PNAS

Supporting Information for

Incomplete transcriptional dosage compensation of vertebrate sex chromosomes is balanced by post-transcriptional compensation.

Nicholas C Lister[1], Ashley M Milton[1], Hardip R Patel[2,3], Shafagh A Waters[4], Benjamin J Hanrahan[1], Kim L McIntyre[1], Alexandra M Livernois[5], Lee Kian Wee[1], Alessa R Ringel[6,7,8], Stefan Mundlos[6,7,9], Michael I Robson[6,7,10], Linda Shearwin-Whyatt[11], Frank Grützner[11], Jennifer A Marshall Graves[12,13]*, Aurora Ruiz-Herrera[14,15], Paul D Waters[1]

*Corresponding author: Jennifer A Marshall Graves
Email: J.Graves@latrobe.edu.au

This PDF file includes:

Materials and Methods

Materials Availability

This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Ethical Guidelines

All tissue collection was approved by the Australian National University Animal Experimentation Ethics Committee (approval number R.CG.14.08) and the Garvan/St Vincents Animal Ethics Committee (#13/35).

Tissue

For both male and female chicken (*Gallus gallus*), three biological replicates were used for heart tissues and two biological replicates were used for liver. For both male and female mouse (*Mus musculus*), two biological replicates were used for heart and liver. For both male and female platypus (*Ornithorhynchus anatinus*), two biological replicates were used for heart and liver.
Primary fibroblast cell cultures were established for male and female individuals of both opossum (*Monodelphis domestica*) and platypus. Opossum fibroblast cell lines were cultured at 35 °C and platypus fibroblasts at 37 °C in 5% CO_2 with DMEM containing 10% *(v/v)* FBS and 10 ml/L PSG.

METHOD DETAILS

RNA extraction

One million fibroblast cells and 25 mg of tissue from either heart or liver was used for RNA isolation using TRIzol reagent (ThermoFisher Scientific).

Protein extraction

Protein was extracted from heart and liver tissues of chicken, mouse and platypus, as well as fibroblasts for platypus and opossum, using RIPA lysis buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 2mM EDTA, 50 mM Tris HCl, pH 8.0). The tissues were homogenized to encourage cell lysis. The lysates were placed on ice for 30 minutes before being centrifuged at 10,000 x *g* for 20 minutes. The residual insoluble fraction of the extraction was subsequently extracted and analyzed on a 12% SDS-PAGE gel. Qubit quantification was performed to determine protein concentration. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE(1) partner repository with the dataset identifier PXD040182.

LC-MS/MS

Chicken, mouse and platypus heart and liver samples were sent to Bioanalytical Mass Spectrometry Facility (BMSF) at the Mark Wainwright Analytical Centre (MWAC) at UNSW. The opossum and platypus fibroblast samples were analysed at the Australian Proteome Analysis Facility (APAF) at Macquarie University.
Samples (dried) were solubilised in 1 ml of 0.25 M TEAB 0.05% SDS and buffer exchanged, using a Viva2 spin column (5 kDa), into 250 µL of 0.25 M TEAB 0.05% SDS. Samples were quantified using Direct Detect (Millipore). 100 µg of each sample (in duplicate) was reduced with TCEP,

alkylated with MMTS and digested with trypsin according to SOP MS-001. The digested samples were iTRAQ labelled, cleaned and fractionated by SCX HPLC. The buffer A was 5 mM phosphate 25% acetonitrile, pH 2.7 and buffer B was 5 mM phosphate 350 mM KCL 25% acetonitrile, pH 2.7. The dried iTRAQ labelled sample was resuspended in buffer A. After sample loading and washing with buffer A, buffer B concentration increased from 10% to 45% in 70 minutes and then increased quickly to 100% and stayed at 100% for 10 minutes at a flow rate of 300 µL/min. The eluent of SCX was collected every 2 minutes at the beginning of the gradient and at 4 minute intervals later.

The labelled sample was resuspended in 100 µl of loading/desalting solution (0.1% formic acid and 2% acetonitrile 97.9% water). The sample (40 µl) was injected onto a peptide trap (Michrome peptide Captrap) for pre-concentration and desalted with 0.1% formic acid, 2% ACN, at 5 µl/min for 10 minutes. The peptide trap was then switched into line with the analytical column. Peptides were eluted from the column using a linear solvent gradient, with steps, from mobile phase A: mobile phase B (98:2) to mobile phase A: mobile phase B (65:35) where mobile phase A was 0.1% formic acid and mobile phase B was 90% ACN/0.1% formic acid at 600 nl/min over a 100 minute period. After peptide elution, the column was cleaned with 95% buffer B for 15 minutes and then equilibrated with buffer A for 25 minutes before the next sample injection. The reverse phase nanoLC eluent was subject to positive ion nanoflow electrospray analysis in an information dependant acquisition mode (IDA). Commercial in confidence In IDA mode a TOFMS survey scan was acquired (m/z 400 - 1500, 0.25 second), with the ten most intense multiply charged ions (counts >150) in the survey scan sequentially subjected to MS/MS analysis. MS/MS spectra were accumulated for 200 milliseconds in the mass range m/z 100 – 1500 with the total cycle time 2.3 seconds.

ChIP-Seq

Platypus fibroblast cells were harvested and fixed with 1% formaldehyde in 10% FCS/PBS for 10 minutes at RT. Crosslinking was quenched with glycine at a final concentration of 125 mM. Fixed cells were washed twice with cold PBS and lysed using fresh lysis buffer (10 mM Tris, pH 7.5, 10 mM NaCl, 5 mM MgCl2, 0.1 mM EGTA with protease inhibitor). Nuclei were centrifuged for 5 min /480g /4 °C, washed with PBS and either snap frozen in liquid N_2 or processed immediately. Chromatin for ChIP-seq was sheared using a Bioruptor until DNA fragment size of 200–500 base pairs was reached.

Samples were processed with the iDeal ChIP-seq kit for histones/transcription factor according to the manufacturer's instructions. For each histone ChIP 5 µg chromatin was used with antibodies against H3K4me1 (1 µg) H3K4me3 (1 µg), H3K27ac (1 µg), and H4K20me1 (5 µg). For H3K4me1, K3K4me3 and H3K27ac, libraries were prepared for sequencing using the KAPA HyperPrep kit and their quality confirmed by Bioanalyzer analysis. Samples were run on NovaSeq 6000 platform.

For H4K20me1, 20 µl of magnetic G beads (Millipore) were added to each ChIP sample and incubated at 4 °C rotating for 2 hours. Beads were allowed to settle in a magnetic rack before removing supernatant. Beads were washed consecutively with 1 ml of low salt buffer, 1 ml of high salt buffer, 1 ml of LiCI buffer, 1 ml of TE buffer, for 5 minutes each at 4 °C. Between each wash, beads were allowed to settle in a magnetic rack before removing supernatant. After the final wash, contents were transferred fresh 1.5 ml tubes. Beads were allowed to settle in a magnetic rack, and

supernatant was removed. Tubes were removed from the magnetic rack and 100 µl of elution buffer was added, followed by agitation at room temperature for 30 minutes. Beads were allowed to settle in a magnetic rack and supernatant (chromatin-antibody complexes) was transferred to a fresh tube. NaCl was added to final concentration of 0.6 M. Total input was thawed and diluted with 4 volumes elution buffer, and NaCl was added to a final concentration of 0.6 M. Samples were incubated at 67 °C overnight to reverse cross-link. Samples were sent to the Beijing Genomics Institute (BGI) for library construction and sequencing on the HiSeq 2000 platform.

Hi-C

Hi-C libraries were prepared as described in a previously published *in situ* protocol(2, 3). Briefly, male platypus fibroblast cells were harvested and fixed with 2% formaldehyde in 10% FCS/PBS for 10 minutes at RT. Crosslinking was quenched by adding glycine to a final concentration of 125 mM. To isolate nuclei the fixed cells were washed twice with cold PBS and lysed using fresh lysis buffer (10 mM Tris, pH 7.5, 10 mM NaCl, 5 mM MgCl2, 0.1 mM EGTA with protease inhibitor). Nuclei were centrifuged 5 min /500g / 4 °C, washed several times with PBS and subsequently digested overnight with the restriction enzyme DpnII. Digested DNA ends were marked with biotin-14-dATP and ligated overnight using T4 DNA ligase. Formaldehyde crosslinking was reversed by incubation in 5 M NaCl for 2 hours at 68 °C, followed by ethanol precipitation. Chromatin was sheared into 300–600 bp fragments using a S-Series 220 Covaris for library preparation, afterwards biotin-filled DNA fragments were pulled down using Dynabeads MyOne Streptavidin T1 beads. DNA ends were repaired using T4 DNA polymerase and the Klenow fragment of DNA polymerase I and phosphorylated with T4 Polynucleotide Kinase NK. Afterwards adaptors were ligated to DNA fragments using the NEBNext Multiplex Oligos for Illumina kit (Illumina). Indexes were added via PCR amplification (3–8 cycles) using the NEBNext Ultra II Q5 Master Mix. PCR purification and fragment size selection were done using Agencourt AMPure XP beads. Library sequencing was done on the Illumina NovaSeq 6000 platform.

QUANTIFICATION AND STATISTICAL ANALYSIS

RNA Sequencing

Chicken and mouse samples were analysed by 1 x 75 bp single-read sequencing on the Illumina NextSeq 500 platform. Library construction and sequencing was performed by Ramaciotti Centre for Genomics (UNSW Sydney, Australia), using the TruSeq Stranded mRNA Library Prep Kit. Opossum and platypus fibroblast samples were sequenced on Illumina HiSeq 2000 by Beijing Genomics Institute (BGI). 2 x 100 bp reads were generated for platypus fibroblasts, while 1 x 75 bp reads were generated for opossum fibroblasts. The resulting raw sequence data is available on the NCBI short read archive as BioProject PRJNA929280. Platypus heart and liver RNA-seq data were obtained from a paper from the Kaessmann group(4). The raw reads from RNA-seq, Hi-C and ChIP-seq were analysed using FastQC(5) to assess read quality.

Sequence mapping

The raw reads from RNA-seq, Hi-C and ChIP-seq were analysed using FastQC(5) to assess read quality. For feature counts for MvF plots, reference genomes were masked for the Y chromosome sequences in platypus and mouse and for the W chromosome in chicken to reduce false positive mappings in the sex lacking these chromosomes for RNA-Seq data. This was not required for the

opossum assembly which was from a female sample. The reference genome versions were GRCg6a for chicken, mOrnAna1.p.v5.1 for platypus, MonDom5 for opossum and GRCm39 for mouse. Subread(6) was used to map the reads to their reference genomes. For RNA-Seq reads, parameters were chosen such that (a) all possible subreads were chosen for mapping (-n 300) and (b) for multi-mapping reads, one random location was chosen for read assignments (--multiMapping -B 1). For Hi-C reads, Subread was run in single-end mode to generate separate forward and reverse .sam files. All parameters were default. Subread was used on ChIP-seq reads in the same fashion, with the exception of specifying a Phred score of phred+64 (-P 6).

Counting mapped reads

FeatureCounts(7) was used to calculate the number of reads assigned to each feature based on gtf files from NCBI. Parameters were chosen such that (a) reads were assigned to all overlapping features if features were overlapping (-O), (b) multi-mapping reads were counted (-M), and (c) strand-specificity (-s 0,1, or 2). For paired-end data, fragments were counted instead of individual reads of a pair (-p).

ProteinPilot workflow

The MS/MS data were submitted to ProteinPilot V5.0.2 (AB Sciex)(8) for processing, using databases obtained from SwissProt. Bias correction was selected. The detected protein threshold (unused ProtScore) was set as larger than 1.3 (better than 95% confidence). The Protein Pilot group file and the protein summaries were exported and male:female ratios calculated.

In platypus, two heart samples each from male (113 and 114) and female (115 and 116), and two liver samples each from male (117 and 118) and female (119 and 121) were analyzed. In chicken, three heart samples each from male (113, 114 and 115) and female (116, 117 and 118), and two liver samples each from male (113 and 114) and female (115 and 116) were analyzed. In mouse, two heart and liver samples each from male (117 and 118) and female (119 and 121) were analyzed. One sample each from platypus fibroblasts were labelled twice for male (115 and 117) and female (114 and 116). One sample each from opossum fibroblasts were labelled twice for male (114 and 116) and female (115 and 117). Male:female ratios for each combination of male and female samples were extracted directly from the protein summary files (e.g. column heading 114:116) to calculate geomeans for each gene.

Hi-C matrix generation and analysis

HiCExplorer(9-11) version 3.4.3 was used for Hi-C matrix generation and analysis. HicBuildMatrix was used to generate 500 kb .h5 genome-wide contact matrices from male forward and reverse read .sam files (--binSize 500000; --restrictionSequence GATC; --danglingSequence GATC). HicSumMatrices was then used to create a merged .h5 genome-wide contact matrix containing the four male replicate samples. Afterwards, hicCorrectMatrix was run to give a corrected genome-wide contact matrix (containing only one resolution/bin size) (--filterThreshold -1.5 3). From this point, multiple hicexplorer tools could be used. Given a .bw file for an active histone mark (H4K20me1), hicPCA was used to generate a Pearson's correlation matrix and .bw eigenvector/principle component files for the corrected matrix (--numberOfEigenvectors 1; --format bigwig; --histonMarkType active). Hicexplorer plotting commands were then used. HicPlotTADs can take a .h5 corrected matrix and .bw eigenvector file and provide figures with tracks showing the Hi-C matrix and eigenvector for a given region of the genome. HicPlotMatrix

was used with the male .h5 corrected matrix file to generate plots of the sex chromosomes (--chromosomeOrder X_1 Y_1 X_2 Y_2 X_3 Y_3 X_4 Y_4 X_5 Y_5; --log1p).

ChIP-seq peak calling

Macs2(12) version 2.2.6 callpeak was used to call peaks from mapped ChIP-seq reads (-f BAM; -g 1.8e9; -B; -q 0.05). Histone peaks were called as either broad or narrow based on reference material available on the ENCODE project website (https://www.encodeproject.org/chip-seq/histone/).

DNA methylation

Percentage methylation was compared with Hi-C and ChIP-seq data. Methylation data originated from previous work(13).

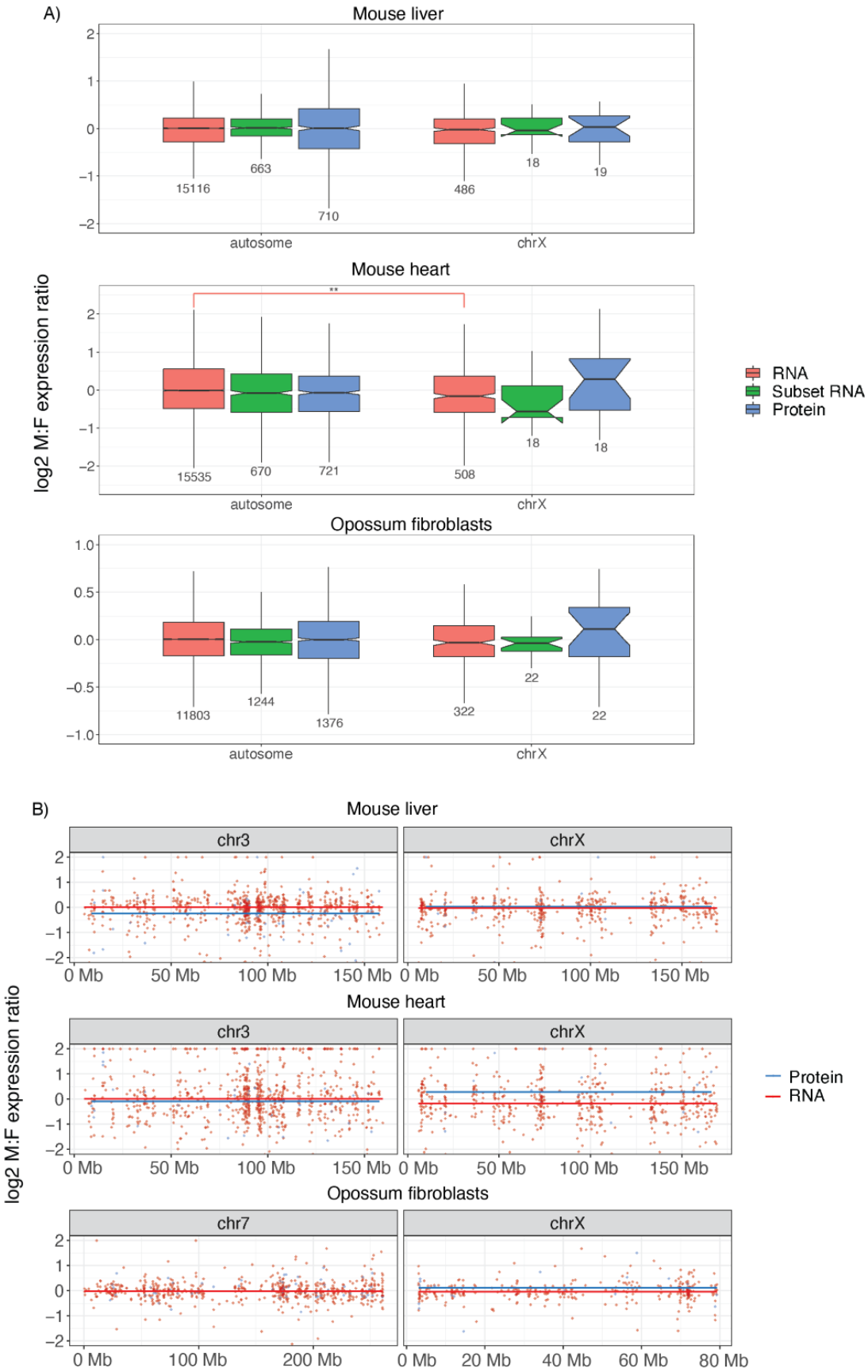

A)
Mouse liver
Mouse heart
Opossum fibroblasts
log2 M:F expression ratio
autosome
chrX
15116
663
710
486
18
19
15535
670
721
508
18
18
11803
1244
1376
322
22
22
**
RNA
Subset RNA
Protein
B)
Mouse liver
Mouse heart
Opossum fibroblasts
chr3
chrX
chr7
0 Mb
50 Mb
100 Mb
150 Mb
200 Mb
20 Mb
40 Mb
60 Mb
80 Mb
log2 M:F expression ratio
Protein
RNA

Fig. S1. Male to female expression ratios of X-borne and autosomal genes in mouse and opossum.

Median male to female (M:F) expression ratios (log2 scale) of autosomal and X borne genes in (**A**) mouse heart (n=2) and liver (n=2), and opossum fibroblasts (n=1). Ratios were calculated for all expressed mRNA (red), the subset of mRNAs sampled in the proteome (green), and proteins (blue). A ratio above zero is higher expression in males, whereas below zero is lower expression. Boxes represent the middle 50% of the data, and whiskers represent 1.5 times the interquartile range. Outliers are not plotted. Median is plotted inside the box, with the number of genes sampled below each boxplot. Mood's median test was used to calculate if ratios for the autosomes and X were statistically different (** $p < 0.001$). (**B**) Whole chromosome plots of M:F ratios for individual genes in the transcriptome (red) and proteome (blue) on the X and a representative autosome in mouse heart and liver, and opossum fibroblasts. Lines are the median M:F ratios for the whole chromosome.

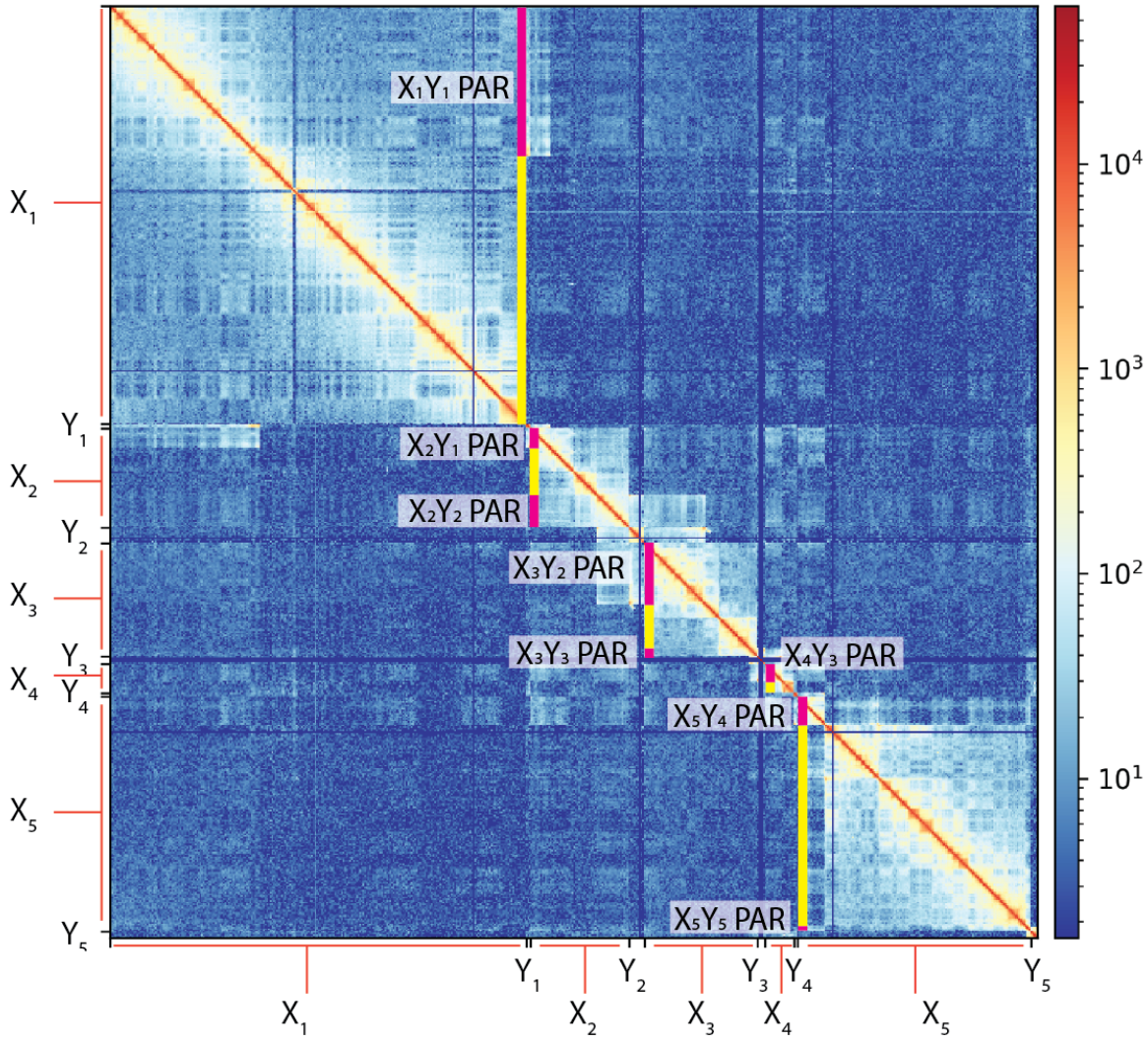


Fig. S2. HiC data of the platypus X chromosomes assembled to define PAR boundaries
The PARs in platypus are assembled on the X chromosome contigs. Y contigs in the assembly only contain Y specific sequence. Hi-C interactions between the PARs on the Xs and Y specific material were used to define PAR boundaries, with proximal PARs having increased contacts with their corresponding Ys compared to distal PAR regions. The HiC contact matrix for X and Y chromosomes in male platypus were generated using 100 kb bins. The heatmap scale indicates the number of contacts between bins. PARs are shown with purple bars and X specific regions by yellow bars.

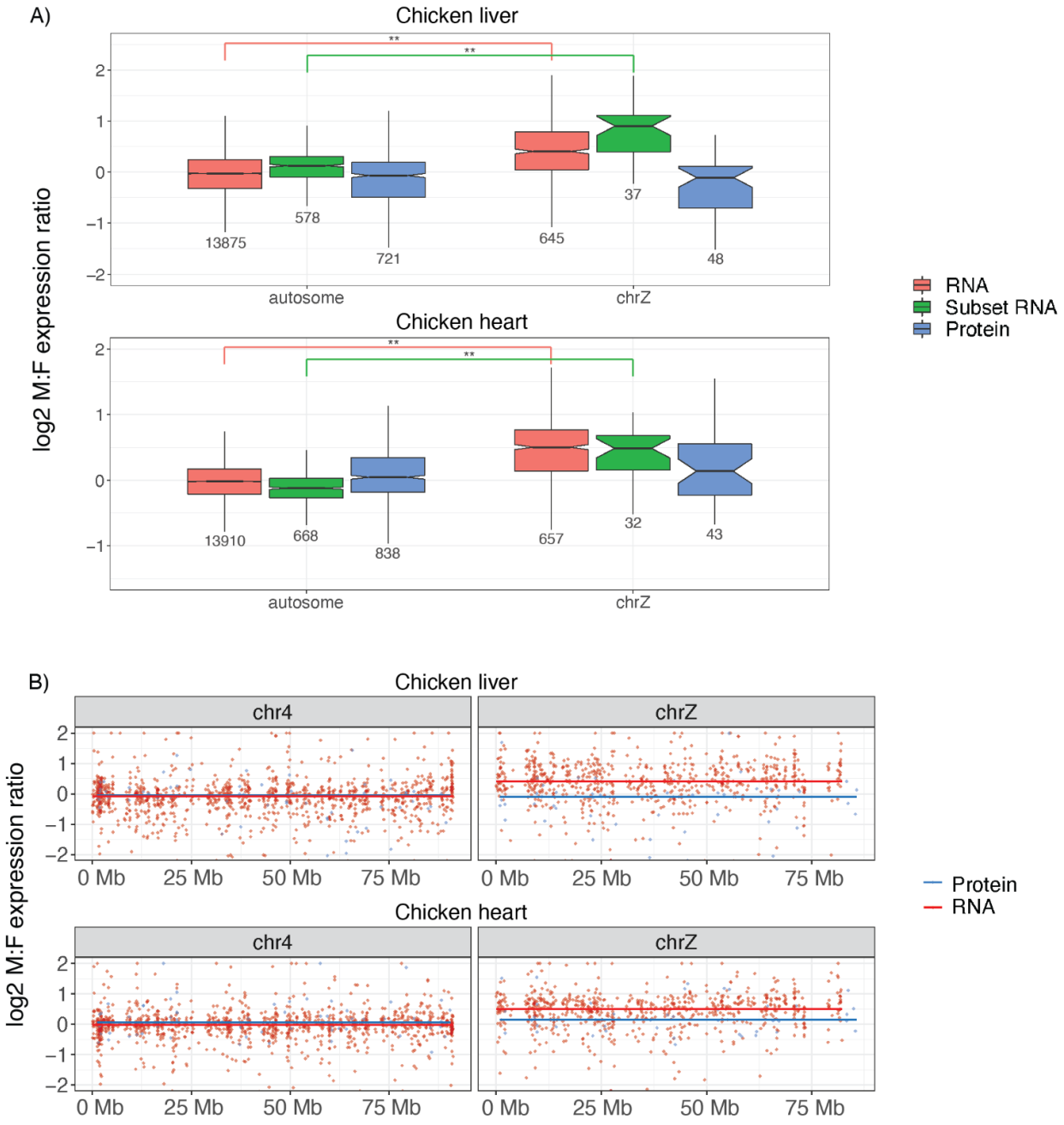


Fig. S3. Male to female expression ratios of Z-borne and autosomal genes in chicken.
Median male to female (M:F) expression ratios (log2 scale) of autosomal and X borne genes in (**A**) chicken heart (n=3) and liver (n=2). Ratios were calculated for all expressed mRNA (red), the subset of mRNAs sampled in the proteome (green), and proteins (blue). A ratio above zero is higher expression in males, whereas below zero is lower expression. Boxes represent the middle 50% of the data, and whiskers represent 1.5 times the interquartile range. Outliers are not plotted. Median is plotted inside the box, with the number of genes sampled below each boxplot. Mood's median test was used to calculate if ratios for the autosomes and X were statistically different (** $p < 0.001$). (**B**) Whole chromosome plots of M:F ratios for individual genes in the

transcriptome (red) and proteome (blue) on the Z and a representative autosome in chicken heart and liver. Lines are the median M:F ratios for the whole chromosome.

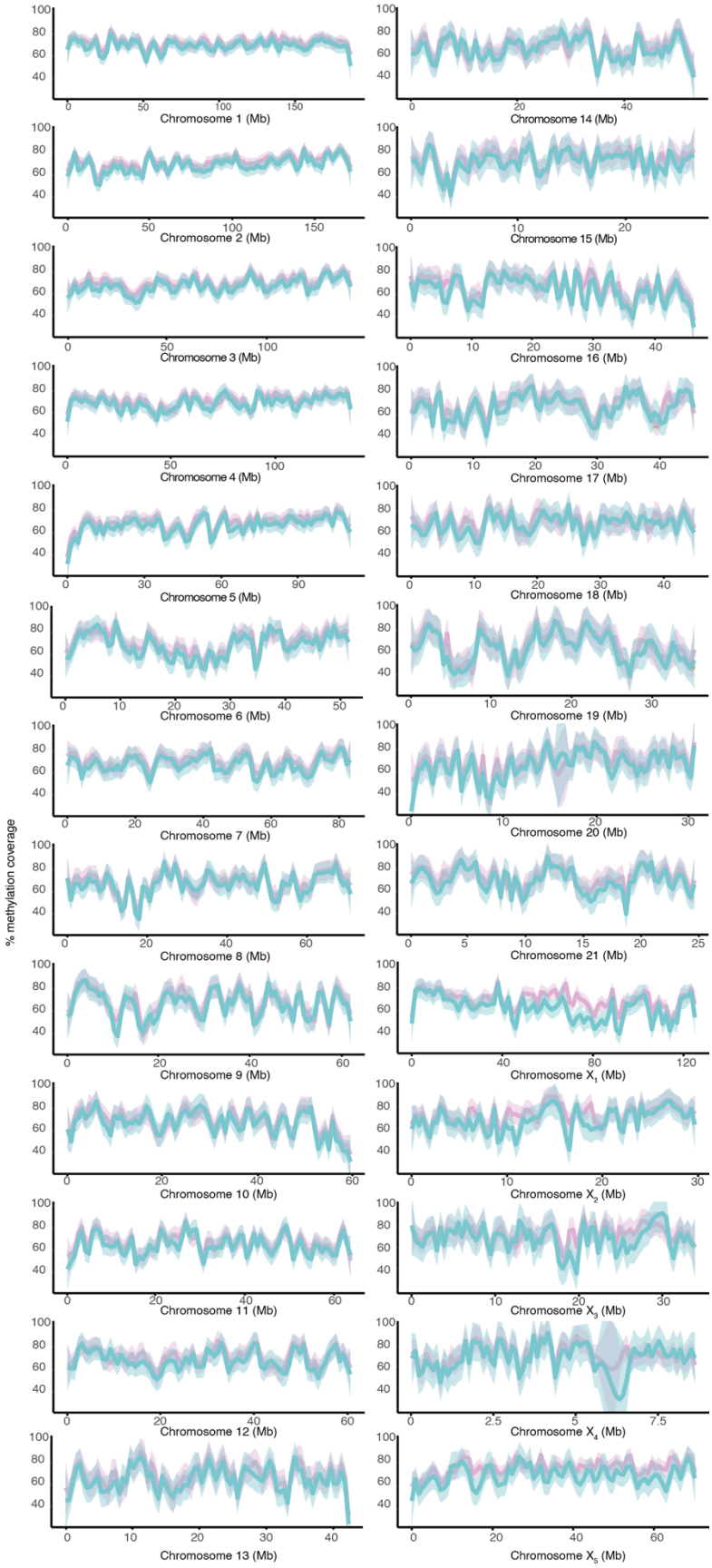
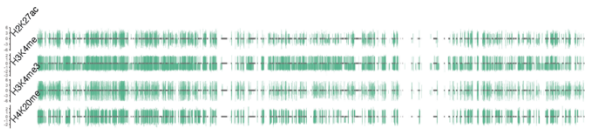
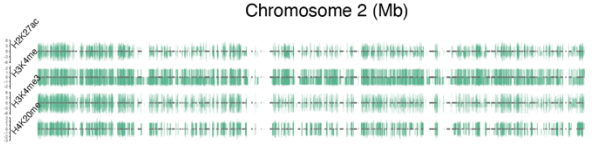

% methylation coverage
Chromosome 1 (Mb)
Chromosome 2 (Mb)
Chromosome 3 (Mb)
Chromosome 4 (Mb)
Chromosome 5 (Mb)
Chromosome 6 (Mb)
Chromosome 7 (Mb)
Chromosome 8 (Mb)
Chromosome 9 (Mb)
Chromosome 10 (Mb)
Chromosome 11 (Mb)
Chromosome 12 (Mb)
Chromosome 13 (Mb)
Chromosome 14 (Mb)
Chromosome 15 (Mb)
Chromosome 16 (Mb)
Chromosome 17 (Mb)
Chromosome 18 (Mb)
Chromosome 19 (Mb)
Chromosome 20 (Mb)
Chromosome 21 (Mb)
Chromosome X1 (Mb)
Chromosome X2 (Mb)
Chromosome X3 (Mb)
Chromosome X4 (Mb)
Chromosome X5 (Mb)

Fig. S4. DNA methylation plots

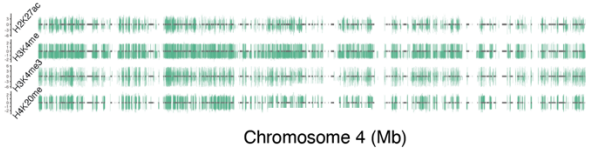
Plots showing percent DNA methylation for all chromosomes in platypus, calculated in 50 kb non-overlapping tiles. Percent DNA methylation is represented as a smoothed line in pink for female and blue for male, with a shaded 95% confidence interval.



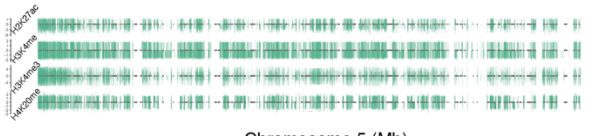
Chromosome 2 (Mb)



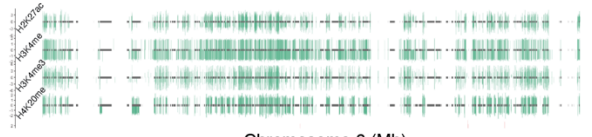
Chromosome 3 (Mb)



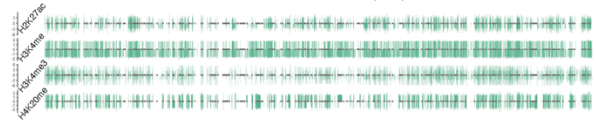
Chromosome 4 (Mb)



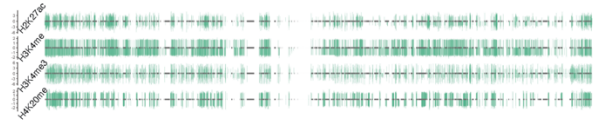
Chromosome 5 (Mb)



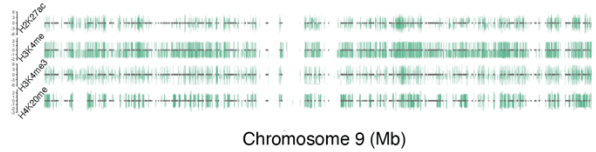
Chromosome 6 (Mb)



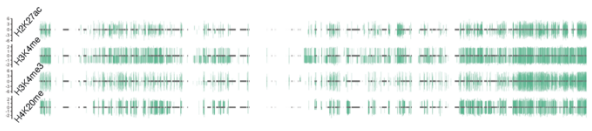
Chromosome 7 (Mb)



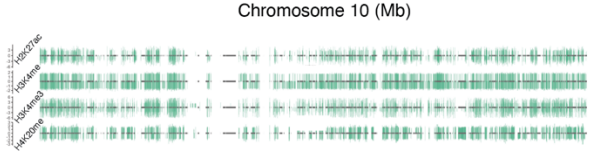
Chromosome 8 (Mb)



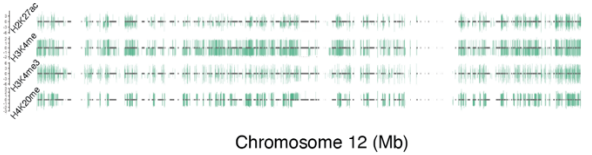
Chromosome 9 (Mb)



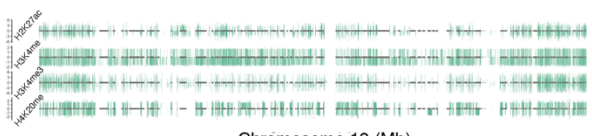
Chromosome 10 (Mb)



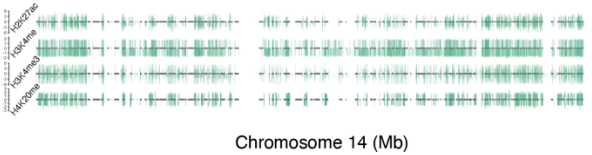
Chromosome 11 (Mb)



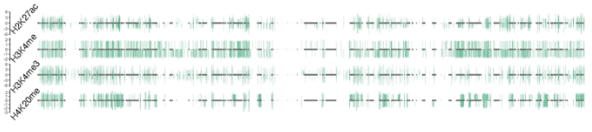
Chromosome 12 (Mb)



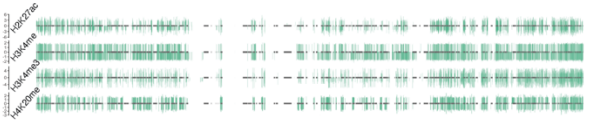
Chromosome 13 (Mb)



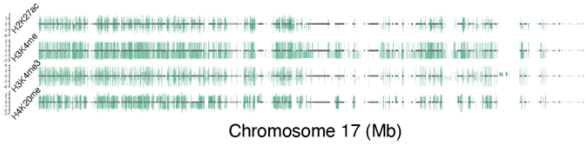
Chromosome 14 (Mb)



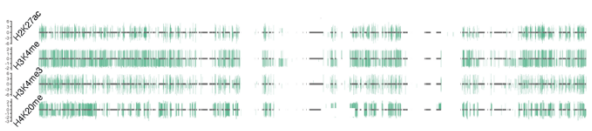
Chromosome 15 (Mb)



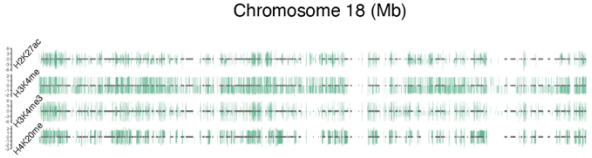
Chromosome 16 (Mb)



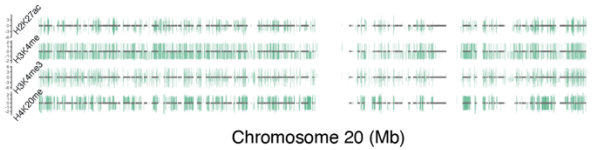
Chromosome 17 (Mb)



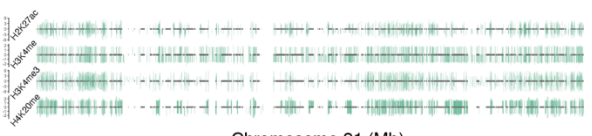
Chromosome 18 (Mb)



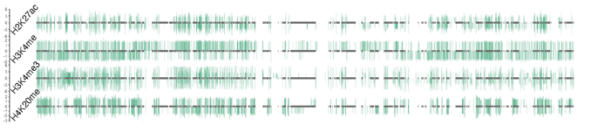
Chromosome 19 (Mb)



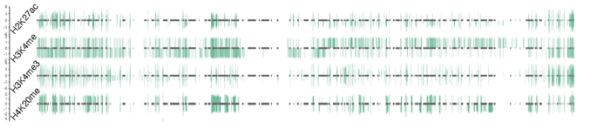
Chromosome 20 (Mb)



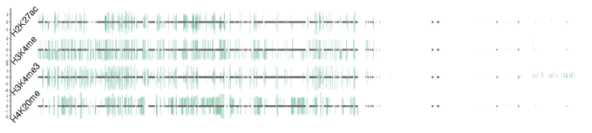
Chromosome 21 (Mb)



Chromosome X_2 (Mb)



Chromosome X_3 (Mb)



Chromosome X_4 (Mb)

Fig. S5. ChIP-seq peaks

ChIP-seq peak tracks for chromosomes 2-21 and X_2-X_4 in platypus. For each chromosome, four active histone marks are shown (from top to bottom): H3K27ac, H3K4me1, H3K4me3 and H4K20me1. For each histone mark, male and female peaks are shown above and below the centre line respectively. ChIP peaks were filtered using a threshold q-value of 0.05 and peak height was displayed up to a fold change of 5 on the y-axis. Gene locations are shown in the centre of the track as grey boxes.

Table S1. PAR boundaries from male 100kb HiC matrix

Chromosome	Start	End
X_1 PAR1 boundary	1	45 100 000
X_2 PAR1 boundary	1	5 700 000
X_2 PAR2 boundary	20 000 000	29 662 106
X_3 PAR1 boundary	1	-
X_3 PAR2 boundary	31 900 000	33 863 336
X_4 PAR1 boundary	1	5 300 000
X_5 PAR1 boundary	1	8 500 000
X_5 PAR2 boundary	69 200 000	70 139 320

Table S2. PAR boundaries from male 500kb HiC matrix

Chromosome	Start	End
X_1 PAR1 boundary	1	45 000 000
X_2 PAR1 boundary	1	5 500 000
X_2 PAR2 boundary	20 000 000	29 662 106
X_3 PAR1 boundary	1	18 500 000
X_3 PAR2 boundary	31 500 000	33 863 336
X_4 PAR1 boundary	1	5 500 000
X_5 PAR1 boundary	1	8 500 000
X_5 PAR2 boundary	69 000 000	70 139 320

Table S3. Antibodies for ChIP-seq

rabbit anti-H3K27ac	Diagenode	C15410174
rabbit anti-H3K4me1	Diagenode	C15410037
rabbit anti-H3K4me3	Merck Millipore	07-473
Rabbit anti-H4K20me1	Abcam	ab9051

Table S4. Reagents for HiC

Biotin-14-dATP-50 nmol	Thermo Fisher Scientific	19524016
cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail	SIGMA-ALDRICH	4693159001
DNA Pol. Large Fragm. (Klenow)	New England Biolabs	M0210L
Dynabeads MyOne Streptavidin T1-10 ml	Thermo Fisher Scientific	65602
NEBNext® Quick Ligation Reaction Buffer (5X)	New England Biolabs	B6058S
NEBNext® Ultra II Q5® Master Mix	New England Biolabs	M0544L
T4 DNA Ligase	New England Biolabs	M0202L
T4 DNA Polymerase	New England Biolabs	M0203L
T4 Polynucleotide Kinase NK	New England BioLabs	M0201
NEBNext Multiplex Oligos for Illumina kit	New England Biolabs	E7500
Quick Ligation™ Kit	New England Biolabs	M2200S

SI References

1. Y. Perez-Riverol *et al.*, The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* **50**, D543-D552 (2022).
2. U. S. Melo *et al.*, Hi-C Identifies Complex Genomic Rearrangements and TAD-Shuffling in Developmental Diseases. *Am J Hum Genet* **106**, 872-884 (2020).
3. S. S. Rao *et al.*, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680 (2014).
4. D. Brawand *et al.*, The evolution of gene expression levels in mammalian organs. *Nature* **478**, 343-348 (2011).
5. S. Andrews (2010) FastQC: a quality control tool for high throughput sequence data.
6. Y. Liao, G. K. Smyth, W. Shi, The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* **47**, e47 (2019).
7. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
8. I. V. Shilov *et al.*, The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Proteomics* **6**, 1638-1655 (2007).
9. J. Wolff *et al.*, Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. *Nucleic Acids Res* **46**, W11-W16 (2018).
10. J. Wolff *et al.*, Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and single-cell Hi-C data analysis, quality control and visualization. *Nucleic Acids Res* **48**, W177-W184 (2020).
11. F. Ramírez *et al.*, High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat Commun* **9**, 189 (2018).
12. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
13. S. A. Waters *et al.*, Landscape of DNA Methylation on the Marsupial X. *Mol Biol Evol* **35**, 431-439 (2018).