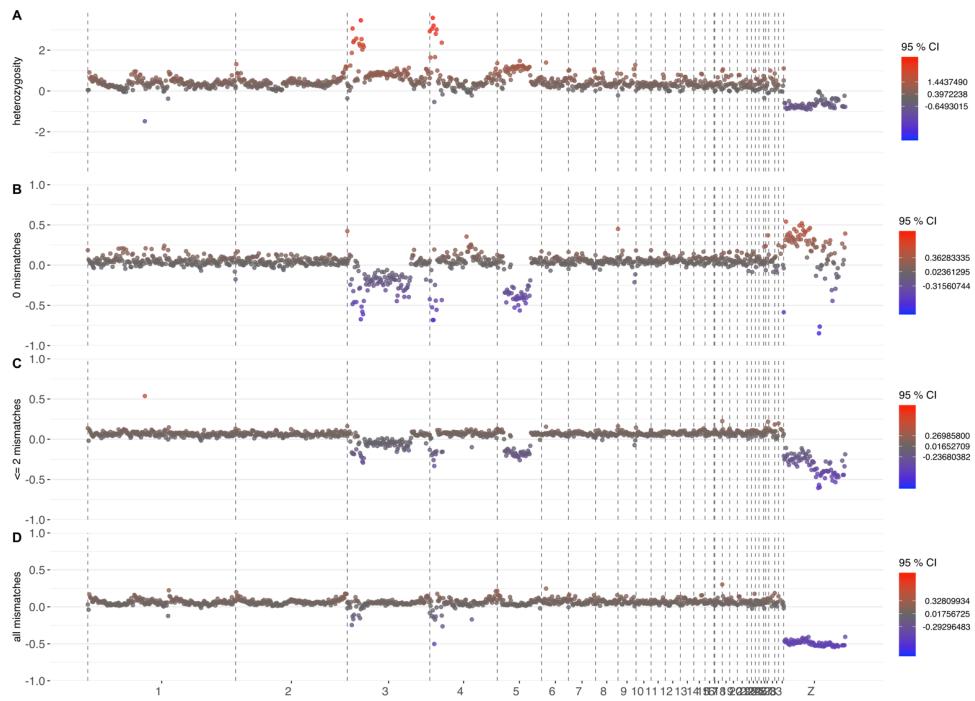


Figure S17. Genome-wide heterozygosity and genome coverage values from 1 male and 1 female Eurasian skylark (*A. arvensis*), analysed using snakefile-synteny and chicken (*G. gallus*) as a synteny-species reference genome. Here plotted along the chromosome positions in the chicken genome. The previously identified sex chromosomes in this species (Z, 3, 4 and 5) are clear outliers.

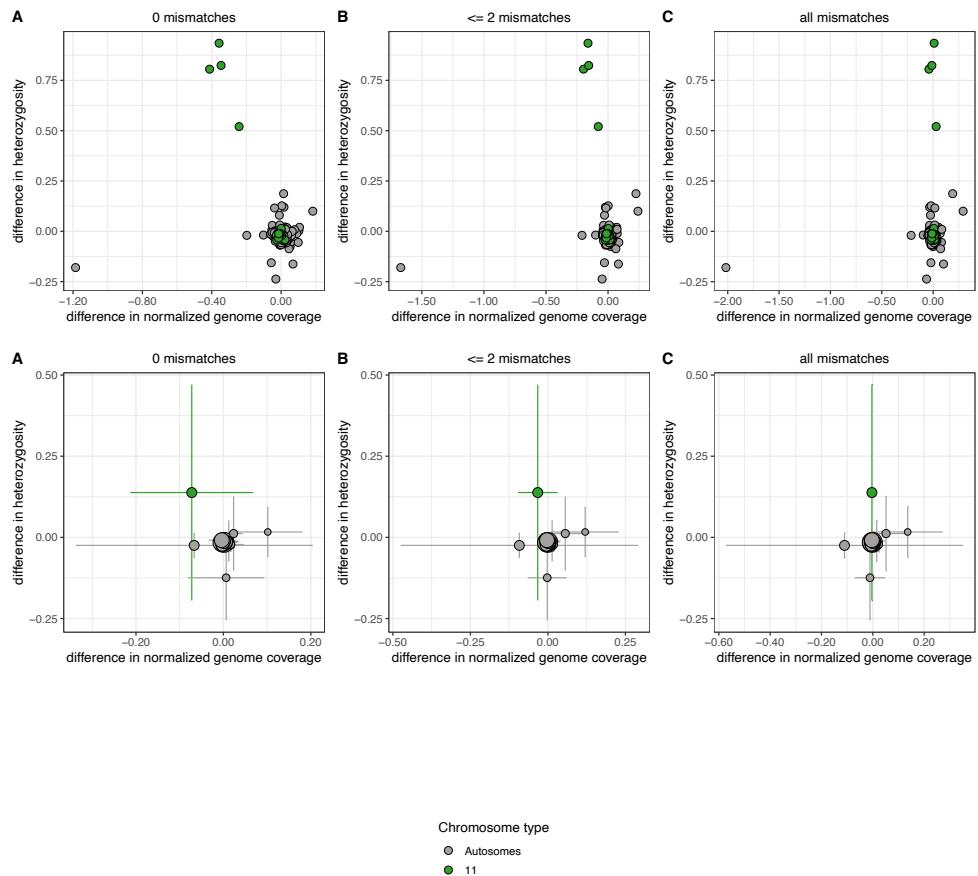


Figure S18. Genome coverage and heterozygosity values (mean \pm standard deviation across 1 Mb windows) per chromosome/scaffold. Data from 3 male ruffs (*C. pugnax*); 2 faeder male (heterozygotic inversion) and 1 resident male (homozygotic inversion). The data was analysed using snakefile-synteny with chicken (*G. gallus*) as a synteny-species reference genome. Chromosome 11, which contain the inversion polymorphism, is an outlier.

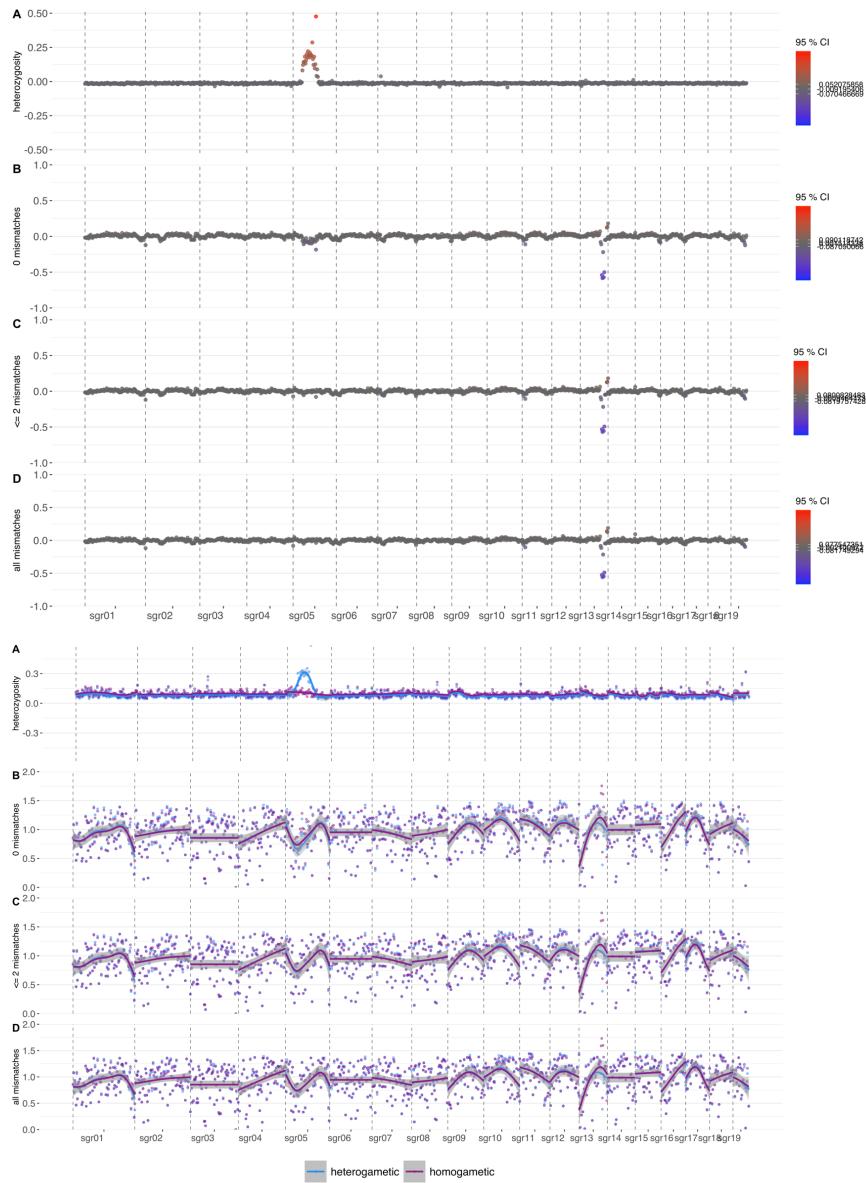


Figure S19. Genome-wide heterozygosity and genome coverage values from 2 male and 1 female turquoise killifish (*N. furzeri*), analysed using snakefile-no-synteny. In the top figure, sgr13 show a reduction in genome coverage in the heterogametic sex compared to the homogametic one. In the bottom figure, however, there is no reduction in genome coverage in the heterogametic sex for sgr13 compared to the genome average. The bottom figure show that the homogametic sex has slightly elevated genome coverage values in this region compared to the rest of the genome, which likely produced the pattern we see in the top figure. When interpreting results provided by this pipeline, it is important to look at all the plots in order to avoid misinterpretations (see Discussion).