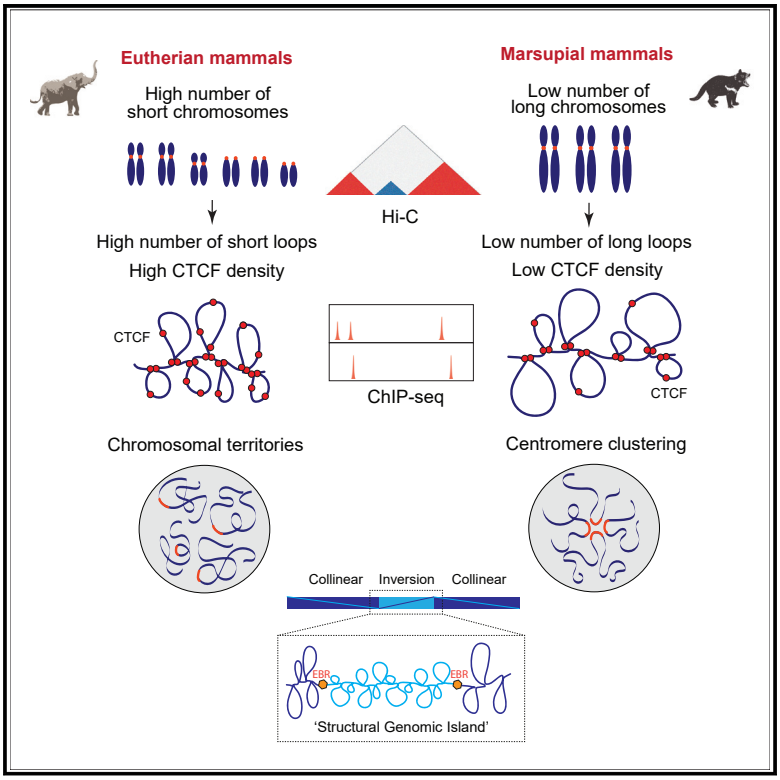
Cell Reports

Article

Principles of 3D chromosome folding and evolutionary genome reshuffling in mammals

Graphical abstract



Highlights

- Vertebrates show different patterns of genome-wide chromosomal interactions
- Long marsupial chromosomes arrange their centromeres in clusters
- Marsupials and afrotherians show contrasting patterns of genome reshuffling
- Evolutionary lineage-specific inversions have distinctive 3D structural features

Authors

Lucía Álvarez-González, Cristina Arias-Sardá, Laia Montes-Espuña, ..., Paul D. Waters, Marta Farré, Aurora Ruiz-Herrera

Correspondence

aurora.ruizherrera@uab.cat

In brief

A major challenge in genome research is to determine why some species have stable genomes whereas others have undergone extensive rearrangement. Álvarez-González et al. describe fundamental principles of 3D chromosome folding in mammals and show that lineage-specific evolutionary genomic reshuffling can influence patterns of higher-order chromatin organization.

Álvarez-González et al., 2022, Cell Reports *41*, 111839
December 20, 2022
https://doi.org/10.1016/j.celrep.2022.111839

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Principles of 3D chromosome folding and evolutionary genome reshuffling in mammals

Lucía Álvarez-González,[1,2] Cristina Arias-Sardá,[3] Laia Montes-Espuña,[1,2] Laia Marín-Gual,[1,2] Covadonga Vara,[1,2] Nicholas C. Lister,[4] Yasmina Cuartero,[5] Francisca Garcia,[6] Janine Deakin,[7] Marilyn B. Renfree,[8] Terence J. Robinson,[9] Marc A. Martí-Renom,[5,10,11,12] Paul D. Waters,[4] Marta Farré,[3] and Aurora Ruiz-Herrera[1,2,13,*]

[1]Departament de Biologia Cel·lular, Fisiologia i Immunologia, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Spain
[2]Genome Integrity and Instability Group, Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Spain
[3]School of Biosciences, University of Kent, Canterbury, Kent CT2 7NJ, UK
[4]School of Biotechnology and Biomolecular Sciences, Faculty of Science, UNSW Sydney, Sydney, NSW 2052, Australia
[5]CNAG-CRG, Centre for Genomic Regulation, The Barcelona Institute of Science and Technology, Baldiri Reixac 4, 08028 Barcelona, Spain
[6]Servei de Cultius Cel.lulars-SCAC, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Spain
[7]Institute for Applied Ecology, University of Canberra, Bruce, ACT 2617, Australia
[8]School of Biosciences, The University of Melbourne, Victoria, VIC 3010, Australia
[9]Evolutionary Genomics Group, Department of Botany and Zoology, Faculty of Science, Stellenbosch University, Private Bag X1, Stellenbosch 7602, South Africa
[10]Centre for Genomic Regulation, The Barcelona Institute for Science and Technology, Carrer del Doctor Aiguader 88, 08003 Barcelona, Spain
[11]ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain
[12]Universitat Pompeu Fabra (UPF), 08002 Barcelona, Spain
[13]Lead contact
*Correspondence: aurora.ruizherrera@uab.cat
https://doi.org/10.1016/j.celrep.2022.111839

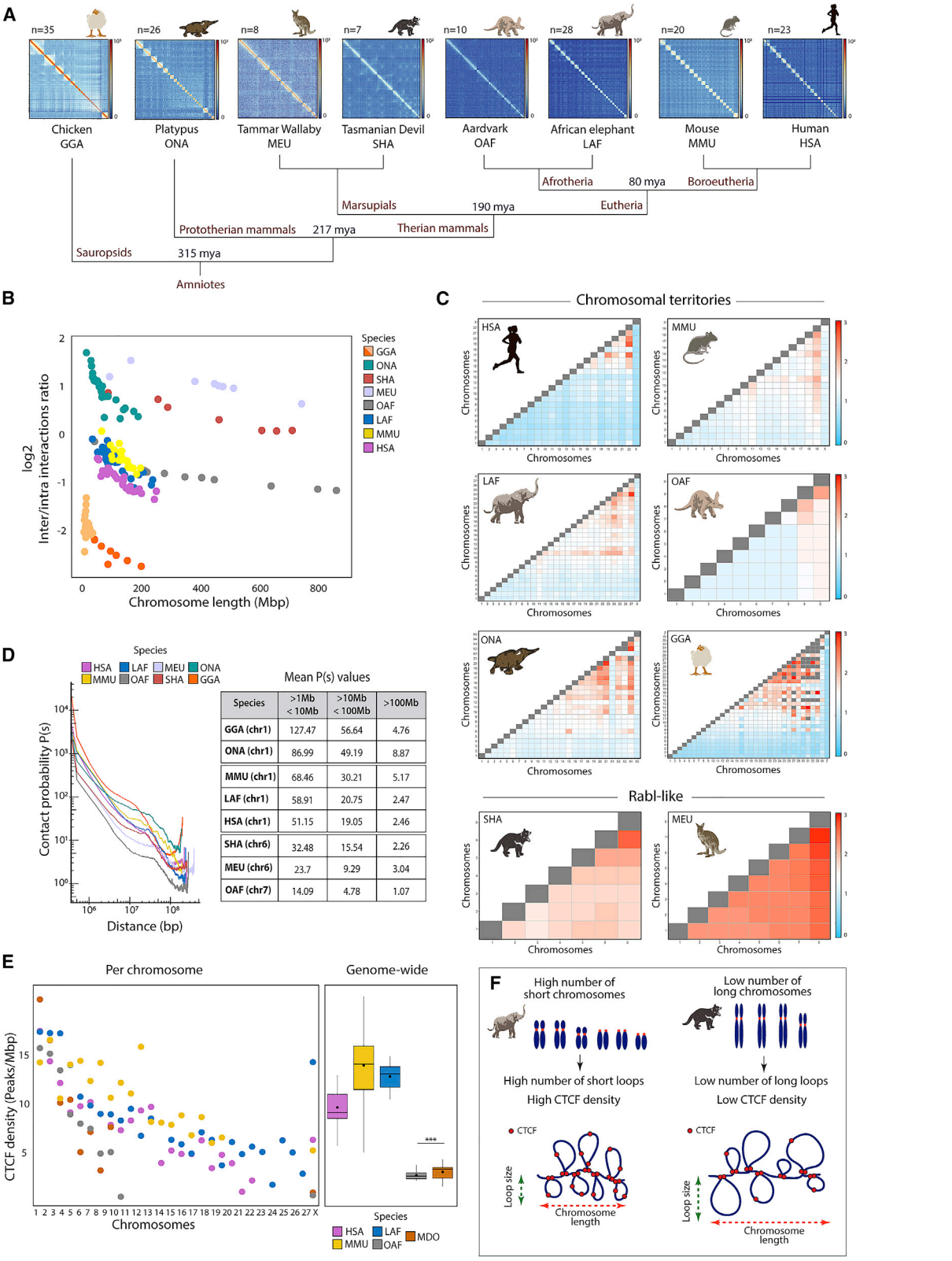
SUMMARY

Studying the similarities and differences in genomic interactions between species provides fertile grounds for determining the evolutionary dynamics underpinning genome function and speciation. Here, we describe the principles of 3D genome folding in vertebrates and show how lineage-specific patterns of genome reshuffling can result in different chromatin configurations. We (1) identified different patterns of chromosome folding in across vertebrate species (centromere clustering versus chromosomal territories); (2) reconstructed ancestral marsupial and afrotherian genomes analyzing whole-genome sequences of species representative of the major therian phylogroups; (3) detected lineage-specific chromosome rearrangements; and (4) identified the dynamics of the structural properties of genome reshuffling through therian evolution. We present evidence of chromatin configurational changes that result from ancestral inversions and fusions/fissions. We catalog the close interplay between chromatin higher-order organization and therian genome evolution and introduce an interpretative hypothesis that explains how chromatin folding influences evolutionary patterns of genome reshuffling.

INTRODUCTION

The evolution of chromatin conformation is fundamental for understanding the mechanism(s) responsible for the origin and plasticity of genome architecture. Distant loci within the genome can interact during the cell cycle to affect function in somatic and germ cells.[1–4] Exploring the similarities and differences of these genomic interactions across diverse phylogroups is central to developing an appreciation of both the dynamics of genome function and, ultimately, the effects on speciation.

Ancestral genome reconstructions have shown that structural changes disrupting synteny preferentially cluster in regions that are prone to break and reorganize—these are referred to as evolutionary breakpoint regions (EBRs).[5,6] It is also known that changes in gene expression caused by genome reshuffling may have a selective advantage through the development of new adaptive characters specific to different mammalian lineages.[7–10] These data suggest that sequence composition is not alone in determining evolutionary plasticity but rather that the occurrence and subsequent fixation of genome rearrangements are multifaceted, involving (1) repetitive elements (i.e., making DNA more susceptible to chromosomal reorganization)[11–13]; (2) functional constrains (i.e., genes related to species-specific phenotypes)[14]; and (3) genome folding dynamics and its effect on gene regulation/function.[3,15,16] This has led to suggestions that the permissiveness of some genomic regions to undergo genomic rearrangements, especially in germ cells, is influenced by chromatin 3D conformation.[4,15–18]



(legend on next page)

Genomes are compartmentalized into different hierarchical layers including chromosomal territories (CTs), compartments (A and B), topologically associated domains (TADs), and looping interactions.[1,19,20] How these different levels of chromatin organization are conserved across species underpins their potential evolutionary genome plasticity. For example, chromatin 3D organization was recently studied in representative species of chordates, plants, and fungi.[21] Two types of 3D genome organization were found at a chromosomal scale: telomeres and centromeres either (1) clustered across chromosomes adopting a Rabl-like configuration or (2) were oriented in a polarized state maintaining individual CTs within the cell. However, little is known about the evolutionary dynamics of 3D genome organization, especially at the root of the three major lineages of mammals, Prototheria (monotremes), Metatheria (marsupials), Eutheria, and their presumptive ancestor.

Monotremes (represented by the duck-billed platypus and the echidnas) are positioned phylogenetically between birds/reptiles and therians and diverged from therian mammals (placentals and marsupials) ≈217 million years ago (mya).[22] Monotremes represent a pivotal group with a mix of reptilian and mammalian morphological, physiological, and karyological features.[23] Marsupials, on the other hand, shared a common ancestor with eutherian mammals ≈190 mya.[22] Karyotyping studies observed a bimodal distribution of diploid chromosome numbers across the marsupial phylogeny, with many species having either a 2n = 14 or a 2n = 22 karyotype.[24]

In eutherians, Afrotheria represents one of the most ancient clades that includes six mammalian orders all with an Afro-Arabian origin. Afrotherian species exhibit extreme morphological diversity and niche preference, which is thought to result from the long period of isolation when Africa was an island continent 105–125 mya.[22] Genome organization within Afrotheria is diverse, with diploid numbers ranging from 2n = 56 in the African (*Loxodonta africana*) and Asian (*Elephas maximus*) elephants to 2n = 20 in the aardvark (*Orycteropus afer*).[25] Given the position of Afrotheria and marsupials near the root of therian mammals and their diverse diploid numbers (reflecting extensive genome reshuffling), the analysis of 3D genome architecture provides a unique opportunity to further understand the mechanisms underpinning mammalian chromosomal evolution.

To explore the principles of 3D genome folding dynamics and evolution in distantly related mammals, a comprehensive computational analysis that included genome-wide chromosome conformation capture (*in situ* Hi-C) in representative therian (eutherian and marsupials) species coupled with comparative genomics, transcriptome sequencing (RNA sequencing [RNA-seq]), and chromatin immunoprecipitation sequencing (ChIP-seq) of CTCF and H3K4me3 was performed. We (1) defined patterns of 3D genome folding in interphase nuclei across species; (2) reconstructed ancestral marsupial and Afrotherian genomes by analyzing whole-genome sequences of 10 species that represent the major therian phylogroups; (3) detected lineage-specific chromosome rearrangements; (4) identified the dynamics and structural properties of therian EBRs by applying integrative computational analyses; and (5) provide evidence of chromosome folding changes due to inversions and fusions/fissions. Our results represent a comprehensive catalog of the close interplay between chromatin higher-order organization dynamics and therian genome evolution.

RESULTS

Patterns of genome-wide chromosomal interactions across vertebrates

We first explored the characteristic features of 3D genome evolution across vertebrates performing *in situ* Hi-C experiments in primary fibroblast cell lines from the Afrotheria (African elephant with 2n = 56 and aardvark with 2n = 20) and Marsupialia (Tasmanian devil with 2n = 14 and tammar wallaby with 2n = 16) (Figure 1A). After filtering the raw Hi-C interactions, an average of 100 million valid interactions were obtained per species (Table S1). The comparison between biological replicates resulted in reproducible Hi-C maps (Pearson correlation, $R^2 > 0.8$, $p < 0.01$; Figure S1). African elephant, aardvark, and tammar wallaby Hi-C data were mapped against their respective reference genome available at the DNA Zoo consortium.[26,27] These data were combined with publicly available Hi-C data for human (2n = 46),[1] mouse (2n = 40),[3] platypus (2n = 52),[28] and chicken (2n = 70)[29] (see STAR Methods and Table S1).

Comparison of Hi-C matrices revealed different patterns of chromosomal interactions as reflected by the log2 inter-/intra-chromosomal interaction ratios in different species (Figure 1B). Chicken (a bird representative) presented the lowest log2 interaction ratios per chromosome (<–1) (Figure 1B). This was followed by eutherian representatives (human, mouse, African

Figure 1. Patterns of chromosome folding

(A) Whole-genome Hi-C contact maps for chicken (*Gallus gallus* [GGA])[29]; platypus (*Ornithorhynchus anatinus* [ONA])[28]; tammar wallaby (*Macropus eugenii* [MEU]; this study, one biological replicate); Tasmanian devil (*Sarcophilus harrisii* [SHA]; this study, two biological replicates); aardvark (*Orycteropus afer* [OAF]; this study, two biological replicates); African elephant (*Loxodonta africana* [LAF]; this study, two biological replicates); mouse (*Mus musculus* [MMU])[3]; and human (*Homo sapiens* [HSA]).[1] Haploid number of chromosomes (n) is shown for each species. Divergence times are based on Meredith et al.[22]

(B) Log2 ratio of inter-/intra-chromosomal interactions according to chromosome length (in Mbp). In the case of chicken (GGA), color intensity differentiates between macro (bold orange) and micro (light orange) chromosomes.

(C) Representation of inter-chromosomal mean interactions per chromosome for human, mouse, African elephant, aardvark, platypus, chicken, Tasmanian devil, and tammar wallaby.

(D) Chromosome-specific contact probability P(s) as a function of genomic distance in all species analyzed. All selected chromosomes are ~200 Mpb in length. Mean P(s) values are represented for three chromosome segments: 1–10, 10–100, and more than 100 Mbp.

(E) CTCF density (expressed as the number of peaks per Mbp) per chromosome and genome wide for human,[30] mouse,[30] African elephant (this study, two biological replicates), aardvark (this study, two biological replicates), and opossum as a marsupial representative.[30] Boxplots are presented as median values (center line); mean values (dot) ± SD. Asterisks represent statistically significant CTCF density between species (two-sided t test, ***$p < 0.001$).

(F) Representation of DNA loop size, chromosome length, and CTCF density for species with high (left panel) and low (right panel) number of chromosomes.

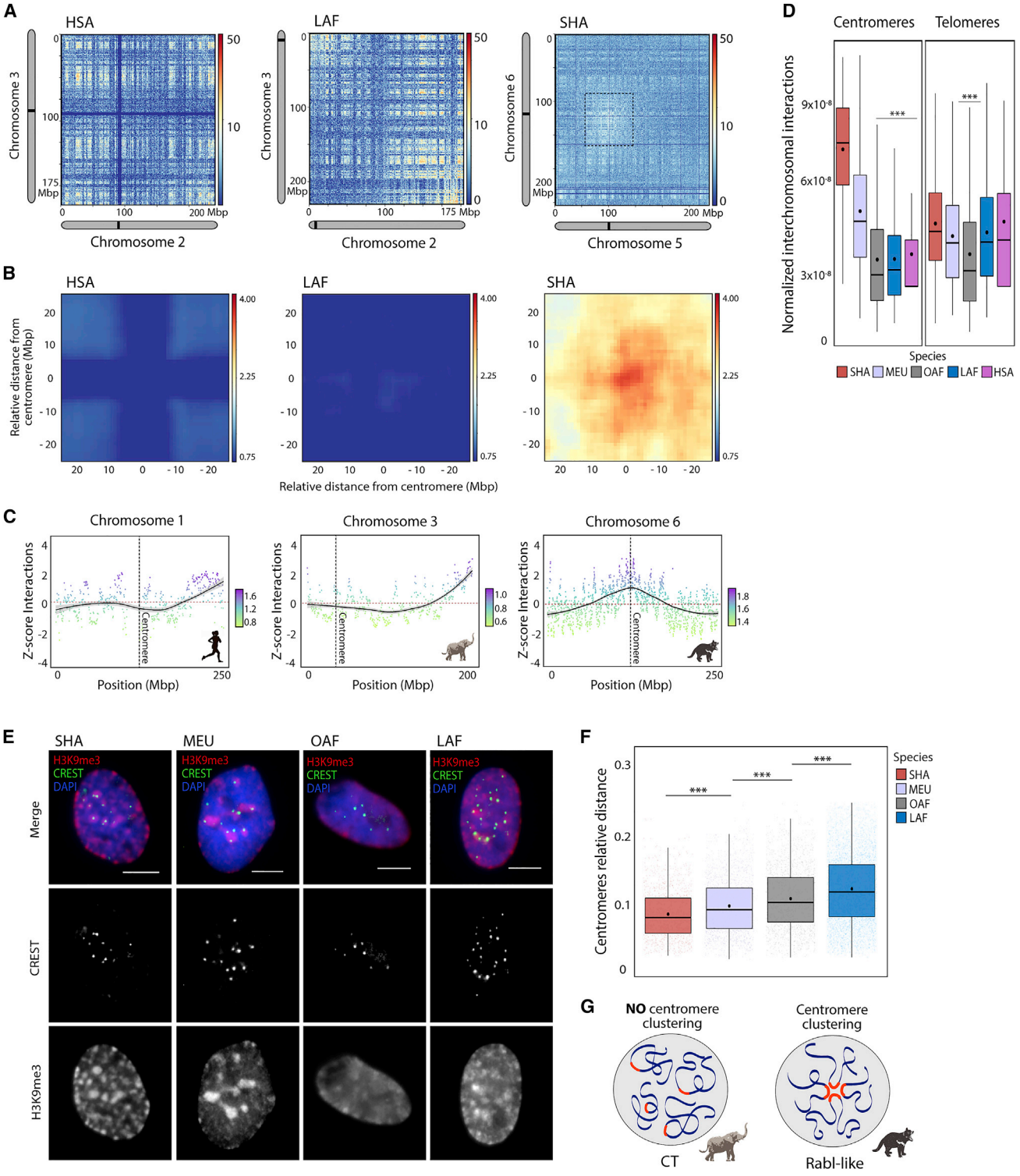


Figure 2. Centromere clustering

(A) Heatmaps showing normalized inter-chromosomal interactions between heterologous chromosomes in human (chromosomes 2 and 3), African elephant (chromosomes 2 and 3), and aardvark (chromosomes 5 and 6). The position of centromeres is presented as black lines for each chromosome.

(B) Centromere genome-wide aggregated contact plots for human, African elephant, and Tasmanian devil. Note the high number of interactions at the centromeric level in the Tasmanian devil.

(legend continued on next page)

elephant, and aardvark), which showed log2 inter-/intra-chromosomal interaction ratios between –1 and 0 (irrespective of chromosome number), suggesting well-defined CTs (Figure 1B). In contrast, log2 inter-/intra-chromosomal interaction ratios increased approximately 2-fold (from 0 to 2) in both marsupial species (tammar wallaby and Tasmanian devil) and the platypus, suggesting a chromosome compartmentalization pattern that was distinct from eutherian mammals. In fact, marsupials had the highest percentage of inter-chromosomal interactions (60% of total), followed by platypus and eutherian mammals (both 40%) and chicken (20%).

Moreover, we detected a general trend for smaller chromosomes (<5% genome size) to show higher inter-chromosomal interactions (mean of inter-chromosomal interactions >1.5) than larger chromosomes (defined as >5% genome size) in all species (Figure 1C). In contrast, marsupials displayed high interactions between all chromosomes, including the X chromosome (Figure 1B), also supporting the view that chromosome compartmentalization is distinct from that observed in other vertebrates.

Diploid numbers determine patterns in chromosome folding

Distance-dependent interaction frequencies represented as probability versus distance (P(s)) curves were compared between species to analyze patterns of chromosome folding (Figure 1D). As previously described,[19] we detected a general decrease in interaction frequencies as genomic distances increased. Visual inspection allowed the delineation of distinct groups based on the decay of contact probability interactions: (1) chicken, (2) platypus, (3) human, African elephant, and mouse, and (4) both marsupial species plus aardvark (Figure 1D). At short genomic distances (<10 Mbp), chicken chromosomes showed the highest contact probability values genome wide (133.05), followed by platypus, human, African elephant, and mouse (values ranged from 95 to 57.64). Interestingly, the three species with the lowest numbers of chromosomes (aardvark, tammar wallaby, and Tasmanian devil) also showed the lowest contact probability values at distances <10 Mbp (from 17.57 to 35.26). At intermediate distances (between 10 and 100 Mbp), platypus diverged from eutherian mammals with a gradual decay in P(s) curves, and marsupials and aardvark showed the lowest interaction frequencies (from 2.85 to 7.51).

We similarly examined whether CTCF peak densities and distributions were related to chromosome number across species. To do so, we performed CTCF ChIP-seq in African elephant and aardvark fibroblasts (see STAR Methods). We detected a total 37,210 CTCF peaks in African elephant and 12,478 peaks in aardvark. These data were combined with publicly available ChIP-seq data for human, mouse, and opossum.[30] We detected a mean density of CTCF peaks/Mbp in eutherian mammals (with high diploid numbers, i.e., 2n = 40–56) that ranged from 8.19 peaks/Mbp in human to 12.25 peaks/Mbp in mouse (Figure 1E). In contrast, CTCF density decreased approximately 4-fold in aardvark (2n = 20) and opossum (2n = 14) (2.97 and 3.55 peaks/Mbp, respectively; two-sided t test, $p < 0.01$). Collectively, these observations suggest that genomes organized into low diploid numbers (and hence longer chromosomes) have reduced contact interaction frequencies at shorter genomic distances (resulting from low CTCF density; Figures 1F and S2).

Divergent centromere clustering in marsupials

Close inspection of genome-wide Hi-C contact maps revealed striking patterns of chromosomal interactions at centromeres. Both Hi-C interaction maps (Figure 2A), aggregate peak analysis (Figure 2B), and *Z* score interaction ratio plots between heterologous chromosomes (Figure 2C) detected higher centromeric inter-chromosomal interactions ($>4 \times 10^{-8}$ chromosome-length-normalized mean interaction value) in marsupials than in therian mammals. This contrasted with interactions detected in the telomeric regions, which were high for all species except for the aardvark ($>3 \times 10^{-8}$ chromosome length normalized mean interaction value; Figure 2D). The pattern of centromeric interaction in marsupials was further demonstrated by the immunodetection of the centromeric constitutive heterochromatin (i.e., H3K9me3 signal) in fibroblast cell cultures (Figures 2E–2F). Both tammar wallaby and Tasmanian devil fibroblasts showed shorter relative distances between the centromeres (i.e., CREST signal; see STAR Methods) of heterologous chromosomes (<0.1 relative distance; two-sided t test, $p < 0.01$) than did aardvark and African elephant (>0.1 relative distance; two-sided t test, $p < 0.01$) (Figure 2E).

Collectively, these results suggest that marsupial heterologous centromeres cluster to adopt a Rabl-like configuration as previously described in the peach potato aphid,[31] yellow fever mosquito,[26] sea urchin,[32] and yeast.[33] This configuration, along with the high inter-/intra-chromosomal interaction ratios (Figure 1B), low interaction values at short and medium genomic distances (Figure 1D), and low CTCF density (Figure 1E), suggest that marsupials have a distinct mechanism of chromosome folding.

In light of these results, it is plausible that the number of inter-chromosomal interactions observed in different taxa is not

(C) *Z* score interactions as a function of genomic distance for representative chromosomes of human, African elephant, and Tasmanian devil. Dashed horizontal lines are centromere locations.

(D) Normalized inter-chromosomal interactions for centromeres (from the centromere up to 3 Mbp, left panel) and telomeres (from the telomere up to 3 Mbp, right panel) for human, mouse, African elephant, aardvark, Tasmanian devil, and tammar wallaby. Boxplots are presented as median values (center line); mean values (dot) ± SD. Asterisks represent statistically significant different interactions between species (two-sided t test, ***$p < 0.001$).

(E) Representative immunofluorescence images of fibroblasts nuclei showing centromeres in green (CREST), H3K9me3 (marker for the constitutive heterochromatin) in red, and chromatin in blue (DAPI). Scale bars, 10 μm.

(F) Boxplots depicting relative distance between centromeres fibroblasts nuclei of African elephant (n = 6,067 measurements); aardvark (n = 1,865 measurements); Tasmanian devil (n = 1,280 measurements); and tammar wallaby (n = 1,675 measurements). Boxplots are presented as median values (center line); mean values (dot) ± SD. Asterisks represent statistically significant differences between species (two-sided t test, ***$p < 0.001$).

(G) Representation of chromosome position inside nuclei of species showing chromosomal territories (CTs) and no centromere clustering (left panel) or centromere clustering (right panel). Centromeres are depicted in orange. Adapted from Hoencamp and collaborators.[21]

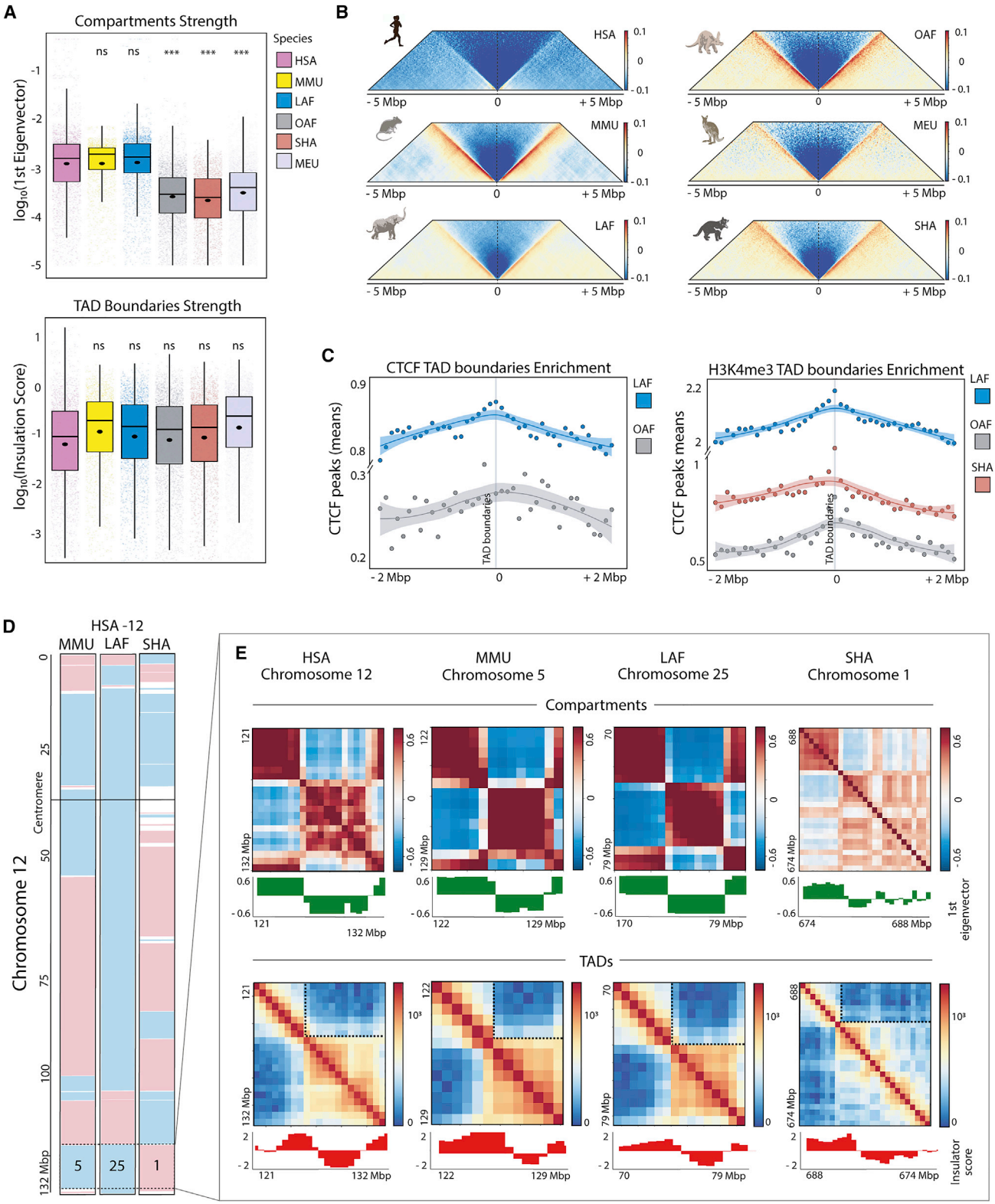


Figure 3. Conservation of the higher-order chromatin organization within mammals

(A) Boxplots depicting first eigenvector values (log scale) and insulator score values (log scale) for human (HSA), mouse (MMU), African elephant (LAF), aardvark (OAF), Tasmanian devil (SHA), and tammar wallaby (MEU). Boxplots are presented as median values (center line); mean values (dot) ± SD. Asterisks represent statistically significant values between species (two-sided t test, ***$p < 0.001$). ns, non-significant.

(legend continued on next page)

directly related to chromosomal size but to chromosomal position within the cell nucleus. That is, within the same phylogenetic grouping, all species have the same CT type and compartmentalization regardless of chromosome number (Figure 2G). Thus, eutherian mammals have their chromosomes organized in territories with small chromosomes positioned toward the center of the cell. This is exemplified by the observation that shorter eutherian chromosomes show high inter-chromosomal interactions, but the inter-/intra-interactions ratios remains the same for all species. In contrast, marsupials show high rates of inter-/intra-interactions ratios (irrespective of chromosome size) with a clear pattern of centromere clustering.

Genomic compartmentalization features within mammals

We scored A/B compartments and TADs detected in the various species. The genome-wide distribution of A/B compartments was consistent across taxa, where ≈50% of the genome corresponded to A compartments ("open") being positively associated with the active histone H3K4me3 (permutation test based on 10,000 permutations, normalized Z score > 0.01, $p < 0.05$). Moreover, average compartment length was consistent between species (mean value ~2.5 Mbp). However, when analyzing compartment strength (first eigenvector values; see STAR Methods), species with longer chromosomes (and low diploid numbers, i.e., tammar wallaby, Tasmanian devil, and aardvark) had significantly lower first eigenvector values (two-sided t test, $p < 0.01$) (Figure 3A).

We identified TADs and examined the robustness of their boundaries at 500 Kbp resolution using TADbit.[34] Depending on the species, between 1,882 and 2,105 TADs of 1 Mbp average length were detected (TAD strength score >6). TAD boundary scores were equivalent between species, and metaborder plots showed clear insulator patterns for all taxa (Figure 3B). All detected TAD boundaries were associated, and enriched, with H3K4me3 and CTCF (permutation test based on 10,000 permutations, normalized Z score > 0.01, $p < 0.05$; Figure 3C).

Conservation of the higher-order chromatin organization within mammals

To test if the higher-order structural organization of mammalian genome architecture was conserved in somatic cells, we established homologous syntenic blocks (HSBs) between single representatives of the boreoeutherians (mouse), afrotherians (African elephant), and marsupials (Tasmanian devil) using the human genome as reference (see STAR Methods; Table S2). At a 300 Kb resolution, a total of 346 HSBs were detected between human and mouse (ranging from 7.7 to 78.3 Mbp in size and covering 83.41% of the human genome; see STAR Methods). In the African elephant and Tasmanian devil, the number of HSBs detected was 230 and 444, respectively, and these covered between 86.15% and 80.91% of the human genome.

We next examined whether conservation of DNA sequence correlated with the conservation of higher-order chromatin organization of the targeted species (Figure 3D) (see STAR Methods). The analysis included a total of 140 HSBs in African elephant, 212 HSBs in Tasmanian devil, and 223 HSBs in mouse covering, respectively, 83.74%, 74.77%, and 79.86% of the human genome. A pairwise evaluation of the higher-order structural conservation of compartments was conducted by comparing first eigenvector values between HSBs, whereas TADs were evaluated by insulator scores (see STAR Methods).

Within species from the same phylogenetic group, we detected that the majority of HSBs (70%) had the same level of higher-order structural conservation at the compartment level. That is, human versus mouse (Boreoeutheria) (73.6% of the genome conserved), African elephant versus aardvark (Afrotheria) (75.1% of the genome conserved), and Tasmanian devil versus tammar wallaby (Marsupialia) (68.7% of the genome conserved) (Figure S3). However, as the evolutionary distance increased, structural conservation decreased between different phylogenetic groups (i.e., humans versus species other than mouse; Figure 3E). In Afrotheria (the African elephant), 60% of compartments were conserved with human, whereas in marsupials, the levels of conservation decreased to 50% (Figure 3F; Table S3). Surprisingly, no differences were observed for TADs at 500 Kbp resolution (>90% of which were conserved in HSBs) irrespective of whether this was for comparisons within the same phylogenetic group or between different phylogenetic groups (Figures 3E and S3; Table S4).

Evolutionary history of marsupial and afrotherian chromosomes

To assess whether patterns of chromosome folding correlated with genomic reshuffling during evolution, we reconstructed ancestral karyotypes and cataloged the evolutionary history of chromosome rearrangements in afrotherians and marsupials.

Two ancestral karyotypes were reconstructed for Marsupialia using five representative marsupial genomes (Tasmanian devil, tammar wallaby, wombat, red kangaroo, and opossum) and the genomes of five outgroup taxa (sloth, African elephant, human, platypus, and chicken) (Figures 4A and S4; Table S2; see STAR Methods). The marsupial ancestral karyotype (MAK) comprised 10 reconstructed ancestral chromosome fragments (RACFs), representing seven ancestral chromosomes that

(B) Metaplots for all TAD boundaries detected in human, mouse, African elephant, aardvark, Tasmanian devil, and tammar wallaby.
(C) Mean number of CTCF peaks and H3K4me3 relative to TAD boundaries positions in African elephant, aardvark, and Tasmanian devil. The data include two biological replicates per species.
(D) Chromosomal synteny between human (HSA) chromosome 12 in mouse (MMU), African elephant (LAF), and Tasmanian devil (SHA). Collinear homologous regions are depicted in light blue and inverted regions in pink.
(E) Zoom in of a structural conserved HSB (two-sided, t test, $p < 0.01$; Tables S3 and S4) human chromosome 12 when compared with mouse (a chromosomal region of chromosome 5), African elephant (a chromosomal region of chromosome 25), and Tasmanian devil (a chromosomal region of chromosome 1). Compartment conservation is represented as a compartment heatmap and first eigenvector distribution. TAD conservation is represented as a contacts heatmap and insulator score distribution.

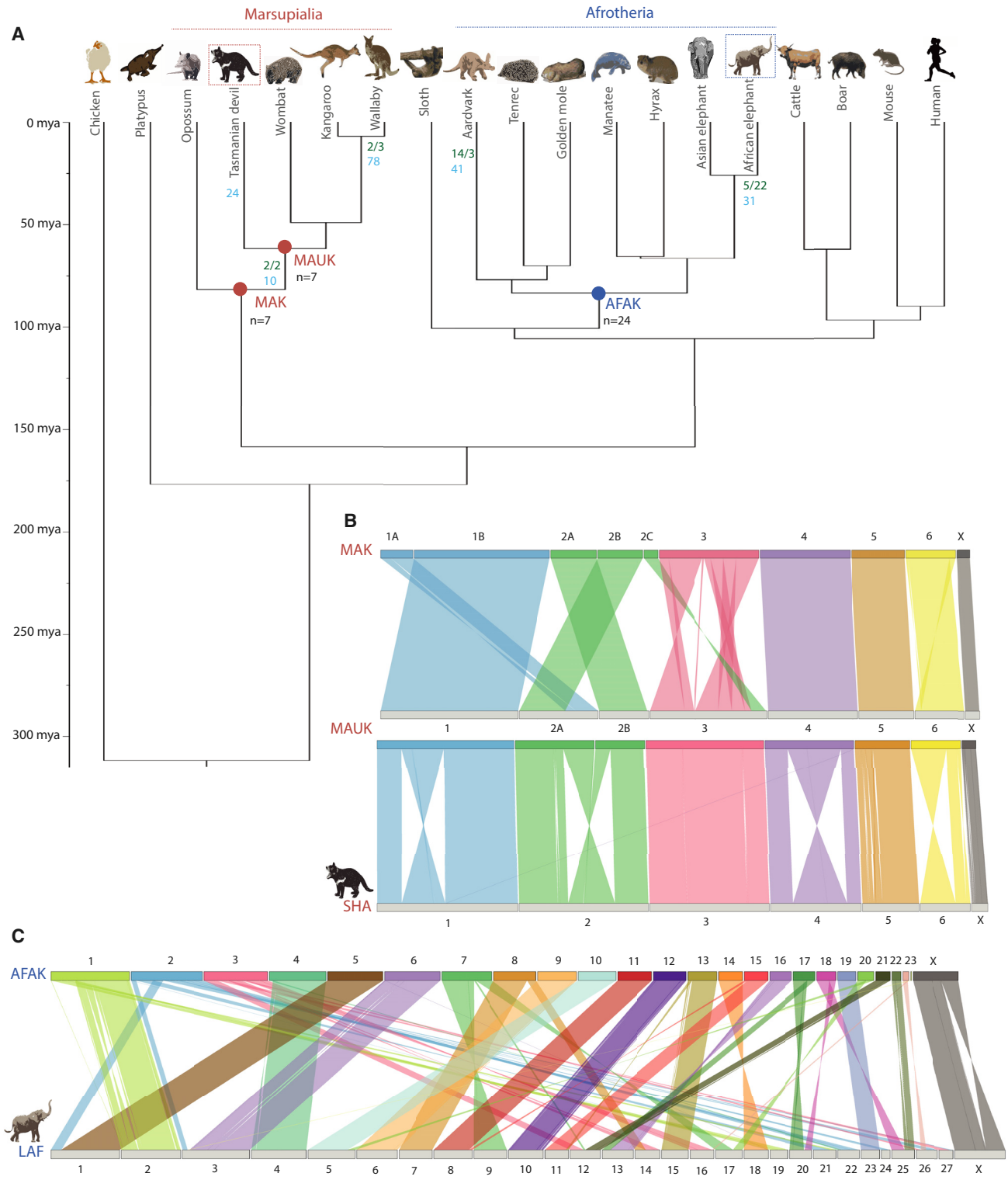


Figure 4. Marsupial and afrotherian ancestral karyotypes and genome reshuffling

(A) Phylogenetic tree of the species included in the analysis. Marsupial species are shown in red, while afrotherians are in blue. The number of intra-chromosomal rearrangements in each branch is shown in pale blue, while inter-chromosomal (fissions/fusions) rearrangements are in dark green.

(B and C) Marsupial ancestral karyotypes (B) and Afrotherian ancestral karyotype (C). Ribbons connecting chromosomes indicate orthologs between ancestors and reference genome chromosomes, with a twist indicating an inversion. MAK, marsupial ancestral karyotype; MAUK, Australian marsupial ancestral karyotype; SHA, Tasmanian devil; LAF, African elephant.

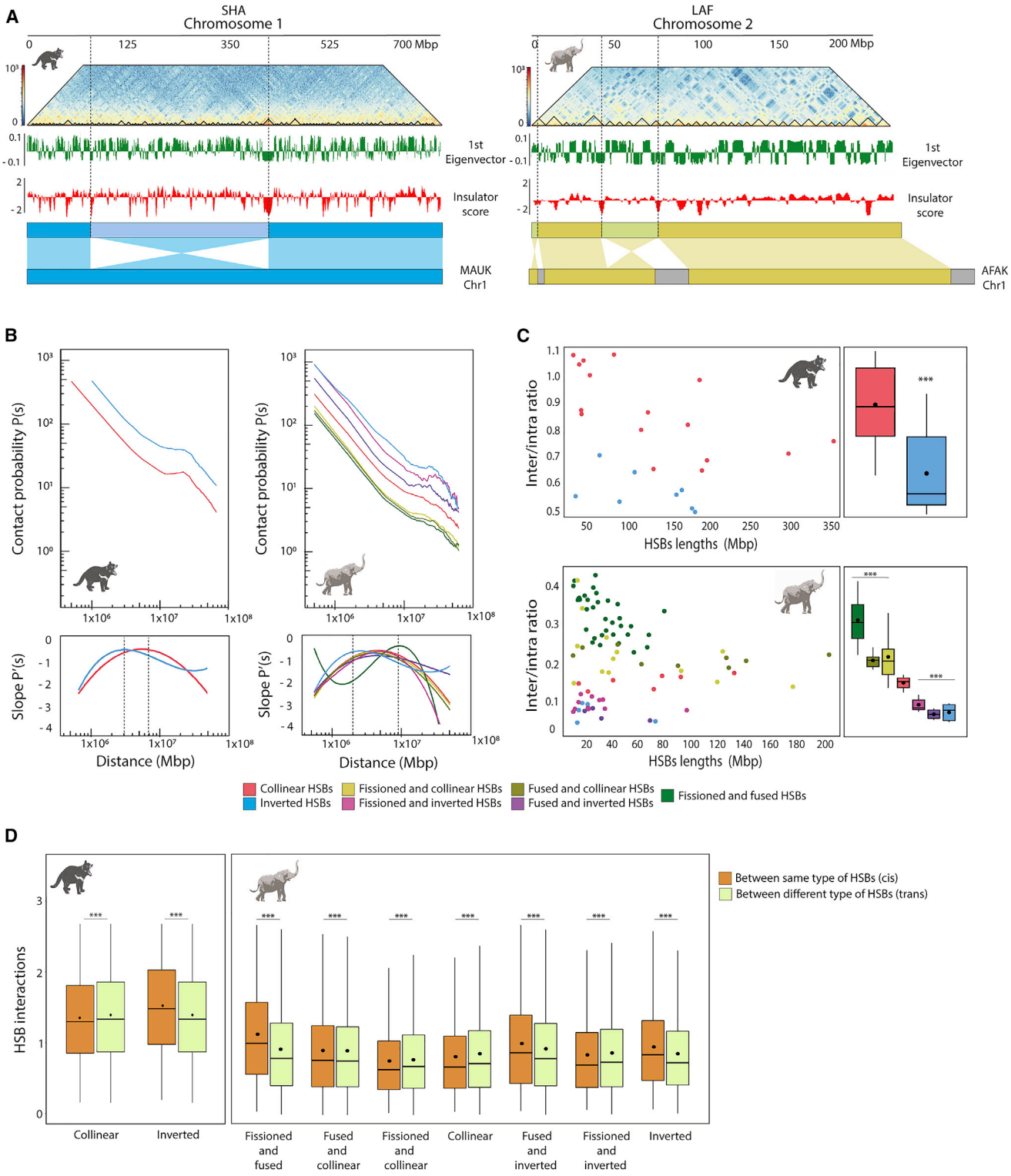


Figure 5. Structural plasticity of lineage-specific chromosomal reorganizations

(A) Chromosome 1 and chromosome 2 region-specific 500 Kbp heatmaps, first eigenvector, insulator score, and ancestral specific reorganizations for Tasmanian devil and African elephant.

(B) Contact probability P(s) as a function of genomic distance and its derivative for each class of reorganized HSBs in Tasmanian devil and African elephant.

(legend continued on next page)

covered 95.87% of the Tasmanian devil reference genome (Figures 4B and S4). Our MAK reconstruction was consistent with previous work,[24] recovering 85.16% of the syntenic associations between MAK and human. Similarly, the ancestor of Australian marsupials (MAUK) contained eight RACFs, representing seven ancestral chromosomes that ranged in size from 71.6 to 698 Mb in size. Our results illustrate that the MAK and MAUK karyotypes were conserved, with three chromosomes maintained as complete syntenic blocks. Ten inversions, two fissions, and two fusions separated MAUK from MAK, whereas 24 inversions occurred between MAUK and the Tasmanian devil, while 77 inversions, three fissions, and two fusions separated MAUK from the tammar wallaby. In summary, these data show that the marsupial lineage is predominantly characterized by intra-chromosomal rearrangements.

We also reconstructed the ancestral karyotype for Afrotheria (AFAK) using the African elephant as the reference genome and six other afrotherian genomes (Asian elephant, aardvark, cape golden mole, rock hyrax, West Indian manatee, and lesser hedgehog tenrec), along with those from three outgroup species (cattle, pig, and human) (Figures 4C and S4; Table S2; see STAR Methods). The AFAK consisted of 24 ancestral chromosomes in 25 RACFs, covering 95.44% of the African elephant genome (Figures 4C and S4). We recovered 71.43% of the syntenic associations between human and African elephant that were previously identified by cross-species chromosome painting,[35] as well as all 23 fissions of human chromosomes, indicating that our reconstruction was accurate. A total of 31 inversions, 22 fissions, and five fusions separate the African elephant from the AFAK. In the aardvark lineage, 41 inversions, three fissions, and 14 fusions occurred after the split from AFAK, showing that afrotherian genomes are predominantly characterized by inter-chromosomal rearrangements (Figure S4).

To analyze the lineage-specific chromosomal reorganizations in the afrotherian and marsupial species, we identified genomic regions that were either collinear or had undergone reorganization from the most recent ancestor (Figure S4). Crucially, for this comparison, all Tasmanian devil chromosomes originated from individual MAUK chromosomes, with no inter-chromosomal rearrangements, and, consequently, more than 65% of the Tasmanian devil genome is fully collinear with that of the MAUK, with 33.58% disrupted due to inversions. The remaining 1.42% corresponded to unplaced sequences (Table S5).

In sharp contrast, the African elephant has retained only three chromosomes (LAF10, LAF23, and LAFX) that are fully collinear with the AFAK; these span 12.60% of the genome. Thus, the majority of African elephant chromosomes originated by either (1) fissions of larger AFAK chromosomes followed by fusions (27.37% of African elephant genome); (2) by fissions only (20.83% of the genome is collinear and 6.94% inverted); or (3) by a fusion of small AFAK chromosomes (24.25% is collinear and 3.54% inverted) (Table S5). These results highlight the different genome reshuffling patterns characterizing marsupials and afrotherian species.

Differences in lineage-specific chromosomal reorganizations were also observed in terms of gene content. Approximately 58% of the Tasmanian devil genes were located in collinear regions and 31.79% in inverted regions. Genes in inverted regions were related to signaling and response to stimuli, as well as major histocompatibility complex (MHC) and MHC class II protein complexes (Figure S5), whereas genes within collinear regions were enriched in Gene Ontology (GO) terms related to metabolic and developmental processes as well as anatomical (structural) development, among others (Figure S5). By contrast, only 10.87% of African elephant genes were located in collinear regions; 27.25% were in chromosomes originating from fission followed by fusion of AFAK chromosomes and 21.83% in chromosomes resulting from fissions only. Genes within collinear regions were enriched in GO terms related to nucleic acid transport. Genes within fission-fusion rearrangements were related to pheromone responses, whereas genes within fissioned regions were enriched in keratin filament and anatomical structure GO terms (Figure S5).

Functional and structural characterization of EBRs

The reconstruction of ancestral genomes allowed us to identify EBRs in both the Tasmanian devil and the African elephant (Figure S6). A total of 34 EBRs were identified within the Tasmanian devil genome, which correlated with chromosome size (R^2 = 0.93, $p < 0.001$; Figure S6A) and TAD boundaries (permutation test based on 10,000 permutations, normalized Z score > 0.01, $p < 0.05$; Figure S6C). Although EBRs were embedded in gene-dense regions (Figure S6E), they were negatively associated with gene position (permutation test based on 10,000 permutations, normalized Z score < -0.01, $p < 0.05$). The same trend was observed in the 45 EBRs identified in the African elephant (Figures S6D–S6F), mirroring previous studies. In both species, no GO term was associated with genes surrounding EBRs. African elephant EBRs, however, were enriched in transposable elements ($p = 0.04$, Z score = 1.82, 1,000 permutations). Interestingly, almost all African elephant EBRs contained afroSINEs and L1-LA elements, which are specific to afrotherians.

Lineage-specific chromosomal reorganizations affect genomic evolutionary plasticity

After ancestral genomes were reconstructed and lineage-specific chromosome reorganizations identified, we analyzed the structural and functional features of both conserved (collinear)

(C) Distribution of inter-/intra-chromosomal interaction ratio according to HSB length (in Mbp) for each class of reorganization in Tasmanian devil and African elephant: collinear, inverted, fissioned and collinear, fissioned and inverted, fused and collinear, fused and inverted, and fissioned and fused. Boxplots are presented as median values (center line); mean values (dot) ± SD. Asterisks represent statistically significant interaction ratio between HSBs (two-sided t test, ***$p < 0.001$).

(D) Boxplots depicting interactions between the same types of HSBs (*cis*) and different types of HSBs (*trans*) in Tasmanian devil and African elephant. Boxplots are presented as median values (center line); mean values (dot) ± SD. Asterisks represent statistically significant interactions between HSBs (two-sided t test, ***$p < 0.001$).

and rearranged genomic regions (inverted, fused, and fissioned) in the Tasmanian devil and the African elephant (Figure 5A).

We detected structural differences between collinear and rearranged regions in both species. As a general trend, lineage-specific inverted regions displayed higher distance-dependent contact probabilities than did collinear regions (Figure 5B). Differences in contact probability were higher in the Tasmanian devil, where blocks presented a mean contact probability value of 2.79 and inverted blocks a mean of 6.85 (2.5× fold increase). This was translated into different estimated slopes (as a proxy of DNA loop size). Inverted regions had a slope maximum of 6.5 Mbp, whereas the maximum value in conserved blocks was 7.5 Mbp, indicating that inverted blocks bear slightly shorter loops and, therefore, an increase in contact probability (Figure 5B).

In the African elephant, mean contact probability values were 14.2 for collinear regions and 20.38 for the inverted blocks (1.5× fold increase). The more complex chromosomal rearrangements between collinear and inverted HSBs (fissions followed by inversions and fusions followed by inversions) showed intermediate values. Finally, HSBs that were maintained as collinear, but either fissioned or fused in the African elephant, presented equally low contact probability values (mean contacts 3.82, a 3.5× fold reduction compared with collinear regions). However, these contacts were never as low as those observed in the Tasmanian devil conserved blocks. A similar tendency was detected in the slopes estimated for each type of reorganization, with a maximum at 2 Mbp for inverted blocks, 4.5 Mbp for collinear blocks, and 8 Mbp for translocated (fused/fissioned) blocks (Figure 5B).

Differences between types of rearrangements were also noted when analyzing inter-/intra-chromosomal interaction ratios (Figures 5C and 5D). In both species, lineage-specific inverted blocks presented the lowest inter-/intra-chromosomal interactions, being lower in African elephant (mean 0.05; two-sided t test, $p < 0.01$) than in Tasmanian devil (mean 0.6; two-sided t test, $p < 0.01$). African elephant blocks that were fissioned and subsequently fused displayed the highest inter-/intra-chromosomal interaction (mean 0.35; two-sided, t test, $p < 0.01$) (Figure 5C). Collectively, these observations point to the presence of distinctive genomic architectural features in lineage-specific chromosomal reorganizations. This was especially relevant for inversions, which presented more intra-chromosomal interactions than surrounding collinear regions (Figure 6).

Remarkably, structural reshuffling did not result in gross changes in TADs and compartments (Figure 3D). This was in line with the fact that EBRs were positively associated with TAD boundaries (multiple permutation test based on 10,000 permutations, normalized Z score > 0.01, $p < 0.05$) in both species (Figure S6). Likewise, EBRs were devoid of genes (multiple permutation test based on 10,000 permutations, normalized Z score > 0.01, $p < 0.05$; Figure S6), suggesting that gross genome reshuffling is less likely to disrupt gene regulation.

DISCUSSION

A major challenge in genome research is to determine why some species have stable genomes whereas others have undergone substantial rearrangement. Here, we describe the fundamental principles of 3D chromosome folding in mammals and show that lineage-specific evolutionary genomic reshuffling can influence patterns of higher-order chromatin organization.

Our data provide evidence for the existence of different chromosome folding patterns within mammals. In the eutherian mammals (Boreoeutheria and Afrotheria) analyzed herein, the chromosomes (irrespective of diploid numbers) were organized into CTs during interphase, mirroring observations in other species.[21] Low inter-/intra-chromosomal interaction ratios observed in human, mouse, African elephant, and aardvark were indicative of relatively highly compacted chromatin as reflected by high distance-dependent interactions and high CTCF density.

In contrast, marsupials (tammar wallaby and Tasmanian devil) showed a distinctive Rabl-like chromosomal distribution with centromeres and heterochromatic regions clustering near the center of the nucleus. This chromosome distribution was accompanied by (1) high inter-/intra-chromosomal interaction ratios and (2) low distance-dependent interactions. These observations, together with the detection of low CTCF genomic density in the opossum (an American marsupial representative),[30] indicate that long marsupial chromosomes (average size 400 Mbp) form a "loose" distribution that extends across the nucleus, with presumably lower numbers of longer loops anchored by their centromeres that are orientated toward the center of the cell. This probably has implications for the position of chromosomes within the nucleus. In this context, our genomic approach supports initial cytogenetics studies reporting a radial configuration of marsupial chromosome inside nuclei.[36]

Previous studies have shown the presence of a Rabl-like configuration in yeast (centromeres forming one large focus in the vicinity of the spindle pole body), wheat (centromeres clusters at one pole), mosquitos, and sea urchins.[21,37] To these species, we can add marsupials—an ancient mammalian clade with genome plasticity and chromosomal diploid number variation that is distinctive from eutherian mammals.[38–40] We can only speculate on the mechanisms responsible of the Rabl-like configuration in marsupials as the function of this chromosome pattern still remains a mystery since its initial first cytological description.[41] Although speculative, it is possible that the Rabl configuration is a relic from anaphase-segregating chromosomes.[42] In this case, larger/longer chromosomes would establish a polarized pattern more readily during interphase with size and the heterochromatin distribution favoring a Rabl configuration.[43] Further research is needed to fully test these hypotheses.

Importantly, our analysis of the ancestral genomic reconstructions showed contrasting patterns of genome reshuffling in marsupials and afrotherians. The Tasmanian devil has retained the same chromosome number as has been proposed for both the MAK and the MAUK (n = 7). Moreover, it has few species-specific inversions. In contrast, the African elephant karyotype has been extensively reorganized (inter- and intra-chromosomally) compared with the AFAK (n = 24), which, in turn, largely resembles the eutherian ancestral karyotypic configuration (n = 23[38]). Based on the homologies shared with the human genome (Figure S3), marsupials have undergone widespread genomic reshuffling after their split from the therian common

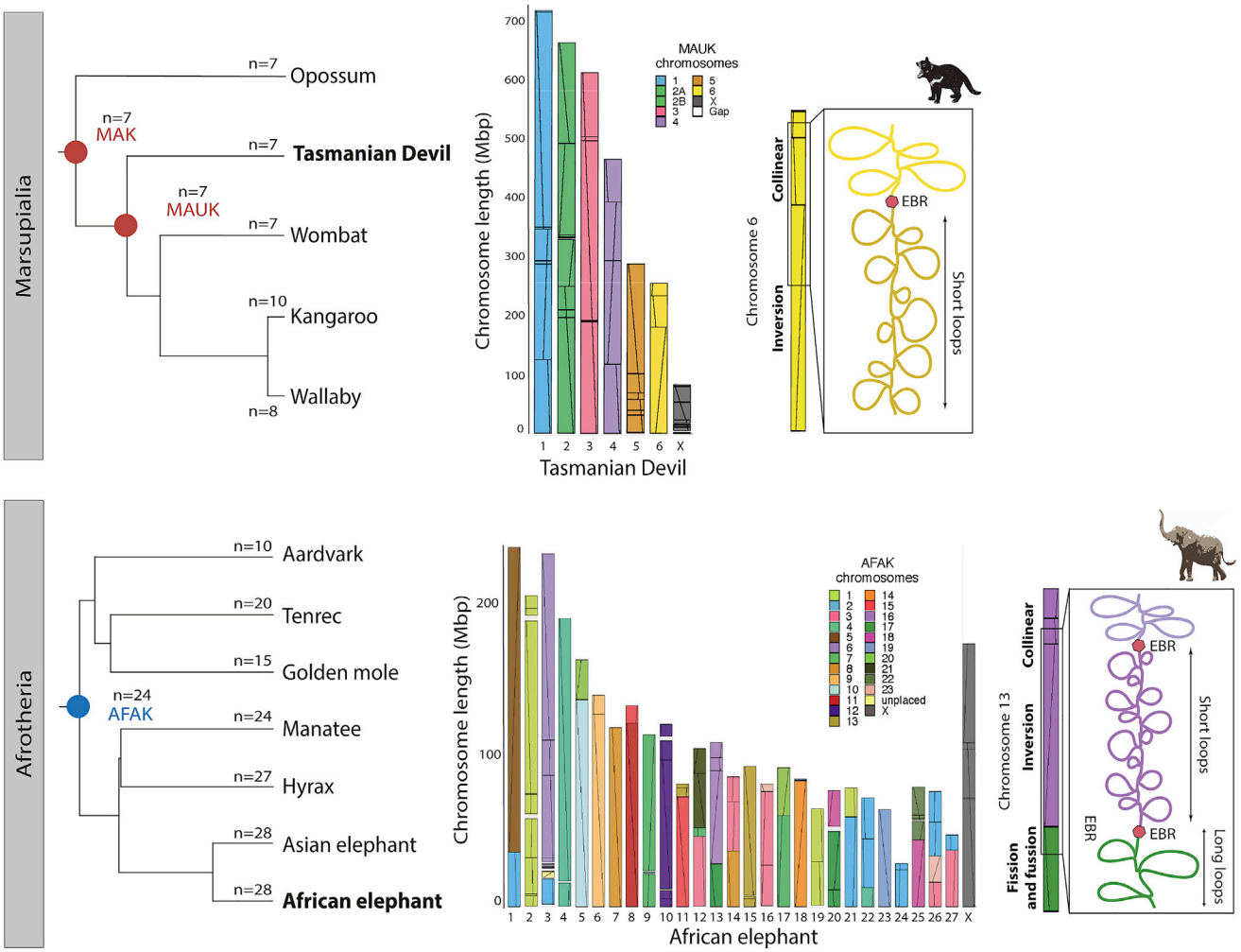


Figure 6. Model on the influence of evolutionary genome reshuffling in chromatin folding

Representation of genomic architectural features detected in lineage-specific chromosomal reorganizations in the Tasmanian devil and African elephant. Left panel: phylogenetic relationship among the afrotherian and marsupial species compared, including the haploid number of chromosomes (n) for each species. Central panel: Tasmanian devil and African elephant chromosome ideograms color coded accordingly to their corresponding MAUK or AFAK chromosomes. Right panel: representation of DNA loops in inverted and fused HSBs in both species, chromosome 6 in the Tasmanian devil, and chromosome 13 in the African elephant. Note that inverted regions have shorter DNA loops than non-inverted regions (either collinear, fissioned, and fusioned). MAK, marsupial ancestral karyotype; MAUK, Australian marsupial ancestral karyotype; AFAK, Afrotheria ancestral karyotype; HSB, homologous synteny blocks; EBRs, evolutionary breakpoint regions.

ancestor. This was later stabilized as reflected by the conserved karyotypes described within the group.[24] In sharp contrast, in afrotherian species (African elephant and aardvark), and probably those boreoeutherians with highly diverse karyotypes (such as n = 3 in the female Indian muntjac and n = 51 in the red viscacha rat[38]), genome evolution involved more complex reorganization (inversions, fusions, and fissions).

Based on the structural plasticity detected in lineage-specific chromosomal reorganizations, we hypothesize that chromosome folding patterns have influence on chromosomal evolution and genome reshuffling in different ways in different phylogroups. In marsupials, the development of centromeric associations probably predated the MAUK radiation and most likely influenced the high occurrence of inversions in the Tasmanian devil and tammar wallaby when compared with the MAUK. As centromeres are anchored toward the center of the nucleus in marsupials, the resulting chromosomal distribution could impose structural constrains that favor intra-chromosomal reorganization rather than inter-chromosomal rearrangements. This is supported by data that suggest centromeres can act as strong topological barriers that prevent contact between the two chromosome arms.[37,44] Should this hold, we can infer that the chromosomal reduction in marsupials, after their split from the therian ancestor, had already resulted in an ancestral configuration of centromere associations that is now reflected in all marsupials.

Likewise, the pattern of chromosome reshuffling observed in modern Afrotheria (that show extensive intra- and inter-chromosomal reorganizations from the AFAK) can be related to

chromosomes being organized into highly compacted CTs, where heterologous chromosomes are more frequently in contact. This would favor inter-chromosomal, rather than intra-chromosomal, reorganization. Importantly, chromatin intermingles at the peripheral regions of CTs in mammals,[45] allowing interactions between non-homologous chromosomes. This can facilitate the rejoining of broken DNA ends of heterologous chromosomes[17] and can explain the excess of inter-chromosomal rearrangements detected.

Equally unexpected was the observation that inversions (irrespective of clade) resulted in different DNA loop sizes and distance-dependent interaction contact frequencies when compared with collinear genomic regions (Figure 6). This was observed for inversions in the Tasmanian devil and African elephant, where chromatin was packaged differently from neighboring, non-reorganized regions on the same chromosome. Inverted regions showed (1) differences in slope (an estimate of DNA loop size), (2) high interactions at short distances, and (3) high intra-chromosomal interactions. Studies in different taxa (i.e., butterflies, mosquitos, pea aphids, fishes, and plants[46–50]) have revealed clusters of differentiated loci (the so-called "genomic islands of divergence"[49,51,52]) often involved in inversions between lineages. In light of our data, we suggest that lineage-specific inversions may also act as "structural genomic islands" by imposing structural constrains (i.e., acting as barriers for genomic contacts) with surrounding regions. In contrast, at higher hierarchical levels (i.e., compartments and TADs), the 3D genome structure of collinear regions has the same level of structural conservation between species as do the inverted regions, thereby extending previous observation in the carnivores[53] to more basal mammals.

It is tempting to speculate that lineage-specific inversions have resulted in new structural features that are distinct from other genomic regions, which have been conserved over tens of millions of years. We suggest that these divergent structural features can result in topological (and hence genetic) barriers that may, at least potentially, have functional implications. Inversions could isolate genes from surrounding regions as they show shorter loops (and hence reduced contacts). This view is consistent with recent evidence provided by divergent expression profiles in inverted regions within Cetartiodactyla.[14] Such topological barriers could be transmitted through the germ line. In fact, it was recently shown that chromosome fusions alter the nuclear architecture in mouse germ cells.[16] This included an increased rate of heterologous interactions that, in turn, alter chromosome axis length and DNA loop size. Importantly, these disturbances in chromosome topology were associated with changes in the recombination landscape, resulting in detectable genomic footprints at a population level.

In conclusion, our study provided an evolutionary view of the 3D genome folding patterns in distantly related mammals. Through an integrative computational analysis of a comprehensive Hi-C dataset and the use of comparative genomics, it was possible to infer the ancestral karyotypic structure of both marsupial and afrotherian genomes. This permitted the reconstruction of lineage-specific chromosome reorganization that captures the deepest divergences of mammals. Crucially, our results suggest that 3D chromosome folding influences the patterns of genome reshuffling that are transmitted to offspring, an observation supported by a recent survey in rodents.[4]

Limitations of the study

As the use of non-model species can be challenging, future studies with a larger sample sizes and greater species representation should be considered a priority. Moreover, understanding the dynamics of chromatin conformation during development in other distantly related species will be fundamental to deciphering the structural plasticity of vertebrate genomes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines
- METHOD DETAILS
  - Primary fibroblast cell lines
  - Immunoflourescence and microscopy
  - *In nuclei* Hi-C
  - African elephant Hi-C assisted assembly
  - Hi-C data processing, binning and normalisation
  - Averaged contact probability P(s) and its derivative
  - Inter-chromosome/intra-chromosome interaction ratio
  - Centromere interaction quantification
  - Centromere aggregate contacts plots
  - First eigenvector and insulator score calculation
  - TAD calling and TAD boundaries
  - ChIP-sequencing
  - ChIP-seq peak calling and annotation
  - African elephant genome annotation
  - RNA-seq analysis
  - Whole-genome pairwise alignments
  - Conservation of the higher-order chromatin organization between species
  - Ancestral karyotype reconstructions
  - Detection of evolutionary breakpoint regions (EBRs)
  - Gene ontology enrichment analysis (GOEA)
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Multi-association and statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.111839.

ACKNOWLEDGMENTS

We thank the specialist and High Performance Computing systems provided by Information Services at the University of Kent. Unpublished genome

assemblies for aardvark (*Orycteropus afer*); African elephant (*Loxodonta africana*); Asiatic elephant (*Elephas maximus*); Cape rock hyrax (*Procavia capensis*); gray short-tailed opossum (*Monodelphis domestica*); red kangaroo (*Macropus rufus*/*Osphranter rufus*); West Indian manatee (*Trichechus manatus*); and wombat (*Vombatus ursinus*) are used with permission from the DNA Zoo Consortium (dnazoo.org). This work was supported by the Ministry of Economy, Industry and Competitiveness (CGL2017-83802-P to A.R.-H.) and the Spanish Ministry of Science and Innovation (PID2020-112557GB-I00 to A.R.-H. and PID2020-115696RB-I00 to M.A.M.-R.). Research funding to P.D.W. (Australian Research Council grants DP180100931, DP210103512, and DP220101429) and T.J.R. (South African National Research Foundation) are gratefully acknowledged. C.V. and L.A.-G. were supported by FPI predoctoral fellowships from the Ministry of Economy and Competitiveness (BES-2015-072924 and PRE-2018-083257). L.M.-G. was supported by an FPU predoctoral fellowship from the Spanish Ministry of Science, Innovation and University (FPU18/03867). C.A.-S. was supported by a GTA fellowship from the University of Kent.

AUTHOR CONTRIBUTIONS

A.R.-H. conceived and devised the study. L.A.-G., C.A.-S., P.D.W., M.M.-R., M.F., and A.R.-H. designed experiments and analysis. L.A.-G., L.M.-G., C.V., Y.C., F.G., and A.R.-H. performed experiments. L.A.-G., C.A.-S., L.M.-E., L.M.-G., M.F., and A.R.-H. analyzed the data. J.D., M.B.R., T.J.R., M.A.M.-R., and A.R.-H. contributed reagents and to data collection. L.A.-G., C.A.-S., P.D.W., M.F., and A.R.-H. wrote the initial first draft of the manuscript with input from all authors. All authors read and approved the final version of the manuscript.

DECLARATION OF INTERESTS

M.A.M.-R. serves as a consultant to Acuity Spatial Genomics, Inc., and receives compensation for these services.

Received: June 18, 2022
Revised: October 1, 2022
Accepted: November 24, 2022
Published: December 20, 2022

REFERENCES

1. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021.
2. Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., and Hadjur, S. (2015). Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. Cell Rep. *10*, 1297–1309. https://doi.org/10.1016/j.celrep.2015.02.004.
3. Vara, C., Paytuví-Gallart, A., Cuartero, Y., le Dily, F., Garcia, F., Salvà-Castro, J., Gómez-H, L., Julià, E., Moutinho, C., Aiese Cigliano, R., et al. (2019). Three-dimensional genomic structure and cohesin occupancy correlate with transcriptional activity during spermatogenesis. Cell Rep. *28*, 352–367.e9. https://doi.org/10.1016/j.celrep.2019.06.037.
4. Álvarez-González, L., Burden, F., Doddamani, D., Malinverni, R., Leach, E., Marín-García, C., Marín-Gual, L., Gubern, A., Vara, C., Paytuví-Gallart, A., et al. (2022). 3D chromatin remodelling in the germ line modulates genome evolutionary plasticity. Nat. Commun. *13*, 2608. https://doi.org/10.1038/s41467-022-30296-6.
5. Ruiz-Herrera, A., Castresana, J., and Robinson, T.J. (2006). Is mammalian chromosomal evolution driven by regions of genome fragility? Genome Biol. *7*, R115. https://doi.org/10.1186/gb-2006-7-12-r115.
6. Larkin, D.M., Pape, G., Donthu, R., Auvil, L., Welge, M., and Lewin, H.A. (2009). Breakpoint regions and homologous synteny blocks in chromosomes have different evolutionary histories. Genome Res. *19*, 770–777. https://doi.org/10.1101/gr.086546.108.
7. Groenen, M.A.M., Archibald, A.L., Uenishi, H., Tuggle, C.K., Takeuchi, Y., Rothschild, M.F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.J., et al. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. Nature *491*, 393–398. https://doi.org/10.1038/nature11622.
8. Ullastres, A., Farré, M., Capilla, L., and Ruiz-Herrera, A. (2014). Unraveling the effect of genomic structural changes in the rhesus macaque - implications for the adaptive role of inversions. BMC Genom. *15*, 530–613. https://doi.org/10.1186/1471-2164-15-530.
9. Farré, M., Narayan, J., Slavov, G.T., Damas, J., Auvil, L., Li, C., Jarvis, E.D., Burt, D.W., Griffin, D.K., and Larkin, D.M. (2016). Novel insights into chromosome evolution in birds, archosaurs, and reptiles. Genome Biol. Evol. *8*, 2442–2451. https://doi.org/10.1093/gbe/evw166.
10. Capilla, L., Sánchez-Guillén, R.A., Farré, M., Paytuví-Gallart, A., Malinverni, R., Ventura, J., Larkin, D.M., and Ruiz-Herrera, A. (2016). Mammalian comparative genomics reveals genetic and epigenetic features associated with genome reshuffling in rodentia. Genome Biol. Evol. *8*, 3703–3717. https://doi.org/10.1093/gbe/evw276.
11. Longo, M.S., Carone, D.M., NISC Comparative Sequencing Program; Green, E.D., O'Neill, M.J., and O'Neill, R.J. (2009). Distinct retroelement classes define evolutionary breakpoints demarcating sites of evolutionary novelty. BMC Genom. *10*, 334–414. https://doi.org/10.1186/1471-2164-10-334.
12. Farré, M., Bosch, M., López-Giráldez, F., Ponsà, M., and Ruiz-Herrera, A. (2011). Assessing the role of tandem repeats in shaping the genomic architecture of great apes. PLoS One *6*, e27239. https://doi.org/10.1371/journal.pone.0027239.
13. Robinson, T.J., Cernohorska, H., Kubickova, S., Vozdova, M., Musilova, P., and Ruiz-Herrera, A. (2021). Chromosomal evolution in Raphicerus antelope suggests divergent X chromosomes may drive speciation through females, rather than males, contrary to Haldane's rule. Sci. Rep. *11*, 3152. https://doi.org/10.1038/s41598-021-82859-0.
14. Farré, M., Kim, J., Proskuryakova, A.A., Zhang, Y., Kulemzina, A.I., Li, Q., Zhou, Y., Xiong, Y., Johnson, J.L., Perelman, P.L., et al. (2019). Evolution of gene regulation in ruminants differs between evolutionary breakpoint regions and homologous synteny blocks. Genome Res. *29*, 576–589. https://doi.org/10.1101/gr.239863.118.
15. Vara, C., and Ruiz-Herrera, A. (2022). Unpacking chromatin remodelling in germ cells: implications for development and evolution. Trends Genet. *38*, 422–425. https://doi.org/10.1016/j.tig.2021.10.007.
16. Vara, C., Paytuví-Gallart, A., Cuartero, Y., Álvarez-González, L., Marín-Gual, L., Garcia, F., Florit-Sabater, B., Capilla, L., Sanchéz-Guillén, R.A., Sarrate, Z., et al. (2021). The impact of chromosomal fusions on 3D genome folding and recombination in the germ line. Nat. Commun. *12*, 2981–3017. https://doi.org/10.1038/s41467-021-23270-1.
17. Farré, M., Robinson, T.J., and Ruiz-Herrera, A. (2015). An Integrative Breakage Model of genome architecture, reshuffling and evolution: the Integrative Breakage Model of genome evolution, a novel multidisciplinary hypothesis for the study of genome plasticity. Bioessays *37*, 479–488. https://doi.org/10.1002/bies.201400174.
18. Deakin, J.E., Potter, S., O'Neill, R., Ruiz-Herrera, A., Cioffi, M.B., Eldridge, M.D.B., Fukui, K., Marshall Graves, J.A., Griffin, D., Grutzner, F., et al. (2019). Chromosomics: bridging the gap between genomes and chromosomes. Genes *10*, 627. https://doi.org/10.3390/genes10080627.
19. Dekker, J., Marti-Renom, M.A., and Mirny, L.A. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat. Rev. Genet. *14*, 390–403. https://doi.org/10.1038/nrg3454.
20. Phillips-cremins, J.E. (2014). Unraveling architecture of the pluripotent genome. Curr. Opin. Cell Biol. *28*, 96–104. https://doi.org/10.1016/j.ceb.2014.04.006.

21. Hoencamp, C., Dudchenko, O., Elbatsh, A.M.O., Brahmachari, S., Raaijmakers, J.A., van Schaik, T., Sedeño Cacciatore, Á., Contessoto, V.G., van Heesbeen, R.G.H.P., van den Broek, B., et al. (2021). 3D genomics across the tree of life reveals condensin II as a determinant of architecture type. Science *372*, 984–989. https://doi.org/10.1126/science.abe2218.

22. Meredith, R.W., Janečka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C., Goodbla, A., Eizirik, E., Simão, T.L.L., Stadler, T., et al. (2011). Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science *334*, 521–524. https://doi.org/10.1126/science.1211028.

23. Grützner, F., and Graves, J.A.M. (2004). A platypus' eye view of the mammalian genome. Curr. Opin. Genet. Dev. *14*, 642–649. https://doi.org/10.1016/j.gde.2004.09.006.

24. Deakin, J.E., and O'Neill, R.J. (2020). Evolution of marsupial genomes. Annu. Rev. Anim. Biosci. *8*, 25–45. https://doi.org/10.1146/annurev-animal-021419-083555.

25. Robinson, T.J., and Seiffert, E.R. (2004). Afrotherian origins and interrelationships: new views and future prospects. Curr. Top. Dev. Biol. *63*, 37–60. https://doi.org/10.1016/S0070-2153(04)63002-X.

26. Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., and Aiden, E.L. (2017). De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science *356*, 92–95. https://doi.org/10.1126/science.aal3327.

27. Dudchenko, O., Shamim, M.S., Batra, S.S., Durand, N.C., Musial, N.T., Mostofa, R., Pham, M., Glenn St Hilaire, B., Yao, W., Stamenova, E., et al. (2018). The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000. Preprint at bioRxiv, 254797. https://doi.org/10.1101/254797.

28. Zhou, Y., Shearwin-Whyatt, L., Li, J., Song, Z., Hayakawa, T., Stevens, D., Fenelon, J.C., Peel, E., Cheng, Y., Pajpach, F., et al. (2021). Platypus and echidna genomes reveal mammalian biology and evolution. Nature *592*, 756–762. https://doi.org/10.1038/s41586-020-03039-0.

29. Fishman, V., Battulin, N., Nuriddinov, M., Maslova, A., Zlotina, A., Strunov, A., Chervyakova, D., Korablev, A., Serov, O., and Krasikova, A. (2019). 3D organization of chicken genome demonstrates evolutionary conservation of topologically associated domains and highlights unique architecture of erythrocytes' chromatin. Nucleic Acids Res. *47*, 648–665. https://doi.org/10.1093/nar/gky1103.

30. Schmidt, D., Schwalie, P.C., Wilson, M.D., Ballester, B., Gonalves, Â., Kutter, C., Brown, G.D., Marshall, A., Flicek, P., and Odom, D.T. (2012). Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. Cell *148*, 335–348. https://doi.org/10.1016/j.cell.2011.11.058.

31. Mandrioli, M., Bandinelli, S., and Manicardi, G.C. (2014). Occurrence of Rabl-like telomere clustering in the holocentric chromosomes of the peach potato aphid *Myzus persicae* (Hemiptera; Aphididae). Cytogenet. Genome Res. *144*, 68–75. https://doi.org/10.1159/000366049.

32. Matsushita, M., Ochiai, H., Suzuki, K.I.T., Hayashi, S., Yamamoto, T., Awazu, A., and Sakamoto, N. (2017). Dynamic changes in the interchromosomal interaction of early histone gene loci during early development of sea urchin. J. Cell Sci. *130*, 4097–4107. https://doi.org/10.1242/jcs.206862.

33. Pouokam, M., Cruz, B., Burgess, S., Segal, M.R., Vazquez, M., and Arsuaga, J. (2019). The Rabl configuration limits topological entanglement of chromosomes in budding yeast. Sci. Rep. *9*, 6795. https://doi.org/10.1038/s41598-019-42967-4.

34. Serra, F., Baù, D., Goodstadt, M., Castillo, D., Filion, G.J., and Marti-Renom, M.A. (2017). Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors. PLoS Comput. Biol. *13*, e1005665. https://doi.org/10.1371/journal.pcbi.1005665.

35. Frönicke, L., Wienberg, J., Stone, G., Adams, L., and Stanyon, R. (2003). Towards the delineation of the ancestral eutherian genome organization: comparative genome maps of human and the African elephant (*Loxodonta africana*) generated by chromosome painting. Proc. R. Soc. Lond. B. *270*, 2639. https://doi.org/10.1098/rspb.2003.2002.

36. Rens, W., O'Brien, P.C.M., Fairclough, H., Harman, L., Graves, J.A.M., and Ferguson-Smith, M.A. (2003). Reversal and convergence in marsupial chromosome evolution. Cytogenet. Genome Res. *102*, 282–290. https://doi.org/10.1159/000075764.

37. Muller, H., Gil, J., and Drinnenberg, I.A. (2019). The impact of centromeres on spatial genome architecture. Trends Genet. *35*, 565–578. https://doi.org/10.1016/j.tig.2019.05.003.

38. Ruiz-Herrera, A., Farré, M., and Robinson, T.J. (2012). Molecular cytogenetic and genomic insights into chromosomal evolution. Heredity *108*, 28–36. https://doi.org/10.1038/hdy.2011.102.

39. Waters, P.D., Patel, H.R., Ruiz-Herrera, A., Álvarez-González, L., Lister, N.C., Simakov, O., Ezaz, T., Kaur, P., Frere, C., Grützner, F., et al. (2021). Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. Proc. Natl. Acad. Sci. USA *118*, e2112494118. https://doi.org/10.1073/pnas.2112494118.

40. Marín-Gual, L., González-Rodelas, L., Pujol, G., Vara, C., Martín-Ruiz, M., Berríos, S., Fernández-Donoso, R., Pask, A., Renfree, M.B., Page, J., et al. (2022). Strategies for meiotic sex chromosome dynamics and telomeric elongation in Marsupials. PLoS Genet. *18*, e1010040. https://doi.org/10.1371/journal.pgen.1010040.

41. Rabl, C. (1885). Rabl. Über Zelltheilung. Morphologisches Jahrbuch. Gegenbaur C., 10, pp. 214–330.

42. Dong, F., and Jiang, J. (1998). Non-Rabl patterns of centromere and telomere distribution in the interphase nuclei of plant cells. Chromosome Res. *6*, 551–558. https://doi.org/10.1023/A:1009280425125.

43. Idziak, D., Robaszkiewicz, E., and Hasterok, R. (2015). Spatial distribution of centromeres and telomeres at interphase varies among *Brachypodium* species. J. Exp. Bot. *66*, 6623–6634. https://doi.org/10.1093/jxb/erv369.

44. Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458–472. https://doi.org/10.1016/j.cell.2012.01.010.

45. Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol. *4*, e138. https://doi.org/10.1371/journal.pbio.0040138.

46. Marques, D.A., Lucek, K., Meier, J.I., Mwaiko, S., Wagner, C.E., Excoffier, L., and Seehausen, O. (2016). Genomics of rapid incipient speciation in sympatric threespine stickleback. PLoS Genet. *12*, e1005887. https://doi.org/10.1371/journal.pgen.1005887.

47. Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., and Jiggins, C.D. (2013). Genome-wide evidence for speciation with gene flow in Heliconius butterflies. Genome Res. *23*, 1817–1828. https://doi.org/10.1101/gr.159426.113.

48. Renaut, S., Grassa, C.J., Yeaman, S., Moyers, B.T., Lai, Z., Kane, N.C., Bowers, J.E., Burke, J.M., and Rieseberg, L.H. (2013). Genomic islands of divergence are not affected by geography of speciation in sunflowers. Nat. Commun. *4*, 1827. https://doi.org/10.1038/ncomms2833.

49. Turner, T.L., Hahn, M.W., and Nuzhdin, S.V. (2005). Genomic islands of speciation in *Anopheles gambiae*. PLoS Biol. *3*, e285. https://doi.org/10.1371/journal.pbio.0030285.

50. Via, S., and West, J. (2008). The genetic mosaic suggests a new role for hitchhiking in ecological speciation. Mol. Ecol. *17*, 4334–4345. https://doi.org/10.1111/j.1365-294X.2008.03921.x.

51. Hejase, H.A., Salman-Minkov, A., Campagna, L., Hubisz, M.J., Lovette, I.J., Gronau, I., and Siepel, A. (2020). Genomic islands of differentiation in a rapid avian radiation have been driven by recent selective sweeps.

Proc. Natl. Acad. Sci. USA *117*, 30554–30565. https://doi.org/10.1073/pnas.2015987117.

52. Wolf, J.B.W., and Ellegren, H. (2017). Making sense of genomic islands of differentiation in light of speciation. Nat. Rev. Genet. *18*, 87–100. https://doi.org/10.1038/nrg.2016.133.
53. Corbo, M., Damas, J., Bursell, M.G., and Lewin, H.A. (2022). Conservation of chromatin conformation in carnivores. Proc. Natl. Acad. Sci. USA *119*, e2120555119. https://doi.org/10.1073/pnas.2120555119.
54. Cortez, D., Marin, R., Toledo-Flores, D., Froidevaux, L., Liechti, A., Waters, P.D., Grützner, F., and Kaessmann, H. (2014). Origins and functional evolution of Y chromosomes across mammals. Nature *508*, 488–493. https://doi.org/10.1038/nature13151.
55. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.
56. Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner (Department of Energy. Joint Genome Institute). https://www.osti.gov/servlets/purl/1241166.
57. Marco-Sola, S., Sammeth, M., Guigó, R., and Ribeca, P. (2012). The GEM mapper: fast, accurate and versatile alignment by filtration. Nat. Methods *9*, 1185–1188. https://doi.org/10.1038/nmeth.2221.
58. Wolff, J., Bhardwaj, V., Nothjunge, S., Richard, G., Renschler, G., Gilsbach, R., Manke, T., Backofen, R., Ramírez, F., and Grüning, B.A. (2018). Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. Nucleic Acids Res. *46*, W11–W16. https://doi.org/10.1093/nar/gky504.
59. Venev, S., Abdennur, N., Goloborodko, A., Flyamer, I., Fudenberg, G., Nuebler, J., Galitsyna, A., Akgol, B., Abraham, S., and Kerpedjiev, P. (2021). open2c/cooltools: v0.4.0.
60. Schwarzer, G. (2007). R Core Team R: A Language and Environment for Statistical Computing (R News), pp. 40–45.
61. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.
62. Kruse, K., Hug, C.B., and Vaquerizas, J.M. (2020). FAN-C: a feature-rich framework for the analysis and visualisation of chromosome conformation capture data. Genome Biol. *21*, 303. https://doi.org/10.1186/s13059-020-02215-9.
63. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. [Online]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
64. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
65. Li, H. (2013). Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. Preprint at arXiv, 1303.3997. https://doi.org/10.48550/arXiv.1303.3997.
66. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.
67. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.
68. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.
69. Alhendi, A.S.N. (2019). CountToFPKM: Convert Counts to Fragments Per Kilobase of Transcript Per Million (FPKM). R Package Version 1 .0.0. .
70. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics *26*, 2204–2207. https://doi.org/10.1093/bioinformatics/btq351.
71. Tempel, S. (2012). Using and Understanding RepeatMasker. Methods Mol. Biol. *859*, 29–51. https://doi.org/10.1007/978-1-61779-603-6_2.
72. Harris, R.S. (2007). Improved Pairwise Alignment of Genomic DNA. Ph.D. Thesis (The Pennsylvania State University).
73. Kolmogorov, M., Armstrong, J., Raney, B.J., Streeter, I., Dunn, M., Yang, F., Odom, D., Flicek, P., Keane, T.M., Thybert, D., et al. (2018). Chromosome assembly of large and complex genomes using multiple references. Genome Res. *28*, 1720–1732. https://doi.org/10.1101/gr.236273.118.
74. Kim, J., Farré, M., Auvil, L., Capitanu, B., Larkin, D.M., Ma, J., and Lewin, H.A. (2017). Reconstruction and evolutionary history of eutherian chromosomes. Proc. Natl. Acad. Sci. USA *114*, e5379–e5388. https://doi.org/10.1073/pnas.1702012114.
75. Kumar, S., Stecher, G., Suleski, M., and Hedges, S.B. (2017). TimeTree: a resource for timelines, timetrees, and divergence times. Mol. Biol. Evol. *34*, 1812–1819. https://doi.org/10.1093/molbev/msx116.
76. Rambaut, A. (2009). Figtree, A Graphical Viewer Of Phylogenetic Trees. http://tree.bio.ed.ac.uk/software/figtree/.
77. Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., and Sherlock, G. (2004). GO::TermFinder–open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics *20*, 3710–3715. https://doi.org/10.1093/bioinformatics/bth456.
78. Gel, B., Díez-Villanueva, A., Serra, E., Buschbeck, M., Peinado, M.A., and Malinverni, R. (2016). RegioneR: An R/Bioconductor package for the association analysis of genomic regions based on permutation tests. Bioinformatics *32*, 289–291. https://doi.org/10.1093/bioinformatics/btv562.
79. Hickford, D., Frankenberg, S., and Renfree, M.B. (2009). The Tammar Wallaby, *Macropus eugenii*: a model kangaroo for the study of developmental and reproductive biology. Cold Spring Harb. Protoc., emo137. https://doi.org/10.1101/pdb.emo137.
80. National Health and Medical Research Council (2013). Australian Code of Practice for the Care and Use of Animals for Scientific Purposes8 (National Health and Medical Research Council).
81. National Health and Medical Research Council (2014). A Guide to the Care and Use of Australian Native Mammals in Research and Teaching (National Health and Medical Research Council).
82. Patel, L., Kang, R., Rosenberg, S.C., Qiu, Y., Raviram, R., Chee, S., Hu, R., Ren, B., Cole, F., and Corbett, K.D. (2019). Dynamic reorganization of the genome shapes the recombination landscape in meiotic prophase. Nat. Struct. Mol. Biol. *26*, 164–174. https://doi.org/10.1038/s41594-019-0187-0.
83. Costantino, L., Hsieh, T.-H.S., Lamothe, R., Darzacq, X., and Koshland, D. (2020). Cohesin residency determines chromatin loop patterns. Elife *9*, e59889. https://doi.org/10.7554/eLife.59889.
84. Yang, F., Alkalaeva, E.Z., Perelman, P.L., Pardini, A.T., Harrison, W.R., O'Brien, P.C.M., Fu, B., Graphodatsky, A.S., Ferguson-Smith, M.A., and Robinson, T.J. (2003). Reciprocal chromosome painting among human, aardvark, and elephant (superorder Afrotheria) reveals the likely eutherian ancestral karyotype. Proc. Natl. Acad. Sci. USA *100*, 1062–1066. https://doi.org/10.1073/pnas.0335540100.
85. Taylor, R.L., Zhang, Y., Schöning, J.P., and Deakin, J.E. (2017). Identification of candidate genes for devil facial tumour disease tumourigenesis. Sci. Rep. *7*, 8761. https://doi.org/10.1038/s41598-017-08908-9.
86. Alsop, A.E., Miethke, P., Rofe, R., Koina, E., Sankovic, N., Deakin, J.E., Haines, H., Rapkins, R.W., and Marshall Graves, J.A. (2005). Characterizing the chromosomes of the Australian model marsupial *Macropus eugenii* (tammar wallaby). Chromosome Res. *13*, 627–636. https://doi.org/10.1007/s10577-005-0989-2.

87. Deakin, J.E., Bender, H.S., Pearse, A.M., Rens, W., O'Brien, P.C.M., Ferguson-Smith, M.A., Cheng, Y., Morris, K., Taylor, R., Stuart, A., et al. (2012). Genomic restructuring in the tasmanian devil facial tumour: chromosome painting and gene mapping provide clues to evolution of a transmissible tumour. PLoS Genet. *8*, e1002483. https://doi.org/10.1371/journal.pgen.1002483.

88. Keilwagen, J., Hartung, F., and Grau, J. (2019). GeMoMa: homology-based gene prediction utilizing intron position conservation and RNA-seq data. Methods in Mol. Biol. *1967*, 161–177. https://doi.org/10.1007/978-1-4939-9173-0_9.

89. Kent, W.J., Baertsch, R., Hinrichs, A., Miller, W., and Haussler, D. (2003). Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. Proc. Natl. Acad. Sci. USA *100*, 11484–11489. https://doi.org/10.1073/pnas.1932072100.

90. Smedley, D., Jacobsen, J.O.B., Jäger, M., Köhler, S., Holtgrewe, M., Schubach, M., Siragusa, E., Zemojtel, T., Buske, O.J., Washington, N.L., et al. (2015). Next-generation diagnostics and disease-gene discovery with the Exomiser. Nat. Protoc. *10*, 2004–2015. https://doi.org/10.1038/nprot.2015.124.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT orRESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-rabbit H3K9me3	Abcam	#ab8898; RRID:AB_306848
Anti-human CREST	provided by M. Fritzler	N/A
Anti-rabbit Cy3	Jackson Immunoresearch	#115-165-003; RRID:AB_2338680
Anti-mouse FITC	Jackson Immunoresearch	#115-095-003; RRID:AB_2338589
Anti-rabbit CTCF	Millipore	#07-729; RRID:AB_441965
Anti-rabbit H3K4me3	Abcam	#ab8580; RRID:AB_306649
Chemicals, peptides, and recombinant proteins		
AmnioMAX™	ThermoFisher Scientific	#11269016
Gentamycin	ThermoFisher Scientific	#15710049
Pencillin-Streptomycin	ThermoFisher Scientific	#151140122
Amphotericin B	ThermoFisher Scientific	#15290018
Formaldehyde	Sigma-Aldrich	#F8775
KaryoMAX™ Colcemid	Gibco	#15212012
PhotoFlo	Kodak	#1464510
Trypsin 0.05%	Gibco	#25300062
NEB2 buffer	New England Biolabs	#B7002S
MboI	New England Biolabs	#R0147M
Proteinase K	New England Biolabs	#P8107S
Phenol/Chloroform/Isoamyl	Sigma-Aldrich	#P2069-400ML
dNTPs	Roche	#11969064001
Biotin-14-dATP	ThermoFisher Scientific	#19524-016
DNA Polymerase I, large (Klenow) Fragment	New England Biolabs	#M0210M
T4 DNA Ligase	New England Biolabs	#M0202M
RNAse A	ThermoFisher Scientific	#EN0531
Dynabeads MyOne Streptavidin T1	ThermoFisher Scientific	#65001
T4 Polynucleotides Kinase	New England Biolabs	#M0201L
T4 DNA Polymerase	New England Biolabs	#M0212M
Klenow Fragment 3' → 5' exo$^-$	New England Biolabs	#M0203L
Nuclease micrococcal from Staphylococcus aureus	Sigma-Aldrich	#N3755
TRIzol	Invitrogen	#15596026
Critical commercial assays		
Dynabeads™ Protein G Immunoprecipitation Kit	ThermoFisher Scientific	#10007D
Truseq ChIP-seq library preparation kit	Illumina	#IP-202-1012
Deposited data		
Elephant Hi-C data	This paper	GEO: GSE206075
Aardvark Hi-C data	This paper	GEO: GSE206075
Tasmanian devil Hi-C data	This paper	GEO: GSE206075
Tammar wallaby Hi-C data	This paper	GEO: GSE206075
Human Hi-C data	Rao et al. (2014)[1]	GEO: GSE63525
Mouse Hi-C data	Vara et al. (2019)[3]	GEO: GSE132054
Platypus Hi-C data	Zhou et al. (2021)[28]	SRA: SRR10530604
Chicken Hi-C data	Fishman et al. (2019)[29]	GEO: GSE96037

(*Continued on next page*)

Continued

REAGENT orRESOURCE	SOURCE	IDENTIFIER
Elephant ChIP-seq data	This paper	GEO: GSE206075
Aardvark ChIP-seq data	This paper	GEO: GSE206075
Tasmanian devil ChIP-seq data	This paper	GEO: GSE206075
Human ChIP-seq data	Schmidt et al. (2012)[30]	ArrayExpress: E-MTAB-437
Mouse ChIP-seq data	Schmidt et al. (2012)[30]	ArrayExpress: E-MTAB-437
Opossum ChIP-seq data	Schmidt et al. (2012)[30]	ArrayExpress: E-MTAB-437
Tasmanian devil RNA-seq data	This paper	GEO: GSE206075
Elephant RNA-seq data	Cortez et al. (2014)[54]	GEO: GSE50747
Experimental models: Cell lines		
Elephant XX/XY fibroblast cell lines	provided by T. J. Robinson	N/A
Aardvark XY fibroblast cell line	provided by T. J. Robinson	N/A
Tasmanian Devil XX fibroblast cell line	provided by J. Deakin	N/A
Tammar wallaby XY fibroblast cell line	This paper	N/A
Software and algorithms		
ImageJ	Schneider et al. (2012)[55]	https://imagej.nih.gov/ij/
Juicer/3D-DNA	Dudchenko et al. (2017)[26]	https://github.com/aidenlab/3d-dna
BBDuk (version 09/2019)	Bushnell (2014)[56]	https://github.com/BioInfoTools/BBMap/blob/master/sh/bbduk.sh
TADbit (version 1.0)	Serra et al. (2017)[34]	https://github.com/3DGenomes/TADbit
GEM3-Mapper (version 3.0)	Marco-Sola et al. (2012)[57]	https://github.com/smarco/gem3-mapper
HiCExplorer (version 3.6)	Wolff et al. (2018)[58]	https://github.com/deeptools/HiCExplorer/blob/master/docs/index.rst
Cooltools	Venev et al. (2021)[59]	https://github.com/open2c/cooltools
Rstats (version 3.6.2)	Schwarzer (2007)[60]	https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html
Bedtools	Quinlan and Hall (2010)[61]	https://github.com/arq5x/bedtools2
FAN-C (version 0.9.1)	Kruse et al. (2020)[62]	https://github.com/vaquerizaslab/fanc
FastQC (version 0.11.9)	Andrews et al. (2014)[63]	https://github.com/s-andrews/FastQC
Trimmomatic (version 0.39)	Bolger et al. (2014)[64]	http://www.usadellab.org/cms/?page = trimmomatic
BWA-MEM (version 0.7.17)	Li et al. (2013)[65]	https://github.com/lh3/bwa
MACS2 (version 2.2.7.1)	Zhang et al. (2008)[66]	https://pypi.org/project/MACS2/
STAR (version 2.7.10a)	Dobin et al. (2013)[67]	https://github.com/alexdobin/STAR
featureCounts	Liao et al. (2014)[68]	https://www.rdocumentation.org/packages/Rsubread/versions/1.22.2/topics/featureCounts
countsToFPKM	Alhendi (2019)[69]	https://github.com/AAlhendi1707/countToFPKM
UCSC Kent Utilities	Kent at al. (2010)[70]	https://github.com/ENCODE-DCC/kentUtils
RepeatMasker (version 4.0.9)	Tempel (2012)[71]	https://github.com/rmhubley/RepeatMasker
LASTZ	Harris (2007)[72]	https://github.com/lastz/lastz
maf2Synteny	Kolmogorov et al. (2018)[73]	https://github.com/fenderglass/maf2synteny
syntenyPlotteR	Farré et al. (2019)[14]	https://github.com/marta-fb/syntenyPlotteR
DESCHRAMBLER	Kim et al. (2017)[74]	https://github.com/jkimlab/DESCHRAMBLER
TimeTree	Kumar et al. (2017)[75]	http://www.timetree.org/

(*Continued on next page*)

Continued

REAGENT orRESOURCE	SOURCE	IDENTIFIER
FigTree	Rambaut (2009)[76]	https://github.com/rambaut/figtree/releases/tag/v1.4.4
GO::TermFinder	Boyle et al. (2004)[77]	https://metacpan.org/pod/GO::TermFinder
RegioneR (version 1.26)	Gel el at. (2016)[78]	https://bioconductor.org/packages/release/bioc/html/regioneR.html
GenoMatriXeR	Álvarez-González et al. (2022)[4]	https://github.com/RMalinverni/GenoMatriXeR

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to, and will be fulfilled by, the Lead Contact: Aurora Ruiz-Herrera (aurora.ruizherrera@uab.cat).

Materials availability

All materials developed in this study will be available from the lead contact upon request.

Data and code availability

- Raw and processed data for all the experiments performed on this paper have been deposited in Gene Expression Omnibus database under de accession code: GSE206075.
- This paper does not report original code. The software used in this described in the key resources table in details.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

Five primary fibroblast cell lines were used in this study. The Tasmanian devil fibroblast cell line (*Sarcophilus harrisii*, one female) was provided by J. Deakin and derived from a skin biopsy that was collected under permit AEECP R.CG.11.06 from Australian National University Animal Experimentation Ethics Committee.

The cryopreserved fibroblast cell lines from aardvark (*Orycteropus afer*, one male) and African elephant (*Loxodonta africana*, one male and one female) were derived from skin biopsies and are part of the South African National Biodiversity Institute (SANBI) Biobank collection (accession number EXT03570_27112020 and catalog numbers 76993, 82,607, 82,608).

The tammar wallaby (*Macropus eugenii*, one male) fibroblast cell line was established from a connective tissue biopsy from one male following standard procedures (see STAR Methods). The animal was held under appropriate permits in open grassy yards in a breeding colony of tammar wallabies of Kangaroo Island South Australia origin. The male was euthanized as previously described.[79] Tissue collection and experiments were approved by the University of Melbourne Animal Experimentation Ethics Committees in accordance with the National Health and Medical Research Council of Australia (2013 and 2014) guidelines.[80,81]

METHOD DETAILS

Primary fibroblast cell lines

Samples of connective tissue were washed in 1xPBS supplemented with an antibiotic-antimycotic solution (100U/ml penicillin, 100μg/ml streptomycin, 50μg/ml gentamicin and 0.25μg/ml amphotericin B) (ThermoFisher Scientific). Cultures were established by disaggregating tissue with a scalpel blade and resuspending cells in AmnioMAX (ThermoFisher Scientific). Cell cultures were incubated at 37°C in 10% CO2. A cell fraction of each cell line was kept for immunofluorescence analysis and quality controls (see below) and the rest were fixated with 1% formaldehyde for 10′ at RT for Hi-C and ChIP-seq (see below).

Quality controls consisted of the analysis of modal karyotypes. Briefly, cells were arrested in metaphase by adding 80μL of Colcemid (10μg/ml) to 10mL of medium for 2h and then trypsinised. Cells were spun down at 600 xg for 7 min and resuspended in 5mL of hypotonic solution (0.075M KCl) for 30 min at 37°C. Chromosomes were then fixed by addition of fixative solution (3:1 methanol/acetic acid) and metaphase spreads were obtained by dropping 15 μL of cell suspension onto a cleaned dry slide. Slides were baked at 65°C for one hour and kept at −20°C until use.

Metaphases were stained homogenously with Giemsa solution for analysis of the modal karyotype. An optical microscope (model Zeiss Axioskop) equipped with a charged coupled device camera (ProgResR CS10Plus, Jenoptik) was used for the microscope

analysis. Good-quality metaphases were captured with the program Progress Capture 2.7.7 and analyzed for each specimen, obtaining the modal karyotype.

Immunoflourescence and microscopy

A total of 50,000 cells were sub-cultured on sterilized slides O/N at 37°C. Slides were then washed twice with 1x PBS and fixed (4% PFA +0.15% Triton X-100 in MiliQ water) for 20 min. Slides were left to dry and washed with twice PhotoFlo 1% and then blocked with PBS-Tween-20 (0.05%). Cells were incubated overnight at 4°C with the following primary antibodies: anti-H3K9me3 (1:300) and anti-CREST (1:50). Primary antibodies were detected with the secondary antibodies: anti-rabbit Cy3 (1:200) and anti-mouse FITC (1:200). Slides were finally mounted with DAPI and analyzed with a fluorescence microscopy (Axiophot, Zeiss) coupled with a ProgRes CS10plus, Jenoptik camera. Representative images were captured with ACO XY (A. Coloma, Open Microscopy) and processed using Photoshop and ImageJ[55].

Distances between centromeres (exemplified as the CREST signal) were measured using the Measure analysis tool of ImageJ. Cells diameters were measured and used to correct all distances so the distances could be comparable between cells. A minimum of 1,000 centromeric distances were measured for each species.

In nuclei Hi-C

Hi-C libraries were generated as previously.[1,3,16] Briefly, confluent fibroblasts washed with PBS and fixed with 1% formaldehyde for 10′ at RT. Cells were then incubated with glycine 0.125M for 5′ at RT and for 15′ at 4°C to stop the crosslinking reaction. Then, 2 mL Trypsin were added, and cells were incubated for 8′ at RT and then washed twice with PBS. Cells were scraped and collected in a tube and centrifuged for 5 min at a maximum speed of 1,800 xg.

Crosslinked cells were resuspended with lysis buffer and incubated on ice for 30′ and then centrifuged for 5′ at 1,800 xg. Pellets were washed with 1x NEB2 buffer (twice) and resuspended with NEB2 buffer with 10% SDS at RT and incubated for 10′ at 65°C with agitation (300 rpm). NEB2 buffer with 10% Triton X-100 solution was added and cells were incubated for 30′ at 37°C. Cells were then centrifuged for 5′ at 1,800 xg (4°C) and washed with 1x NEB2 buffer twice. An aliquot for ND (Non-Digested) control was taken from the sample to be processed and incubated at 37°C together with the digested sample. 400 U of MboI were added to the rest of the samples and chromatin was digested O/N at 37°C with agitation. Proteinase K (10 mg/mL) was added and the aliquoted samples were incubated for 45′-60′ at 65°C followed by Phenol:Chloroform purification.

Samples were centrifuged for 5′ at 1,800 xg and then washed with 1x NEB2 buffer, twice. After the second wash, samples were directly resuspended with the reparation mix (1x NEB2 buffer, 0.05 mM dCTP, 0.05 mM dTTP, 0.05 mM, 0.05mM biotin-dATP, 50U Klenow). Samples were incubated for 45′ at 37°C and for 10′ at 65°C and centrifuged for 5′ at 1,800 xg and then resuspended with ligation buffer [1x NEB T4 ligase buffer, 0.83% Triton X-100, 0.1 mg/mL BSA, 5 μL ligase (2000 U/μl), 963 μL H2O]. Samples were incubated at 16°C for at least 4h or O/N with mixing, then centrifuged for 5′ at 1,800 xg and resuspended in 1x NEB2 buffer. Samples were incubated with RNAseA (10 mg/mL) for 15 min at 37°C. The mix was incubated with Proteinase K (10 mg/mL) at 65°C O/N to reverse the cross-link. Samples were cooled to RT and purified with Phenol/Chloroform/Isoamyl alcohol. DNA content was measured on a Qubit. Samples were sonicated: 20 s time ON, 60 s time OFF, 8 cycles. Samples were then loaded in an electrophoresis gel of 1.2% agarose to check fragment size.

Samples were incubated for 30′ with rotation at RT with Dynabeads MyOne Streptavidin T1 beads and 2x Binding Buffer (10 mM TrisHCl, 1mM EDTA, 2M NaCl). Beads were washed twice with Binding Buffer and resuspended in the end repair mix [1x NEB T4 DNA ligase buffer with 10 mM ATP, 25 mM dNTP mix, 10U/μl NEB T4 PNK, 3U/μl T4 DNA polymerase I, 5U/μl NEB DNA polymerase I (Klenow)]. Samples were incubated for 30′ at RT. Beads were washed with 1x Binding Buffer, twice. Beads were subsequently resuspended in the dATP attachment master mix (1x NEBuffer 2, 0.5mM dATP, 5U/μl NEB Klenow exo minus). Samples were incubated at 37°C for 30'. The beads were washed with 1x Binding Buffer, twice and resuspended with 1x NEB Quick ligation buffer. Libraries were quantified using a Qubit fluorometer and submitted for Illumina sequencing (paired end 75bp each side on HiSeq 4000).

African elephant Hi-C assisted assembly

The HiC data from the African elephant generated on this work was used to upgrade the assembly (Table S2). This was done using the Juicer/3D-DNA pipeline described in Dudchenko and collaborators (2017) with default parameters (https://www.dnazoo.org/assemblies/Loxodonta_africana).[26,27]

Hi-C data processing, binning and normalisation

BBDuk (version 09/2019)[56] was used for quality check and trimming of the raw fastq sequences. Setting a minimum read length of 35 bp and a minimum Phred quality score of 20, adapters and low-quality reads were removed while preserving their longest high-quality regions. HiC data processing was conducted using TADbit (version 1.0)[34]. Fastq high-quality reads were mapped against the reference genome of each species (Table S1 and S2 using GEM3-Mapper (version 3.0)[57]. Reads were mapped using windows from 15bp to 75bp in 5bp steps.

Pairs of mapped read-ends were filtered in order to keep only valid pairs. We used the following filters provided by TADbit to remove artifacts: "self-circle," "dangling-end," "error," "extra dangling-end," "too short," "too large," "duplicated," and "random

breaks.'' The maximum molecule length parameter was set at 2 times the 99.9 percentile of the insert size distribution, returned by the ''insert_size'' from TADbit. The maximum distance of a read to a cleavage site was set to the 99.9 percentile of the insert size distribution.

An in-house script was used for binning and normalization of the data as previously described.[3] The script uses a TADbit module to read the artifacts-filtered map files and bin then into a square matrix of 50 Kbp. Using the tool 'hicMergeMatrixBins' from HiCExplorer (version 3.6)[58], 500 and 50 Kbp matrices were created and corrected using 'hicCorrectMatrix' that uses ICE (Iterative Correction and Eigenvector decomposition) method. For comparison purposes, all processed matrixes at 50 Kbp and 500 Kbp were normalized to a total of 100,000,000 interaction counts by scaling the sum of all interactions within the matrix. The tool 'hicCorrelate' from HiCExplorer was used to compute pairwise correlations between Hi-C replicates (based on the Pearson correlation method).

Averaged contact probability P(s) and its derivative

Using as an input the corrected 500 Kbp matrices normalized to 100M counts, both contact probability vs distance curves (P(s)) were calculated genome-wide and by chromosome using the HiC analysis tools Cooltools.[59]The derivative of the contact probability vs distance curve (herein slope), was also calculated using Cooltools. Average loop sizes were then estimated as highest point depicted in contact probability slopes distribution according to previous studies.[82,83]

Inter-chromosome/intra-chromosome interaction ratio

Intra- and inter-chromosomal interactions were obtaining by converting the corrected 500 Kbp matrices to normalized 100M counts matrices from h5 to ginteractions, using the tool 'hicConvertMatrix' from HiCExplorer. Then, the mean number of inter- and intra-chromosomal interactions per chromosome was calculated to obtain the inter-chromosome/intra-chromosome interaction ratio using Rstudio. The function *lm model* (fitting linear model) from stats (version 3.6.2)[60] was used to calculate the linear regression.

Centromere interaction quantification

Centromeric regions were identified in the genomes of aardvark, African elephant, Tasmanian devil and tammar wallaby based on the decay of intra-chromosomal interactions, the Ns content and chromosomal morphology.[35,40,84–87] Intra-chromosomal interaction frequency decay was calculated by transforming intra-chromosomal interaction into *Z* score interactions for each chromosome. The *Z* score interactions were then smoothed by locally weighted regression (loess) and plotted using ggplot2. The Ns content of each matrix bin was done using bedtools nuc.[61]Once centromeric regions were defined, inter-chromosomal interactions were represented using ggplot2 from Rstudio.

Centromere aggregate contacts plots

The Hi-C submatrices of the centromeric region of each species and its aggregate were calculated and plotted using the HiCExplorer function 'hicAggregateContacts' on genome-wide and inter-chromosomal interactions mode. As input 500Kbp corrected and normalized HiC matrices were used.

First eigenvector and insulator score calculation

The HiC analysis tools package, FAN-C (version 0.9.1)[62] was used to obtain first eigenvector and insulator score values. To calculate the first eigenvector the tool 'fanc compartments' was used with normalized 500Kbp matrices and default parameters. 'fancplot' was used to generate the compartment matrices. For the insulator score, the tool 'fanc insulation' was used with normalized 500Kbp matrices and the following windows: 0.5Mbp, 1Mbp, 1.5Mpb, 2Mbp. The analysis of mean values of the first eigenvector were used to analyze compartment strength. The higher the first eigenvector value, the higher the compartment signal.

TAD calling and TAD boundaries

TADs were called using the tool find_tads from TADbit (version 1.0)[34] with default parameters and normalized 500Kbp matrixes. Only TAD boundaries with a TADbit score >6 were considered for the following analysis.

ChIP-sequencing

For chromatin immunoprecipitation, antibodies for CTCF and H3K4me3 were used. Two biological replicates of H3K4me3 ChIP-sequencing (ChIP-seq) were performed using primary fibroblast (30 million cells for each replicate) from African elephant, aardvark and Tasmanian devil. Two biological replicates of CTCF ChIP-sequencing (ChIP-seq) were performed using primary fibroblast (30 million cells for each replicate) of African elephant and aardvark.

Crosslinked cells (as described in *in nuclei* Hi-C) were incubated on ice in lysis buffer (0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA, 10 mM Tris-HCl) and lysis wash buffer (200 mM NaCl, 1 mM EDTA, 0.5mM EGTA, 10 Mm Tris-HCl). Samples were then digested using MNase I to fragment the chromatin below 200bp. Fragmented chromatin was diluted in cold binding buffer (from Dynabeads Protein G Immunoprecipitation Kit). 50 μL of each sample was kept as input control and the remaining was divided in two aliquots, where antibodies for CTCF and H3K4me3 were added (one in each aliquot) and then incubated overnight at 4°C. The antibody-chromatin pull-down was done using Protein G Dynabeads, following the manufacturer's protocol. Finally, proteins were digested and the DNA purified with phenol:chloroform:isoamyl alcohol. Quality-check of the inmunoprecipate was done using Qubit

and Bioanalyzer Agilent. Libraries were constructed using commercial Truseq ChIP-seq library preparation kit and sequenced by paired-end Illumina Novaseq 6000.

ChIP-seq peak calling and annotation

Quality check was performed using FastQC (version 0.11.9)[70] and reads were trimmed using Trimmomatic (version 0.39).[64] Trimmed fastq files were then aligned to the reference genome using bwa-mem (version 0.7.17).[65] Alignments with a MAPQ >30 were extracted and used as input for MACS2 (version 2.2.7.1)[66] with either the default settings of *callpeak* to produce narrow peaks (CTCF) or with – *broad - -broad-cut-off 0.05* to produce broad peaks (H3K4me3). Since two replicates were sequenced, we selected the peaks in both replicates using bedtools intersect.[61]

In African elephant 37,210 CTCF peaks were common between replicates (44,898 peaks on replicate #1 and 40,014 on replicate #2) and 67,874 H3K4me3 peaks were common between replicates (71,940 peaks on replicate #1 and 74,949 on replicate #2). In aardvark, a total of 12,478 CTCF peaks were common between replicates (16,012 peaks on replicate #1 and 14,835 on replicate #2) and 28,819 H3K4me3 peaks were common between replicates (32,525 peaks on replicate #1 and 33,001 on replicate #2). Finally, in the Tasmanian devil, a total of 43,899 H3K4me3 peaks were common between replicates (49,454 peaks on replicate #1 and 50,565 on replicate #2).

African elephant genome annotation

To annotate genes in the African elephant genome we used Gene Model Mapper (GeMoMa).[88]Because GeMoMa is a homology-based gene prediction tool, we started with two sets of reference gene annotations: tenrec (GCF_000313985.2_ASM31398v2) and African elephant (GCF_000001905.1_Loxafr3.0). GeMoMa was run using default parameters without RNA-Seq data. We then combined both independent annotations using the GAF command from GeMoMa with a weighting of 10 and 20 for tenrec and African elephant genomes, respectively and a combined sumWeight $\geq$20. Finally, we used gffCompare to assess the differences between the annotations and identify truncated genes. After manual curation, a total of 20,641 genes were annotated in the African elephant genome.

RNA-seq analysis

Fibroblast RNA was TRIzol (Invitrogen) extracted from Tasmanian devil cultured fibroblasts according to manufactures instructions. Illumina stranded mRNA libraries were prepared and sequenced (NovaSeq 6000 SP 2 × 100bp) at the Ramaciotti Center for Genomics (UNSW Sydney, Australia).

Raw African elephant RNA-seq data was obtained from Cortez and collaborators.[54] Raw reads were mapped with STAR (version 2.7.10a)[67] using the end-to-end alignment mode and gene expression quantification was performed with featureCounts.[68] Counts were transformed into FPKMs using the R package countToFPKM.[69]

Whole-genome pairwise alignments

The genomes of 20 mammals were downloaded from NCBI or DNA Zoo (Table S2). Each genome, assembled either at chromosome or HiC-scaffold level, was filtered to remove unplaced scaffolds using faFilter tool from UCSC Kent Utilities.[70] Repetitive elements were masked using RepeatMasker (version 4.0.9)[71] with the parameters -xsmall -species mammal.[68] Three sets of pairwise alignments were done, using human, Tasmanian devil or African elephant as reference genomes. All alignments were generated using LASTZ[72] with the parameters: -minScore = 1000, -linearGap = medium, C = 0, E = 30, K = 3000, L = 3000, O = 400. LASTZ alignments were extended into chains and nets with axtChain (parameters: -minScore = 1000 -verbose = 0 -linearGap = medium), followed by chainAntiRepeat, chainSort, chainPreNet, chainNet and netSyntenic, all with default parameters.[89] In Tasmanian devil, to bypass the hardcoded limits of UCSC Kent tools, LASTZ alignment files bigger than 50Gb were split into smaller files, concatenated into chains using axtChain with the above paraments and then merged using chainMergeSort. In all cases, Y chromosomes were omitted due to their low-quality assembly.

Homologous synteny blocks (HSBs) were then constructed using maf2Synteny[73] at three resolutions of detection (100 Kbp, 300 Kbp and 500 Kbp). The 300 Kbp resolution was chosen for further analysis because it achieved a balance between higher coverage of the reference genome and number of rearrangements detected. HSBs were visualised using syntenyPlotteR.[14]

Conservation of the higher-order chromatin organization between species

The higher-order structural conservation between HSBs was evaluated by pairwise comparisons of the first eigenvector and insulator score values. We included in the comparative analysis HSBs >2 Mbp (spanning a minimum of 4 HiC 500Kbp bins) as the average TAD size was 1 Mbp in the species analyzed.

Pairwise comparisons were done by testing whether first eigenvector and insulator score mean values were equivalent between HSBs of two given species using a two-sided t test with a confidence level of 99%. Pairwise comparisons with a p value >0.01 were considered conserved (null hypothesis). The corresponding first eigenvector and insulator score values were obtained by extracting the values of each block out of the genome-wide data using bedtools intersect.[63] For eigenvector comparisons, all positive values (A compartments) were transformed into 1 and all negatives (B compartments) into −1.

Ancestral karyotype reconstructions

Reconstructed ancestral chromosome fragments (RACFs) for the marsupial and afrotherian ancestors were generated using DESCHRAMBLER[74] with a syntenic fragment resolution of 300 Kbp and a minimum adjacency score of 0.0001. We modified the hardcoded constriction of 500,000,000 bp chromosome size to 1 Gbp in DESCHRAMBLER to allow for Tasmanian devil chromosome sizes. Both phylogenetic trees used in the analysis were estimated using the divergence times between branches defined in TimeTree[75] and visualized using FigTree.[76]

Two ancestors of the marsupial lineage were reconstructed: MAK, the ancestral karyotype of all marsupial species; and MAUK, the ancestral karyotype of Australian marsupials (Australidelphia ancestor). The number of RACFs produced from DESCHRAMBLER was higher than the number of chromosomes suggested in previous studies24. Consequently, the adjacent RACFs in each of the reconstructed ancestors were manually merged using both reference genome and other reconstructed ancestors, which were most closely related. This process was started on the MAK ancestor using Tasmanian devil as a point of reference, before working back in evolutionary time using the closest related ancestors as a point of reference.

Only one ancestor of the Afrotheria lineage was reconstructed (AFAK). RACFs were manually curated and merged following the same steps as above. Plots of the RACFs were produced with the R package syntenyPlotteR.[14]

Detection of evolutionary breakpoint regions (EBRs)

Evolutionary breakpoint regions (EBRs) in all ancestors were identified by considering the Tasmanian devil or African elephant genomes as a reference. EBRs were counted from the coordinates of manually merged RACFs. In marsupials, when one EBR was shared for both ancestors but each of them presented a different size, the coordinates chosen to define that EBR were the ones corresponding to the minimum size. EBRs were classified into well-defined EBRs (EBRs ranging from 0 Kbp to 50 Kbp) and no well-defined EBRs (EBRs ranging from 50 Kbp to 300 Kbp), and the EBRs higher than 300 Kbp were discarded of our analysis. The EBRs were phylogenetically classified depending on the ancestral lineage in which they occurred. EBRs were further separated by the type of rearrangement they delimited into inversion EBRs or inter-chromosome EBRs, if they demarcated inversions or were the result of fusion or fission events, respectively.

Further characterization of the EBRs included multipermutation test analysis using different sets of structural and functional marks, such as: TSS, H3K4me3, expressed genes, A compartments, compartment boundaries and TAD boundaries.

Gene ontology enrichment analysis (GOEA)

Unique protein coding gene IDs of Tasmanian devil genome mSarHar1.11 were identified using Ensembl BioMart.[90] For African elephant, our new gene annotation was used. We then used bedIntersect tool to select genes within the different sets of regions analyzed.[72] The basic version of the GO annotations (go-basic.obo) was downloaded from the Gene Ontology Project website (Ashburner et al., 2000; The Gene Ontology Consortium, 2021). The GOEA was performed using a modified version of GO:TermFinder Perl module.[77] We considered significantly enriched only those terms with a corrected p-value ≤ 0.05 and a calculated false discovery rate (FDR) $\leq 1\%$. Plots were created with ggplot2 (version 3.3.5) in R (version 4.1.0).

QUANTIFICATION AND STATISTICAL ANALYSIS

Multi-association and statistical analysis

In the case of multi-association analysis, statistical association between different genomic features was evaluated using the RegioneR R package version 1.26[78]. Based on the regioneR package we used a series of functions to allow the calculation of associations between multiple regionsets. Due to the implementation of multiple comparisons, the p value calculation was adjusted using Benjamini – Hochberg procedure.

The value of the Z score for association with an adjusted p value greater than 0.005, was considered as 0. The Z score was subsequently normalised by dividing it by the square root of n where n is the number of regions present in the permuted regionset. All the permutations were performed using randomizeRegions and NumOverlaps respectively as randomization and evaluation function.[4]

Information regarding statistical significance is provided in the legends to the figures. Data are presented as mean ± standard error of the mean. In the figures, *** indicates $p < 0.001$. The threshold for statistical significance was $p = 0.05$.

Cell Reports, Volume *41*

Supplemental information

Principles of 3D chromosome folding and evolutionary genome reshuffling in mammals

Lucía Álvarez-González, Cristina Arias-Sardá, Laia Montes-Espuña, Laia Marín-Gual, Covadonga Vara, Nicholas C. Lister, Yasmina Cuartero, Francisca Garcia, Janine Deakin, Marilyn B. Renfree, Terence J. Robinson, Marc A. Martí-Renom, Paul D. Waters, Marta Farré, and Aurora Ruiz-Herrera

SUPLEMENTARY FIGURES

Figure S1

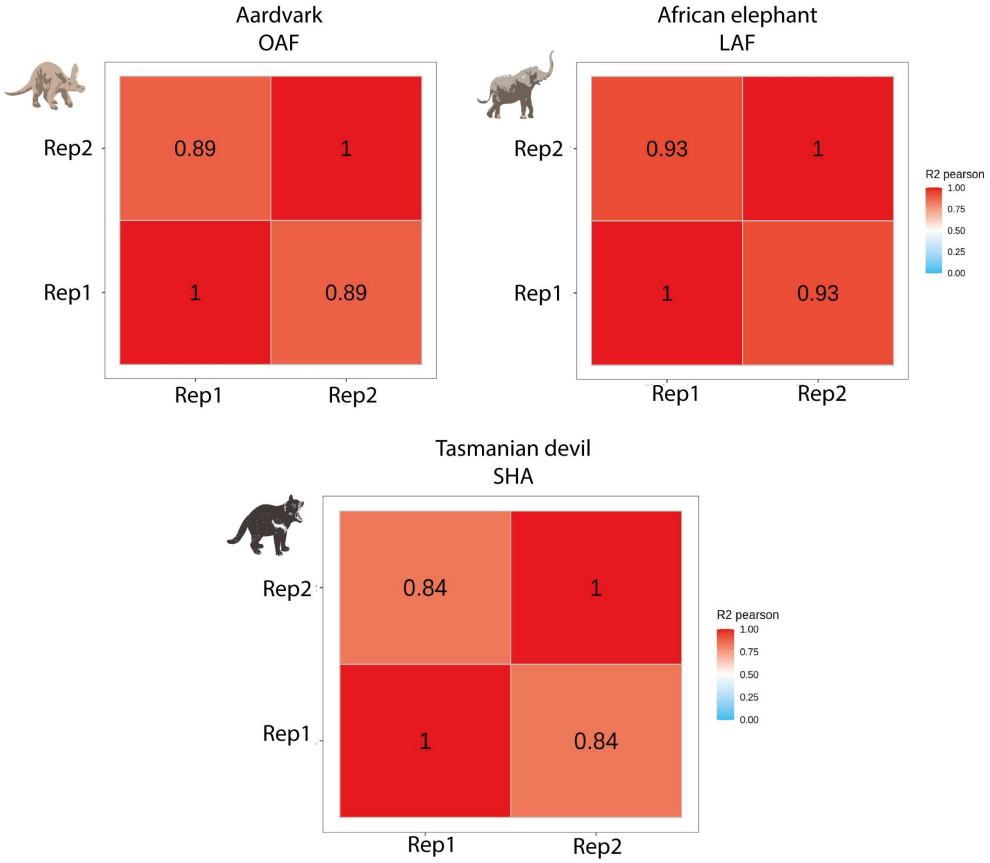


Figure S1. Pearson correlation heatmaps showing reproducibility of the Hi-C maps between biological replicates, **Related to Figure 1.** Two biological replicates for each species are indicated (Rep1 and Rep2).

Figure S2

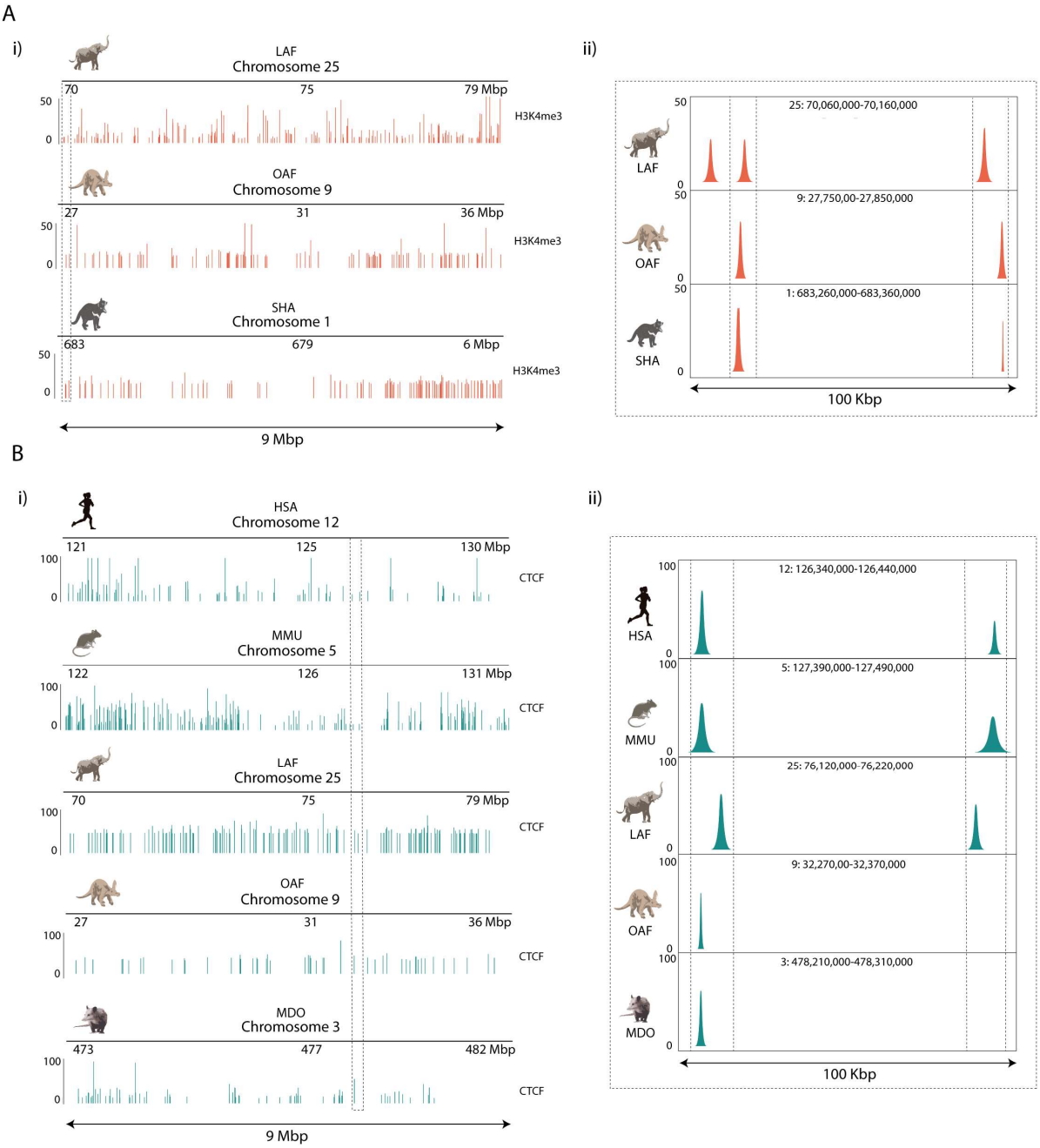


Figure S2. Genomic tracks showing H3K4me3 and CTCF occupancy in specific orthologous regions, Related to Figure 1. (A) H3K4me3 occupancy in the African elephant (LAF) chromosome 25, aardvark (OAF) chromosome 9 and Tasmanian devil (SHA) chromosome 1. Left panel (i) corresponds to a 9 Mbp region in each chromosome, whereas right panel (ii) represents a 100 Kbp zoom-in in the same species. (B) CTCF occupancy in the human (HSA) chromosome 12, mouse (MMU) chromosome 5, African elephant (LAF) chromosome 25, aardvark (OAF) chromosome 9 and opossum (MDO) chromosome 3. Left panel (i) corresponds to a 9 Mbp region in each chromosome, whereas right panel (ii) represents a 100 Kbp zoom-in in the same species.

Figure S3

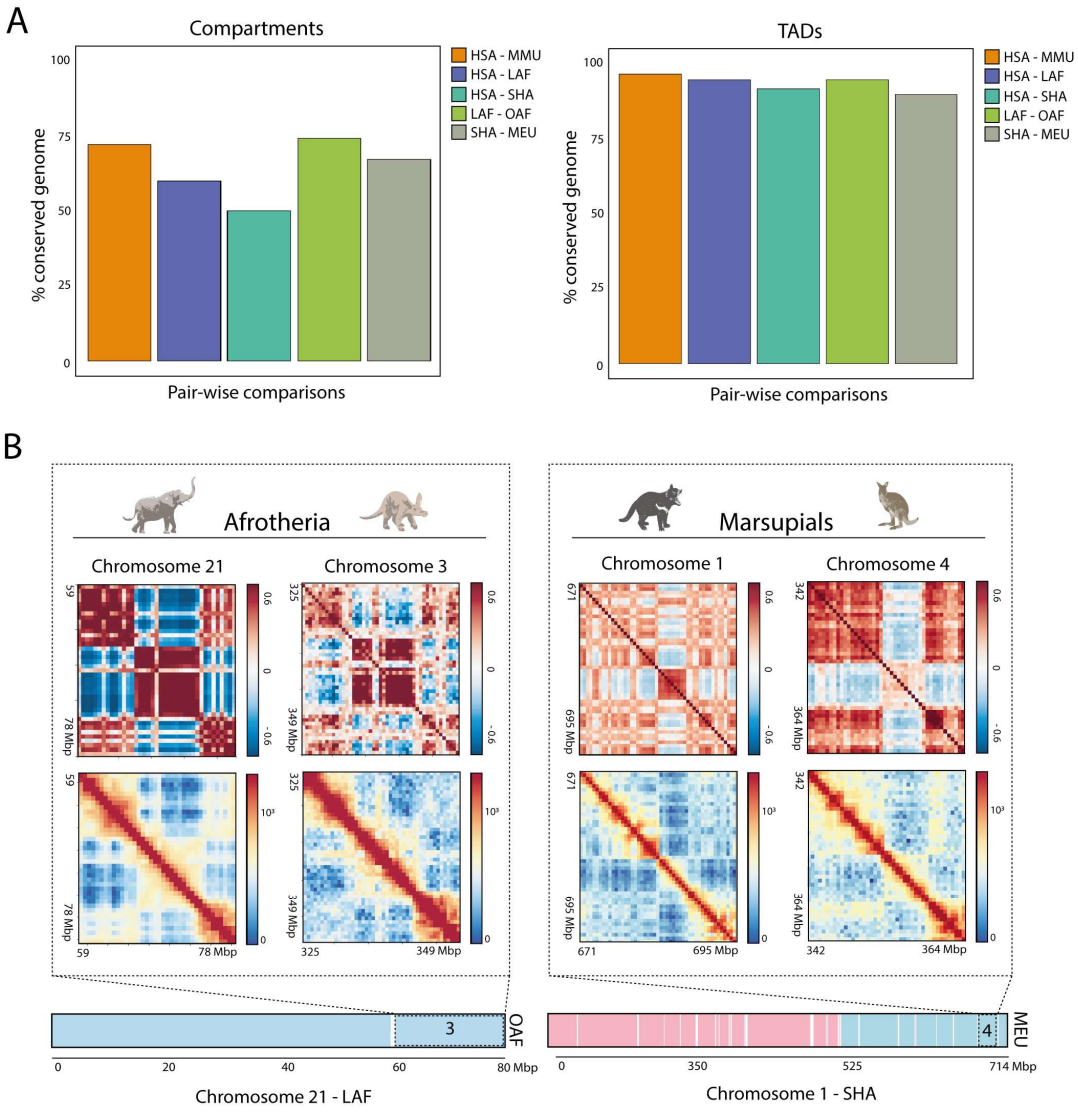


Figure S3. Conservation of the higher-order chromatin organization within mammals, Related to Figure 3. (A) Percentage of compartments and TADs conserved between pair-wise comparisons. Human (HSA), mouse (MMU), African elephant (LAF), aardvark (OAF) and Tasmanian devil (SHA), aardvark (OAF) and tammar wallaby (MEU). (B) Zoom-in of a structural conserved HSB from elephant (LAF) chromosome 21 compared to aardvark (OAF) (a region of chromosome 3) and a conserved HSB Tasmanian devil (SHA) chromosome 1 compared to Tammar wallaby (MEU) (a region of chromosome 4). Compartment conservation is represented as compartment heatmaps and 1st eigenvector distribution, and TAD conservation is represented as contact heatmaps and insulator score distribution.

Figure S4

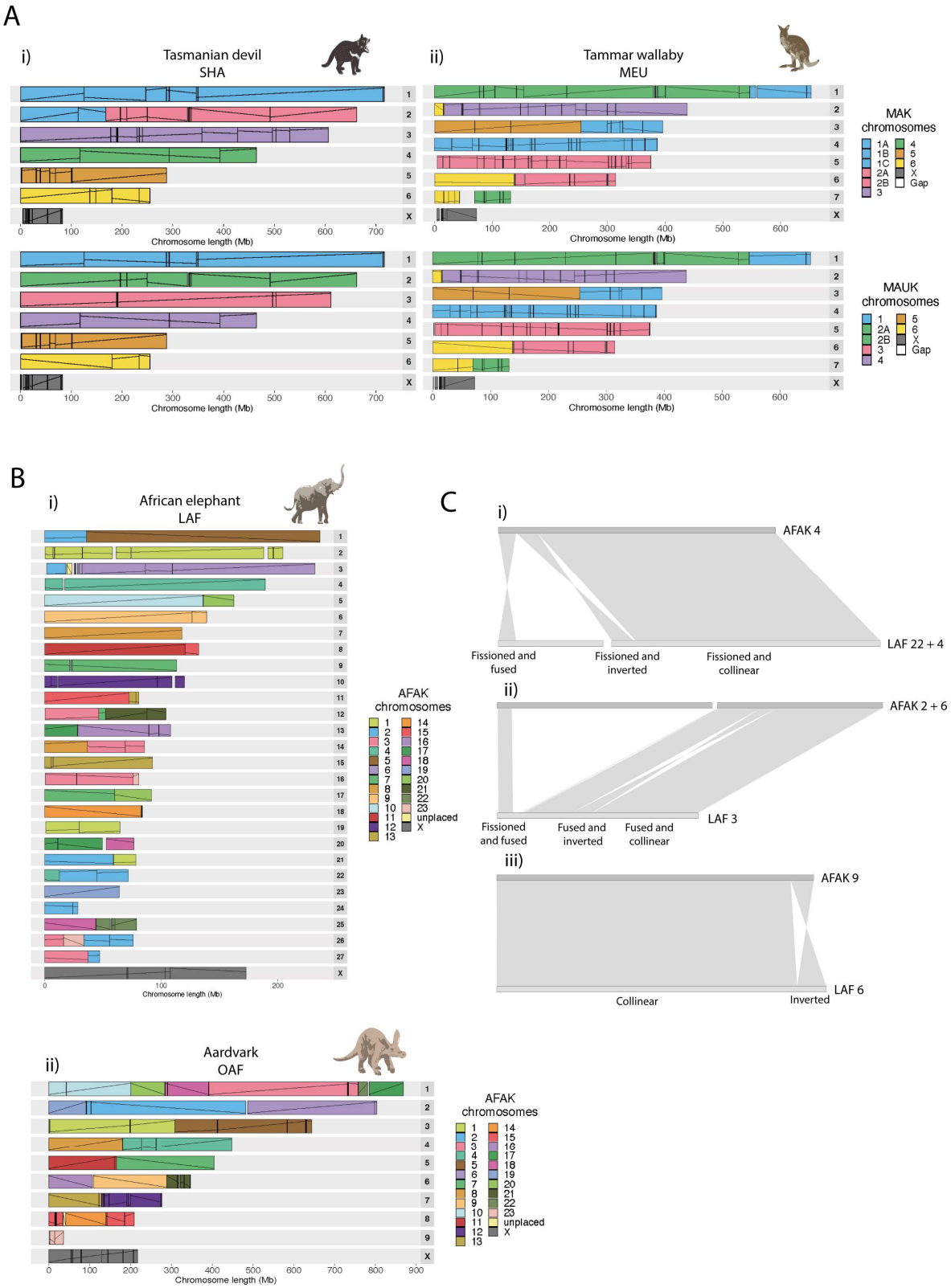


Figure S4. Ancestral genomic reconstructions, Related to Figure 4: (A) Ancestral karyotype reconstructions for marsupials. i) Tasmanian devil chromosome ideograms colour-coded accordingly to their corresponding MAK (marsupial ancestral karyotype) or MAUK (Australian marsupial ancestral karyotype) chromosomes and ii) Tammar

wallaby chromosome ideograms colour-coded with MAK or MAUK. (B) Ancestral karyotype reconstruction for afrotherians. i) African elephant chromosome ideograms colour-coded accordingly to their corresponding AFAK (afrotheria ancestral karyotype) chromosomes and ii) aardvark chromosome ideograms colour-coded accordingly to their corresponding AFAK chromosomes. (C) Types of chromosome rearrangements identified in African elephant genome when compared to AFAK. i) Fission of an ancestral chromosome, leading to a fusion with another chromosome, an inversion or maintained collinear. ii) Fusion of two ancestral chromosomes, leading to a fission, an inversion or collinear, iii) One ancestral chromosome is conserved as a single chromosome in the extant species.

Figure S5

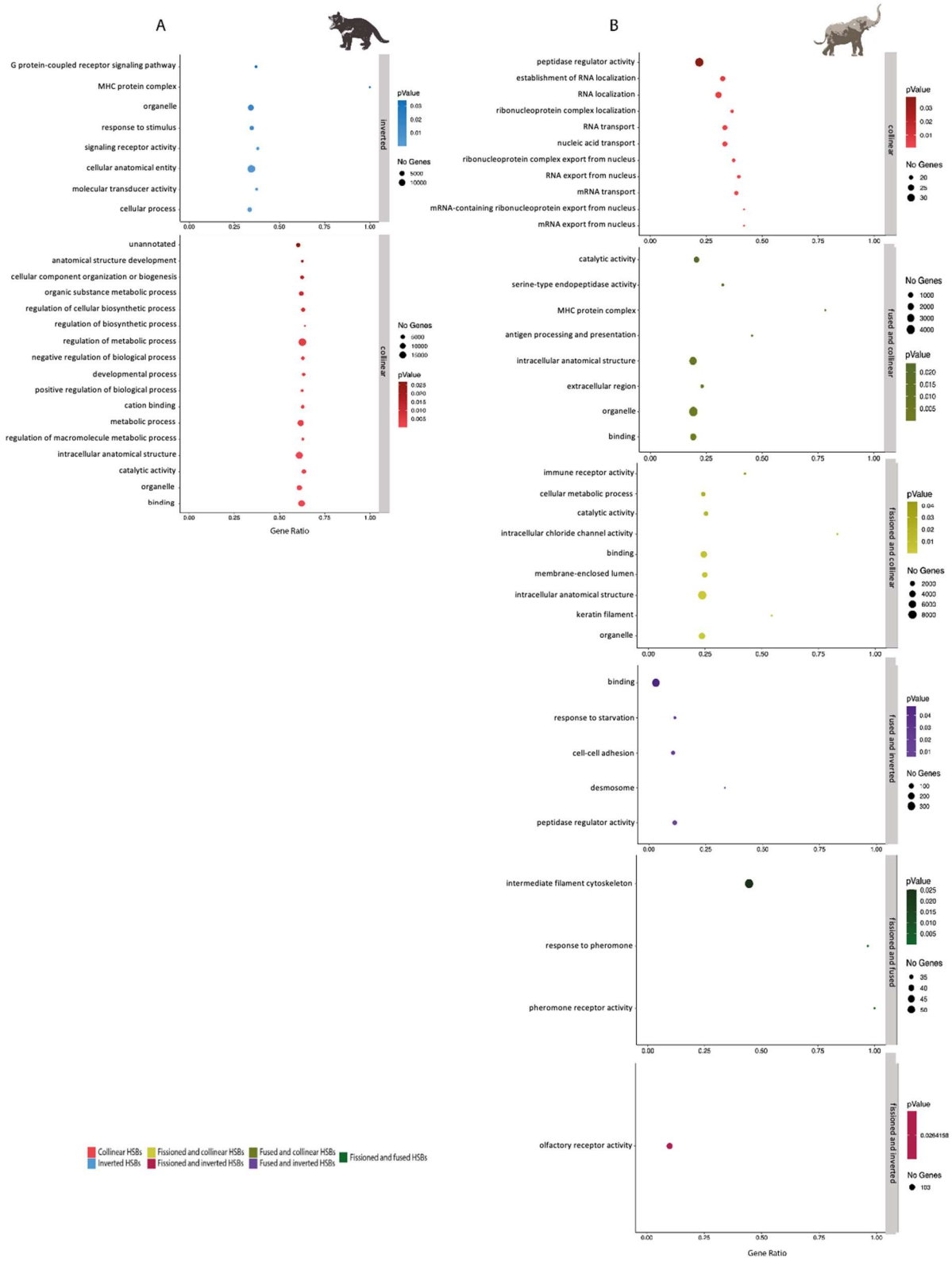


Figure S5. Gene ontology enrichment analysis, Related to Figure 4: (A) the Tasmanian devil and (B) the African elephant based on chromosome rearrangement type: collinear, inverted, fissioned and collinear, fissioned and inverted, fused and collinear, fused and inverted and fissioned and fused.

Figure S6

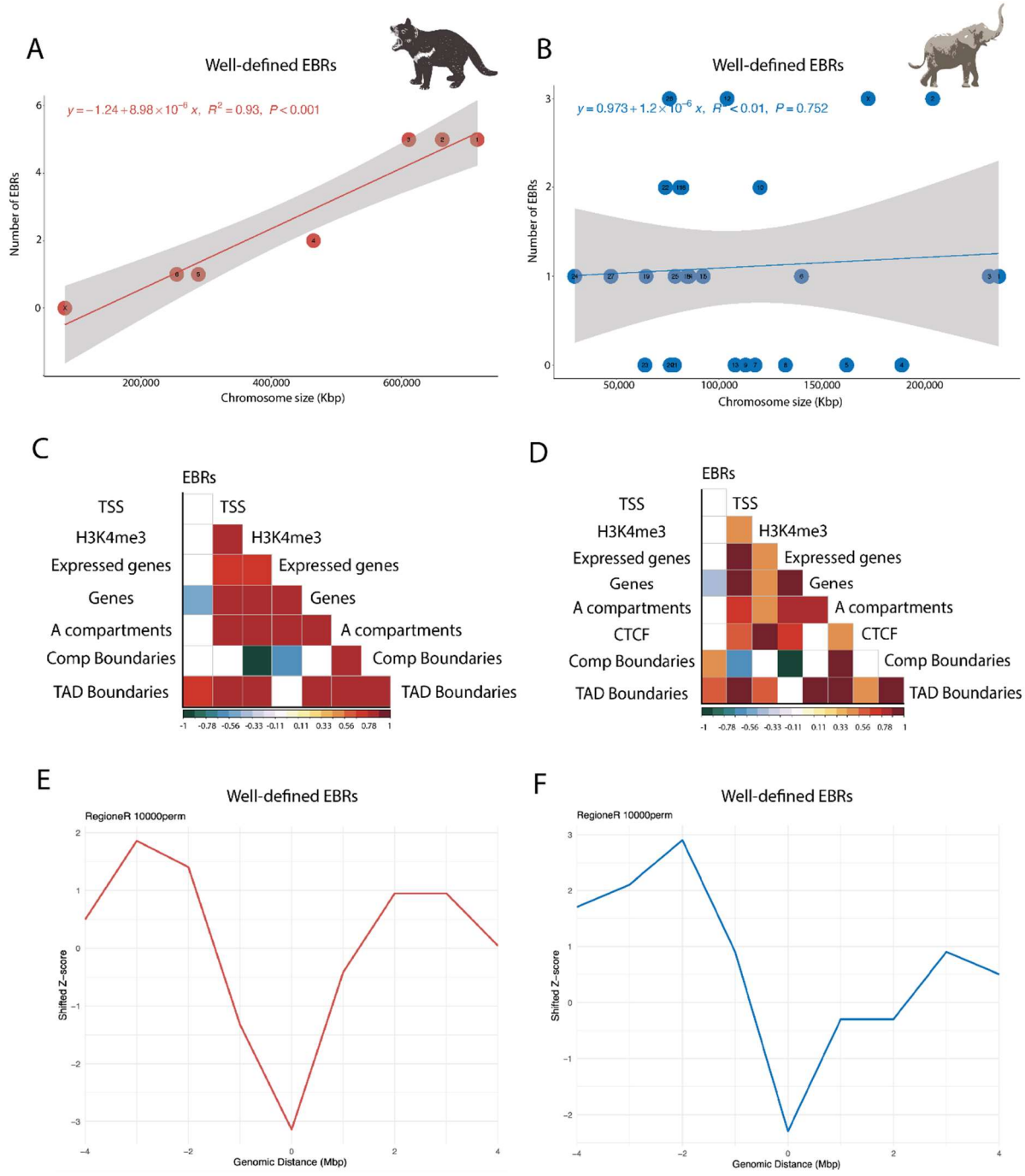


Figure S6. Characteristics of evolutionary breakpoint regions (EBRs), Related to Figure 4. EBRs detected in (A) the Tasmanian devil and (B) the African elephant. (C-D) Heatmaps obtained by regioneR (multicomparison) displaying correlations between EBRs and transcription start sites (TSS), H3K4me3 peaks, genes and expressed genes, A compartments, compartment boundaries and TAD boundaries in (C) the Tasmanian devil and (D) the African elephant. (E-F) Z-score analysis obtained by regioneR indicating the correlation between EBRs and Tasmanian devil (E) and African elephant (F) genes.