

# Patterns of chromosome evolution in ruminants

Cristina Arias-Sardá  | Sarah Quigley  | Marta Farré 

School of Biosciences, University of Kent,  
Canterbury, UK

## Correspondence

Marta Farré, School of Biosciences,  
University of Kent, Canterbury, Kent CT2  
7NJ, UK.

Email: [m.farre-belmonte@kent.ac.uk](mailto:m.farre-belmonte@kent.ac.uk)

Handling Editor: Kay Lucek

## Abstract

Studying when and where gross genomic rearrangements occurred during evolution is key to understanding changes in genome structure with functional consequences that might eventually lead to speciation. Here we identified chromosome rearrangements in ruminants, a clade characterized by large chromosome differences. Using 26 genome assemblies, we reconstructed five ancestral karyotypes and classified the rearrangement events occurring in each lineage. With these reconstructions, we then identified evolutionary breakpoints regions (EBRs) and synteny fragments. Ruminant karyotype evolution is characterized by inversions, while interchromosomal rearrangements occurred preferentially in the oldest ancestor of ruminants. We found that EBRs are depleted of protein coding genes, including housekeeping genes. Similarly, EBRs are not enriched in high GC regions, suggesting that meiotic double strand breaks might not be their origin. Overall, our results characterize at fine detail the location of chromosome rearrangements in ruminant evolution and provide new insights into the formation of EBRs.

## KEYWORDS

Comparative genomics, Genomics/Proteomics, Mammals, Molecular Evolution, Chromosome evolution

## 1 | INTRODUCTION

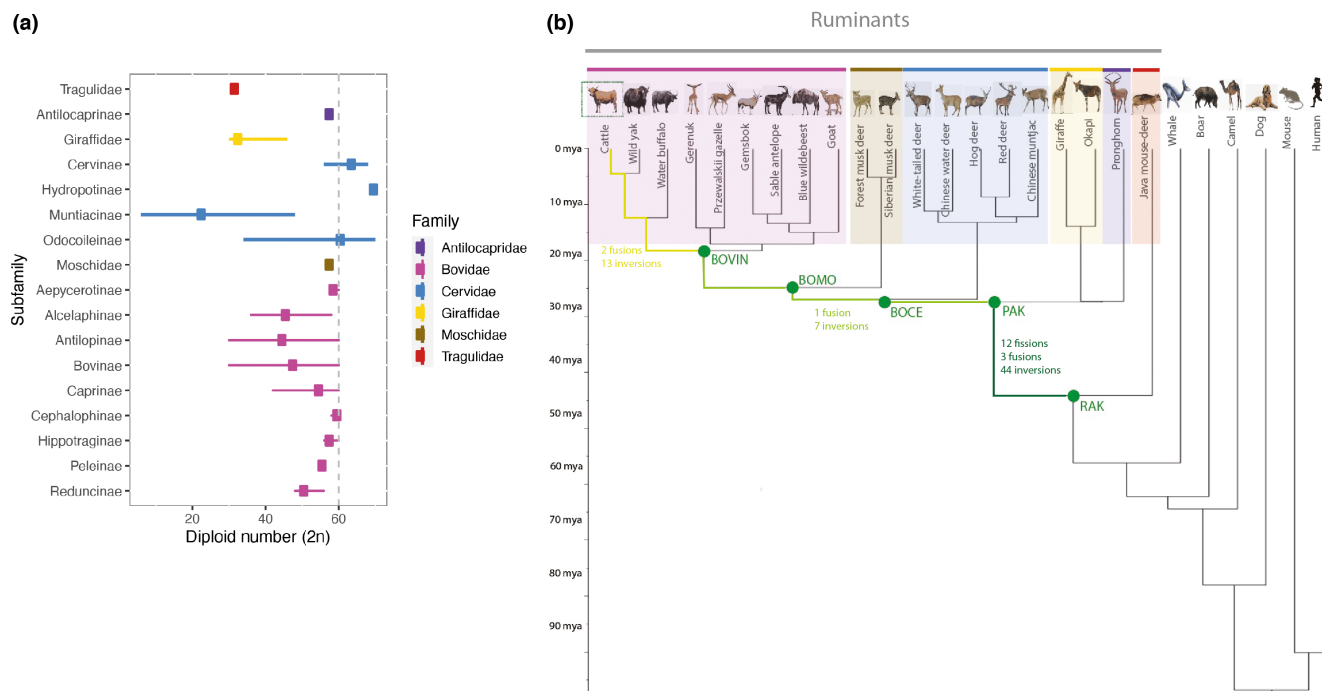
Ruminantia, the largest suborder within the Cetartiodactyla order of mammals, includes 234 species (Burgin et al., 2018; Mammal Diversity Database, 2022) in two infraorders, Tragulina and Pecora, that diverged ~50 million years ago (Fernández & Vrba, 2005; Zachos et al., 2001). Ruminant species belong to six families (Tragulidae, Antilocapridae, Giraffidae, Cervidae, Moschidae and Bovidae), with around 23% of species being endangered or critically endangered in the IUCN Red List (IUCN, 2022) (Table S1). Ruminants exhibit extreme morphological and ecological characteristics such as a multichambered stomach, cranial appendages (headgear) (Davis et al., 2011), specialized dentition, a highly cursorial locomotion and a wide range of body size variations. Ruminants are distributed across extensive habitats, including different altitudes, latitudes

and ecological environments (Wang et al., 2019). Alongside, ruminants show a wide variation in diploid chromosome numbers, from  $2n=6/7$  to  $2n=70$  in the Indian muntjac (*Muntiacus muntjak*) and white-tailed deer (*Odocoileus virginianus*), respectively, with most of the species having  $2n=60$  chromosomes, especially in the Bovidae family (Figure 1a) (Graphodatsky et al., 2020). This disparity in chromosome number is due to chromosome rearrangements and not changes in ploidy.

Chromosome rearrangements (CRs) can facilitate adaptation (Joron et al., 2011; Mérot et al., 2021), change gene expression patterns of nearby genes (Berdan et al., 2021; Farré et al., 2019; Lazar et al., 2018) and alter recombination (Farré et al., 2013; Navarro & Barton, 2003) which could eventually lead to speciation (reviewed in (Damas et al., 2021)). Comparative genomic studies in mammals and birds have shown that syntenic fragments (SFs, regions

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.



**FIGURE 1** Ruminant species. (a) Chromosome numbers (2n) of ruminant families and subfamilies. (b) Phylogenetic tree of the species included in the analysis. The different shades of green in the branches indicate the ancestral karyotypes reconstructed and the number and type of chromosome rearrangements between each ancestor. BOCE, ancestral karyotype of the Cervidae, Moschidae and Bovidae; BOMO, ancestral karyotype of Moschidae and Bovidae; BOVIN: bovid ancestral karyotype; PAK, pecoran ancestral karyotype; RAK, ruminant ancestral karyotype.

of the genomes with the same order of markers) and evolutionary breakpoint regions (EBRs, localized at the end of SFs) occurred in distinct genomic contexts (Farré et al., 2016; Larkin et al., 2009; Murphy et al., 2005). SFs are enriched for evolutionary conserved sequences and genes related to basic organismal development (Farré et al., 2016; Larkin et al., 2009); instead, EBRs are clustered in regions with transposable elements (Farré et al., 2011, 2019; Larkin et al., 2009) and collocate with genes related to lineage-specific biology (Farré et al., 2019; Kim et al., 2017). However, the evolutionary forces that lead to the appearance and fixation of CRs are still unclear.

CRs have been widely studied in ruminants, mostly using cytogenetic and genetic maps (Kulemzina et al., 2009, 2011; Slate et al., 2002). However, the resolution of these approaches is limited, hindering the identification of EBRs and SFs. Recently, with the implementation of DESCHRAMBLER (Kim et al., 2017), we reconstructed the ancestral karyotype of three ruminant ancestors using nine ruminant genome assemblies. We showed that ruminant karyotypes are populated with multiple inter- and intrachromosomal rearrangements that might affect gene expression and regulation in the lineage leading to cattle (Farré et al., 2019). Here, we expand on our previous work by generating five new in silico karyotype reconstructions using 20 ruminant genomes including at least one representative of each family (Figure 1b). This allows us to refine the identification and timing of CRs in ruminant evolution. Looking at EBRs and SFs we also point to possible mechanisms in the formation

and fixation of CRs. Our results significantly increase knowledge of ruminant genome evolution and will facilitate greater understanding of the role of chromosome rearrangements in adaptation and speciation in this lineage, as well as the possible mechanisms leading to their formation.

## 2 | MATERIALS AND METHODS

### 2.1 | Genome selection and filtering

Details of all publicly available genome assemblies, assembly quality and known diploid number in all extant ruminant species can be found in Table S1 (Dudchenko et al., 2017; Foley et al., 2023; Graphodatsky et al., 2020). A total of 20 species with genomes assembled at chromosome-level or with a scaffold N50 ≥ 1 Mbp, and representing all ruminant families and subfamilies, were included for further analysis (Table 1). Moreover, six outgroup species (human, mouse, dog, pig, camel and sperm whale) were also included in the analysis. All genome assemblies were filtered using faFilter from UCSC Kent Utilities (Kent et al., 2010; Kuhn et al., 2013) to remove all scaffolds smaller than 10,000 bp ensuring unplaced scaffolds were not included in the analysis. Genome assemblies obtained from NCBI (Table 1) were already soft-masked and no further filters were done, while genomes downloaded from DNA zoo were soft-masked using RepeatMasker version 4.0.9

TABLE 1 Genome assemblies used in this study.

Common name	Scientific name	Assembly ID/GeneBank accession number	Assembly level	Scaffold N50 (Mbp)	Size (Gb)	<sup>a</sup> % SFs
Pronghorn	<i>Antilocapra americana</i>	AntAmePen_v2_BIUU_UCD/GCA_004027515.2	Scaffold	18.8	2.96	94.63
Wild yak	<i>Bos mutus</i>	ASM764659v3/GCA_007646595.3	Scaffold	16.6	2.76	94.44
Water buffalo	<i>Bubalus bubalis</i>	UOA_WB_1/GCA_003121395.1	Scaffold	117.2	2.66	91.84
Goat	<i>Capra hircus</i>	ARS1/GCA_001704415.2	Chromosome	N/A	2.92	93.88
Blue wildebeest	<i>Connochaetes taurinus</i>	BWD_HiC/N/A	Chromosome	N/A	2.65	95.91
Sable antelope	<i>Hippotragus niger</i>	Sable_antelope_Masurca.scf_HiC/N/A	Chromosome	N/A	2.6	95.99
Gerenuk	<i>Litocranius walleri</i>	GRK_HiC/N/A	Chromosome	N/A	2.98	94.99
Gemsbok	<i>Oryx gazella</i>	UCDavis_Ogaz_1/GCA_003945745.1	Chromosome	N/A	2.74	84.57
Przewalski's gazelle	<i>Procapra przewalskii</i>	PLG/GCA_006410515.1	Scaffold	5.5	2.69	89.3
Hog deer	<i>Axis porcinus</i>	ASM379854v1_HiC/N/A	Chromosome	N/A	2.68	95.87
Red deer	<i>Cervus elaphus</i>	CerEla1.0/GCA_002197005.1	Chromosome	N/A	3.44	95.24
Chinese water deer	<i>Hydropotes inermis</i>	NPU_HINE_1.0/GCA_006459105.1	Scaffold	13.8	2.53	94.25
Chinese muntjac	<i>Muntiacus reevesi</i>	CIJ_HiC/N/A	Chromosome	N/A	2.6	95.74
White-tailed deer	<i>Odocoileus virginianus</i>	Ovir.te_1.0_HiC/N/A	Chromosome	N/A	2.38	95.67
Giraffe	<i>Giraffa camelopardalis</i>	Giraffe.scafSeq.fill.gapcloser/N/A	Chromosome	N/A	2.44	95.81
Okapi	<i>Okapia johnstoni</i>	ASM166083v1_HiC/N/A	Chromosome	N/A	2.89	95.59
Forest musk deer	<i>Moschus berezovskii</i>	Is35.final.genome_HiC/N/A	Chromosome	N/A	2.73	95.84
Siberian musk deer	<i>Moschus moschiferus</i>	MosMos_v2_BIUU_UCD/GCA_004024705.2	Scaffold	11.7	2.11	90.87
Java mouse-deer	<i>Tragulus javanicus</i>	ASM402496v2/GCA_004024965.2	Scaffold	14.1	2.59	92.95

<sup>a</sup>Coverage syntenic fragments (SFs) of the pair-wise alignments of each species to cattle genome.

(Tempel, 2012) with the same options for all species (-pa -64 -xsmall -species cattle).

## 2.2 | Establishing ancestral karyotypes

To define ancestral karyotypes, we used DESCHRAMBLER, a tool to reconstruct ancestral karyotypes (Kim et al., 2017). This tool takes as input a phylogenetic tree and pairwise alignments of the species included in the analysis.

### 2.2.1 | Pairwise alignments

Pairwise alignments were performed using the cattle genome (*Bos taurus*) as a reference. We used lastZ version 1.04.00 (Harris, 2007) with the following parameters: -minScore=1000, C=0, E=30, K=3000, L=3000, O=400. The output of lastZ was then converted into chain and net files using the following UCSC Kent Utilities: axtChain (parameters: -psl -verbose=0 -minScore=1000 -linearGap=medium), chainAntiRepeat, chainSort, chainPreNet chainNet and netSyntenic. Syntenic fragments (SFs) of 150-Kbp

and 300-Kbp resolution were represented in Evolution Highway (EH) images using syntenyPlotter (Farré et al., 2019). The coverage of the net files was calculated to assess reliability of the alignment (Table 1).

### 2.2.2 | Phylogenetic tree

A phylogenetic tree including all 20 ruminant and six outgroup species was obtained using TimeTree divergence time estimates (Kumar et al., 2017), written in newick format and plotted using FigTree (Rambaut, 2009).

### 2.2.3 | DESCHRAMBLER analysis

We started the reconstruction running DESCHRAMBLER (Kim et al., 2017) with a resolution of 300Kbp and minimum adjacency score of 0.01. Five different ancestors were reconstructed: ruminant ancestor (RAK), pecoran ancestor (PAK), Cervidae+Moschidae+Bovidae ancestor (BOCE), Moschidae+Bovidae ancestor (BOMO) and Bovidae ancestor

(BOVIN). The reconstructed ancestral chromosome fragments produced by DESCHRAMBLER were manually curated and merged to reconstruct ancestral chromosomes using previous studies as reference (Farré et al., 2019). For RAK and PAK reconstruction, the number of ancestral chromosome fragments reconstructed by DESCHRAMBLER was higher than the number of chromosomes previously proposed by these studies. Previous studies have shown that occasionally ancestral chromosomes are more fragmented than what has been suggested by other methodologies, in part because the use of scaffold assemblies of descendant species, where the exact tips of chromosomes are not known, and/or in part because of ambiguous cases resulting from insufficient evidence of adjacency (Kim et al., 2017). We manually curated our reconstructions by merging ancestral fragments following the structure of FISH-based reconstructions (Farré et al., 2019), and orientating of these merged ancestral fragments into ancestral chromosomes by minimizing the number of rearrangements. In some instances, ancestral chromosome fragments were not merged into larger ancestral chromosomes, due to the inability to manually resolve complex rearrangements. In these cases, ancestral chromosomes were formed by more than one fragment, and labelled accordingly to their size. Visualizations of ancestral karyotypes were made using syntenyPlotter (Farré et al., 2019).

### 2.3 | Detection of chromosome rearrangements (CRs) and evolutionary breakpoint regions (EBRs)

We classified CRs between ancestors into fusions, fissions, inversions and complex rearrangements (when more than one type of CR occurs). We then defined evolutionary breakpoint regions (EBRs) as the breaks of synteny at the edges of CRs in each of the ancestors. EBRs were then classified by size into well-defined EBRs (ranging from 0bp to 50Kbp) and not defined EBRs (ranging from 50Kbp to 300Kbp). When the break of synteny was >300Kbp, these regions were considered gaps and discarded from our analysis. The EBRs were then phylogenetically classified depending on the ancestral lineage in which they occurred into ancestor specific or reference species specific. Finally, EBRs were further separated by the type of rearrangement they delimited into inversion EBRs or interchromosomal EBRs, if they demarcated inversions or were the result of fusion or fission events respectively.

Syntenic fragments (SFs) forming CRs were classified as collinear, inversion, fusion or fission SFs depending on whether they have been conserved syntenic during evolution (collinear SFs) or are the result of any type of chromosome rearrangement that change their position during evolution (Figure S1).

### 2.4 | Identification of housekeeping genes

Normalized gene expression data in transcripts per million (TPM) were downloaded on 23/05/23 from the Gene Expression Atlas

(<https://www.ebi.ac.uk/gxa/home> (Moreno et al., 2022)) for cow (Merkin et al., 2012), pig (Eory, 2017) and sheep (Jiang et al., 2014), whereas gene expression data for human were downloaded from <https://gtexportal.org/home/datasets> GTEx release V8 (dbGap Accession phs000). Only tissues present in all four species were used for further analysis. We considered as housekeeping genes those genes expressed at a constant or stable level in almost all tissues of the dataset, following previous publications (Eisenberg & Levanon, 2013). First, we labelled genes as expressed if they have >10 TPM in over 90% of the tissues in each species (Papatheodorou et al., 2019). Then, we calculated whether genes were stably expressed across tissues using the nonparametric metric Gini coefficient, traditionally used in economics for determining income inequality, and previously used to assess expression inequality for transcriptomic data (Jiang et al., 2018). The Gini coefficient provides a value between 0 and 1 for each gene, representative of the expression equality of a given gene. The closer the value to 1 the greater the expression inequality becomes. We used the function Gini from D<sub>ES</sub>cTools v 0.99.31 (Signorell, 2023). Only genes with a Gini coefficient  $\leq 0.4$  were considered housekeeping genes.

To cross-compare housekeeping genes across species, we downloaded single copy orthologous genes for the four species from Ensembl (Cunningham et al., 2022). We considered mammalian ancestral housekeeping genes those orthologs that were labelled as housekeeping in all species, while ruminant ancestral housekeeping genes only those orthologs that were labelled as housekeeping in cattle and sheep only. Finally, any housekeeping genes in the cattle genome not identified as either mammalian or ruminant ancestral or housekeeping in any other species were named cattle-specific housekeeping genes. An UpSetR plot to visualize this relationship was created using UpSetR (version 1.4.0) in R (version 4.2.2).

### 2.5 | Statistical analysis

To assess the relationship between EBRs and SFs with genomic and transcriptomic features, we downloaded the unique protein coding genes and transposable elements from the cattle genome ARS-UCD1.2 using Ensembl BioMart. We calculated the fraction of GC content in windows of 10Kbp in the cattle genome using bedtools nuc (Quinlan & Hall, 2010). Windows with at least 60% GCs were labelled as high GC content. This cut off was chosen because it represents two times the 95% confidence interval of the mean % GC in 10Kbp windows of the genome. We then used *RegioneR* R package version 1.26 with 1000 permutations in each dataset to evaluate the overlap of these features, with a  $p$ -value  $\leq 0.05$  considered statistically significant.

We assessed the density of these features in regions surrounding the well-defined EBRs. First, we calculated the density of the genomic features (as number of bases of the feature) in non-overlapping 10Kbp windows in the cattle genome using bedtools coverage (Quinlan & Hall, 2010). Then 30 10Kbp windows up- and down-stream of each EBR midpoint were selected, and the mean

density of each window was calculated using a custom script. We finally plotted the density in a composite plot using ggplot2 (version 3.3.5) in R (version 4.1.0).

An additional analysis was performed to study the correlation of the EBRs and the TEs content of cattle. Genomic Association Test (GAT) version 1.3.4 (Heger et al., 2013) was used to compute the significance of overlap between EBRs and different families of TEs. A  $p$ -value  $\leq 0.05$  was considered statistically significant.

### 3 | RESULTS

To assess the functional contribution of chromosomal rearrangements to the evolution of ruminants, we first reconstructed ruminant ancestral karyotypes using 20 ruminant genomes including at least one representative of each ruminant family (Table 1, Figure 1). Five ruminant ancestral karyotypes were reconstructed: RAK, the deepest node, ancestral to all ruminants; PAK, the pecoran ancestral karyotype; BOCE, the ancestral karyotype that included Cervidae, Moschidae and Bovidae families; BOMO, the common ancestral karyotype including Moschidae and Bovidae families; and BOVIN, the shallowest node, ancestral of the Bovidae family.

#### 3.1 | Reconstructing ancestral karyotypes

We used DESCHRAMBLER to reconstruct the five ancestors with a resolution of 300 Kbp and an adjacency score of 0.01. All ancestral reconstructions showed high genome coverage of the reference genome (cattle), ranging from 88.6% to 93.93% in the deepest node RAK and the shallowest node BOVIN respectively. DESCHRAMBLER generated a total of 32 and 40 ancestral chromosome fragments for PAK and RAK respectively, while 30 ancestral chromosome fragments for BOCE, BOMO and BOVIN were generated. We then manually curated the ancestral chromosome fragments to increase coverage and minimize overestimation of CRs (see Methods). After the curation, a total of 27 ancestral chromosome fragments were obtained in RAK and 30 in the rest of ancestors (Table 2).

Most of the ancestral chromosome fragments represent entire ancestral chromosomes, except in six ancestral chromosomes (Figure 2). RAK, the deepest node, presents a haploid number of  $23+X$  within 27 fragments, with RAK10 and RAKX comprised of

two and three fragments each respectively. PAK consists of 30 ancestral chromosome fragments, with PAKX split into two fragments, showing the same haploid number ( $28+X$ ) that the Pecora ancestor reported previously (Farré et al., 2019). BOCE and BOMO contain 30 ancestral chromosome fragments with a haploid number of  $28+X$  in both ancestors, due to the fragmentation of BOCEX and BOMO11. Finally, BOVIN, the shallowest node, presents 30 ancestral chromosome fragments, with BOVIN24 fragmented into two pieces. Our BOVIN reconstruction presents a lower haploid number ( $28+X$ ) than previously reported ( $29+X$ ), due to BOVIN1 including the fusion of cattle chromosomes 9 and 14 (Farré et al., 2019).

#### 3.2 | Detection of chromosome rearrangements (CRs) and evolutionary breakpoint regions (EBRs) in ruminants

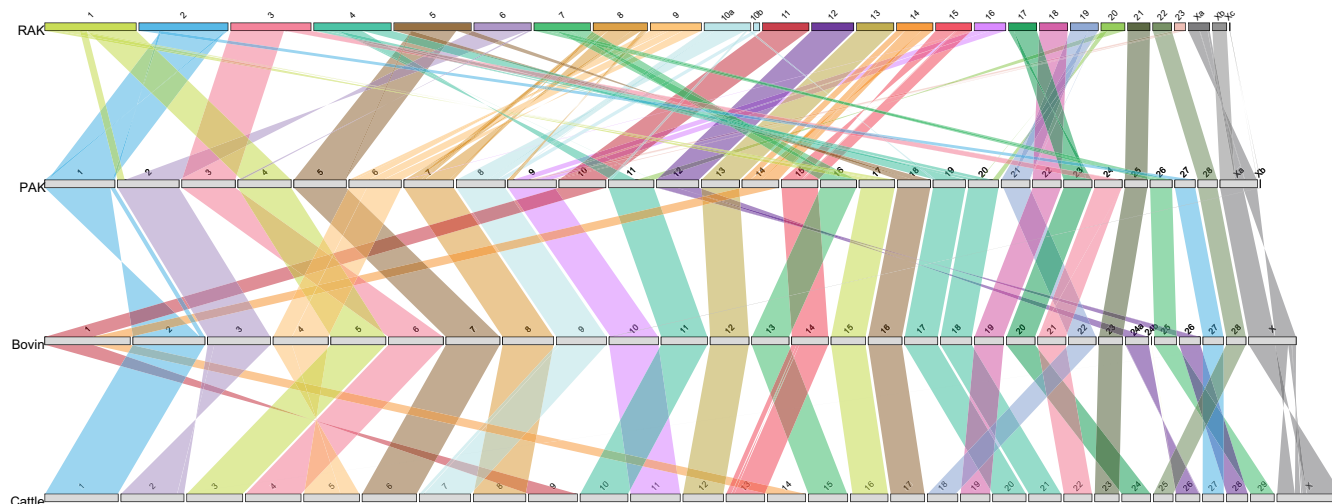
Ancestral karyotypes provide the basis to identify and date CRs that occurred during ruminant evolution. To this end, we cross-compared SFs in each ancestral reconstruction and used parsimony to delineate the CRs between RAK and PAK, PAK and BOVIN; and BOVIN and cattle (Figure 2 and Table S2). Only six chromosomes were maintained collinear between RAK and PAK (RAK 11, 12, 13, 18, 21 and 22), while a total of 59 CRs occurred, with 12 fissions, three fusions, and 44 inversions (Figure 2), representing a CR rate of 0.94 interchromosomal rearrangements/million years (My), 2.75 inversions/My, and a total of 3.69 CRs/My. Contrariwise, only eight CRs differentiated PAK from BOVIN, with one fusion and seven inversions (Figure 2), indicating that intrachromosomal rearrangements were the main structural change between PAK to BOVIN genome. Additionally, twenty-one ancestral chromosomes remained collinear between PAK and BOVIN (PAK chromosomes 2–5, 7, 8, 11, 13, 15–28). This translates to 0.1 and 0.7 interchromosomal rearrangements and inversions/My, respectively, and a total of 0.8 CRs/My between PAK and BOVIN. Finally, the youngest ancestral karyotype reconstructed, BOVIN (Figure 2), differs from cattle in two fusions and 13 inversions, with 0.11, 0.72 and 0.83 of interchromosomal, inversion and total CRs/My of evolution respectively. Overall, only four RAK chromosomes, homologous to cattle chromosomes BTA12, 19, 23 and 25, were maintained completely syntenic for 50 million years of evolution (Figures S2 and S3).

We then classified SFs in the cattle genome depending on the type of rearrangement that occurred since the divergence of RAK

TABLE 2 DESCHRAMBLER results. No. of chromosomes ( $n$ ) indicates the haploid number of our reconstructions.

Ancestor	No. RACFs	No. chrom ( $n$ )	Max. Size (kb)	Min. Size (kb)	Total size (Gb)	Ref. coverage <sup>a</sup>
RAK	27	$23+X$	192,591	1398	2.41	88.6
PAK	30	$28+X$	151,319	855	2.45	90.35
BOCE	30	$28+X$	151,973	1579	2.47	91.05
BOMO	30	$28+X$	154,317	7360	2.49	91.18
BOVIN	30	$28+X$	184,490	365	2.53	93.83

<sup>a</sup>Coverage against cattle genome (size = 2.71Gb). RACFs: reconstructed ancestral chromosome fragments.



**FIGURE 2** Comparison of reconstructed ancestral karyotypes. Colours indicate different chromosomes, with ribbons showing syntenic relationships. Inversions are shown when ribbons are twisted. BOVIN, bovid ancestral karyotype; PAK, pecoran ancestral karyotype; RAK, ruminant ancestral karyotype.

into: (i) collinear, when SFs have been conserved between species, (ii) inversion, (iii) fission, (iv) fusion or (v) a combination of some of the previous cases (Figure S1). A total of 48.84% of cattle genome has been maintained collinear, 8.77% inverted, and 40.48% into interchromosomal rearrangements (specifically, 17.74% fissioned and collinear SFs, 21.37% fissioned and fused SFs, and 1.67% fissioned and inverted SFs). These results highlight the important role that interchromosomal rearrangements have played in genome reshuffling during Ruminantia species evolution.

CRs are by definition delimited by evolutionary breakpoint regions (EBRs). Looking at all CRs placed in the reference genome context, we found a total of 56 EBRs in cattle (Table S3). To study their location and genomic characteristics, we first divided them into well-defined if their genomic size is 0–50Kbp or not defined if their size is 50–300 Kbp. A total of 32 and 24 were well defined and not defined EBRs respectively. The EBRs were then classified depending on whether they were cattle or ancestor specific. Within the 32 well defined EBRs, three were cattle-specific, three BOVIN-specific, two PAK-specific and 24 RAK-specific. Within the 24 not defined EBRs, five were cattle-specific, one PAK-specific and 18 RAK-specific. EBRs, including all types, are not uniformly distributed in cattle chromosomes (Figure S4). Cattle chromosome 11 and 21 showed a significantly higher density of EBRs (0.056, 0.057 EBRs/Mbp respectively) than the genomic mean of 0.021 EBRs/Mbp ( $Z$ -score 2.09,  $p$ -value = .03; and  $Z$ -score 2.16,  $p$ -value = .03, respectively, Figure S4). Contrariwise, no type of EBRs were located in cattle chromosomes 4, 6, 9, 12, 19, 23, 25 and 27.

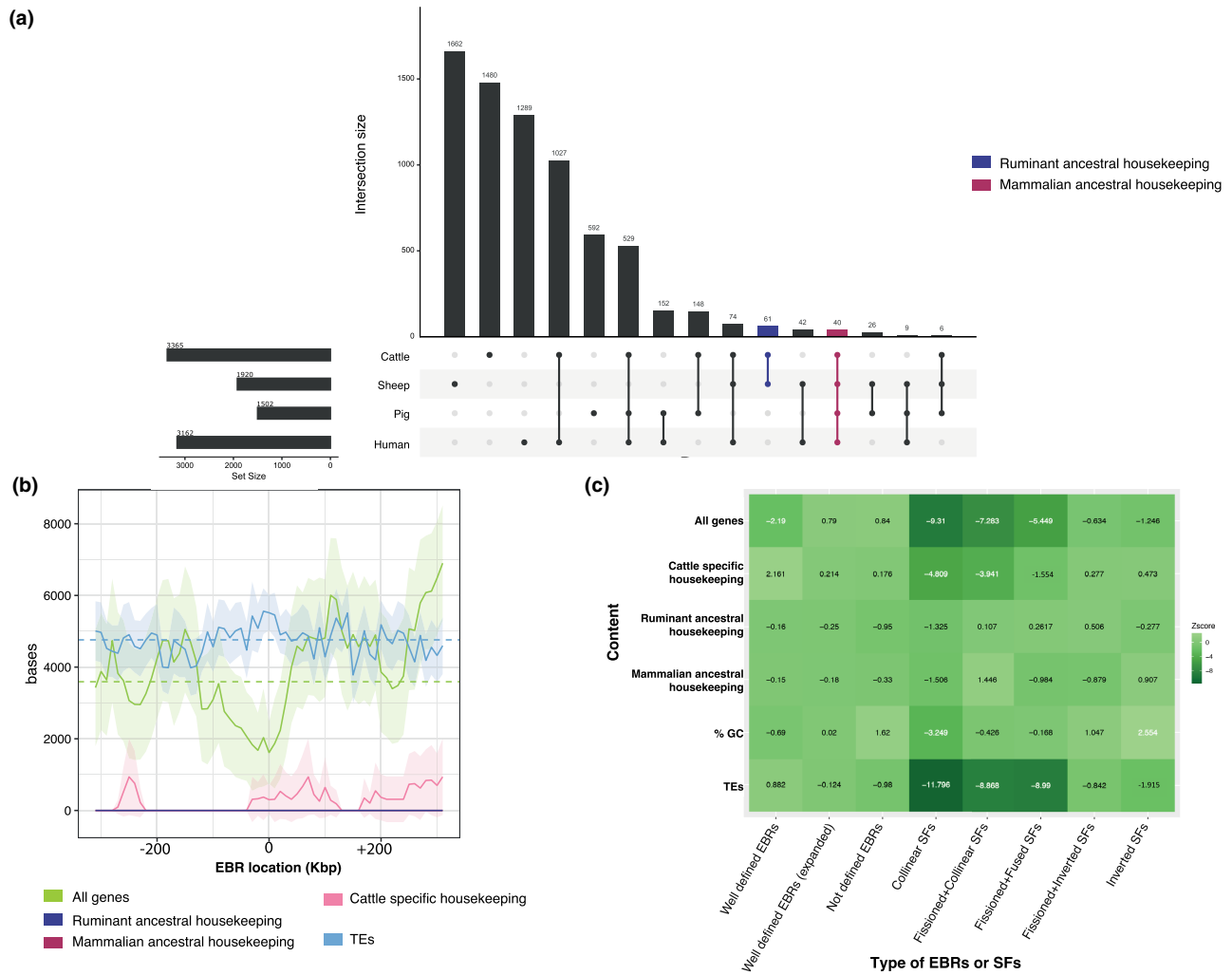
### 3.3 | Assessing the non-random distribution of CRs and EBRs

Seeing the pattern that EBRs and CRs are not uniformly distributed across ruminant chromosomes, we studied whether they occur in

non-random locations of the genome. We previously hypothesized that DNA sequence composition, the 3D structure of the chromatin within the nucleus and the effect on gene expression were key elements in determining the genomic distribution of EBRs, as part of the Integrative Breakage Model (Farré et al., 2015, 2019). Here, we tested this model by first looking at the genomic and genic content within and nearby EBRs. To do so, several permutation tests were performed over three different sets of EBRs: (i) well-defined EBRs; (ii) not defined EBRs; and (iii) regions near well-defined EBRs ( $\pm 30$ Kbp) (Figure 3). Only 8 out of 21,861 protein coding genes annotated in cattle are found within well-defined EBRs, representing a clear negative association between genes and well-defined EBRs ( $Z$ -score =  $-2.191$  and  $p$ -value = .0195, Figure 3b,c). In not defined EBRs, 294 protein-coding genes were found, while 48 genes were in surrounding regions of well-defined EBRs; however, these overlaps are not statistically significant ( $Z$ -score = .84 and  $p$ -value = .278;  $Z$ -score = .79 and  $p$ -value = .259, respectively, Figure 3b,c). Therefore, these results indicate that well defined EBRs do not co-localize with genes (Figure 3b).

Previous publications suggested that CRs might be the result of non-allelic homologous recombination between segmental duplications or transposable elements. Here, we used permutation testing to assess the overlap of EBRs and transposable elements. Although well-defined EBRs are not enriched in transposable elements overall ( $Z$ -score = .882,  $p$ -value = .259), several ruminant-specific transposable elements are significantly associated with EBRs, including ERV1, ERV2 and BovB (Figure S5).

Because CRs must start with an incorrectly repaired double strand break (DSB) and to be passed to the next generation this must happen in the germline, it has been suggested that EBRs might co-localize with crossovers in meiosis. Here, we used high GC content as a measure of past recombination levels linked to biased gene conversion to assess if EBRs are related to meiotic crossovers. A total of 2353 windows contained more than 60% of GC, representing 8.66%



**FIGURE 3** Association of evolutionary breakpoint regions (EBRs) with several genomic and transcriptomic data. (a) Overlap of housekeeping genes in four mammalian species. The number of housekeeping genes in each species is shown in the left barplot as 'set size'. The x-axis represents the number of genes being housekeeping for the different overlap combinations of species. Different combinations of overlap are represented by lines interlinking dots, while the bar chart on top shows the degree of overlap as 'intersection size'. In red, we show the mammalian ancestral housekeeping genes, while in blue the ruminant ancestral housekeeping genes. (b) Composite plot showing the density of genomic features in regions surrounding well defined EBRs. Dark coloured lines indicate the mean density (number of bases of each feature in 10Kbp windows) up- and down-stream of the midpoint of EBRs, while coloured shading shows the 95% confidence intervals. Dotted lines indicate the genome-wide mean of each feature. (c) Heatmap plotting the z-score of the statistical association of EBRs and SFs with genomic and transcriptomic data. Shading indicates the z-score value, while white numbers show significant associations ( $p$ -value  $\leq 0.05$ ).

of the cattle genome. Contrary to our expectation, EBRs are not associated with high GC windows in the cattle lineage (well-defined EBRs  $Z$ -score =  $-0.691$ ,  $p$ -value =  $.6214$ , not defined EBRs  $Z$ -score =  $1.62$ ,  $p$ -value =  $.124$ , Figure 3c); whereas collinear SFs are negatively associated with high GC content ( $Z$ -score =  $-3.249$ ,  $p$ -value =  $.003$ ) and inverted SFs are enriched ( $Z$ -score =  $2.554$ ,  $p$ -value =  $.023$ ).

The Integrative Breakage Model predicts that only those reorganizations not disturbing essential genes and/or gene expression would likely be the only ones fixed within populations. To test this expectation, we defined housekeeping genes in four species, including cattle, sheep, pig and human (Figure 3a). This was performed using publicly available gene expression data, filtered to include only tissues from organs found in the datasets for all species (brain, colon,

heart, kidney, liver, lung, skeletal muscle, spleen and testis) improving comparability of the results. Available samples were taken from various structures within the organs thus leading to variable tissue quantities across the species: nine, 13, 17, and 24 samples from cow, sheep, pig and human respectively. We then obtained single copy orthologous genes for all four species from Ensembl (Cunningham et al., 2022). We defined mammalian ancestral housekeeping genes, as those orthologs being housekeeping in all four species; or ruminant ancestral housekeeping, as those orthologs considered housekeeping in only cattle and sheep (Figure 3a). A total of 40 and 61 housekeeping genes were labelled as mammalian or ruminant ancestral, respectively, while 1480 were cattle-specific (Figure 3a and Table S4).

Overall, ~2% of all housekeeping genes co-located with EBRs while ~98% were within SFs (Table S5). Mammalian and ruminant ancestral housekeeping genes are not found in any type of EBRs, and only one and 30 cattle-specific housekeeping genes co-locate with well-defined and not defined EBRs, respectively, although not statistically significant (Figure 3c). Mammalian and ruminant ancestral housekeeping genes are mainly located in collinear SFs (52.5% and 37.7% respectively) and fissioned-fused SFs (27.5% and 36.06%, respectively, Tables S4 and S5). But because collinear and fissioned-fused SFs represent 48.84% and 21.37% of the cattle genome, these overlaps are not statistically significant (collinear SFs  $Z$ -score = -1.32,  $p$ -value = .16, and  $Z$ -score = -1.5,  $p$ -value = .13 for ruminant and mammalian ancestral genes, respectively, Figure 3c). Instead, 50.81% of cattle-specific housekeeping genes are located in collinear SFs; representing a significant depletion in these genomic regions ( $Z$ -score = -4.81,  $p$ -value = .001). Similarly, a negative association between cattle-specific housekeeping genes and fission+collinear SFs was found ( $Z$ -score = -3.94,  $p$ -value = .001, Figure 3c).

## 4 | DISCUSSION

Understanding the mechanisms and impact of genomic rearrangements has been an active area of research since the first comparative cytogenetic maps were described. Now, with the availability of chromosome-scale genome assemblies generated by large consortiums (Foley et al., 2023; Rhie et al., 2021), this area of research is again at the forefront of evolutionary biology. Here, we assessed the patterns of chromosome evolution in ruminants, a clade characterized by extreme differences in diploid numbers (Figure 1a). We first reconstructed five ancestral karyotypes using 26 genome assemblies: the ancestor of all ruminant families (RAK), the ancestor of Pecoran species (PAK), the ancestor of Cervidae, Moschidae and Bovidae families (BOCE), the ancestor of Moschidae and Bovidae families (BOMO), and the ancestor of Bovidae family, the shallowest node (BOVIN). After manual curation, RAK, the deepest node, comprised 23+X chromosomes with RAK10 and RAKX split into two and three fragments respectively (Figure 2). This fragmentation, also present in PAKX and BOVIN24, is most likely due to a lack of support from the outgroup species, making DESCHRAMBLER unable to join these fragments. Poor alignment with the outgroup species and/or misassemblies in ingroup genomes in these areas might also cause this issue. Our reconstructions mostly agree with *in silico* ancestral karyotypes previously published (Damas et al., 2022; Farré et al., 2019), except in the fusion of cattle chromosomes 9 and 14 that we reported in BOVIN. This is probably caused by a lower number of bovid species assembled at chromosome level used in previous publications.

These reconstructed ancestral karyotypes allowed us to identify CRs and EBRs between each ancestor leading to cattle. Our results indicate that chromosome evolution in ruminants is characterized by a first burst of chromosome rearrangements, with a high proportion of interchromosomal ones between RAK and PAK, followed by

a period of intrachromosomal reshuffling leading to a final stasis during the past 44 million years (Figures 1b and 2), in line with previous publications (Farré et al., 2019; Kulemzina et al., 2014). For a CR to happen, a DSB must be incorrectly repaired using an unexpected DNA template (reviewed in (Bursed et al., 2022)). The higher rate of CRs between RAK-PAK (3.69 CRs/My), compared to the most recent branches, might be explained by changes in the DNA repair machinery. As shown in parrots, the loss of genes involved in the repair of DSBs led to an increase of CRs in this species (Huang et al., 2022). Studies on the evolution of the DNA damage repair machinery in ruminants would help to disentangle this issue.

Even if changes in the DSB repair machinery might be linked to the increase in the CR rate, these might not be solely responsible for the type of CR. Instead, the 3D structure of the chromatin within the nucleus might play a role in the potential type of CRs because the template used for repair must be in close physical proximity of the DSB and the repair machinery (Álvarez-González, Arias-Sardá, et al., 2022). We recently showed that species whose chromosomes are individually compacted within the nucleus (forming the so-called chromosome territories) tend to present more interchromosomal rearrangements, while species whose centromeres are all clustered together present more intrachromosomal rearrangements (Álvarez-González, Arias-Sardá, et al., 2022). Interestingly, the two types of 3D structures are found within the ruminant clade, with most species showing chromosome territories except for Indian muntjac (Hoencamp et al., 2021), making it tempting to speculate that a change in the chromatin 3D structure within the nucleus might have occurred during ruminant evolution.

We observed that CRs and EBRs are not uniformly distributed in cattle chromosomes, with four chromosomes being maintained completely collinear for more than 50 million years (Figure 2). However, the number of EBRs might be underestimated, since EBRs occurring at the ends of cattle chromosomes cannot be detected in our analysis when using cattle as a reference.

EBRs are not associated to regions with high GC content (Figure 3b,c). High GC content is related with recombination hotspots in meiosis due to GC-biased gene conversion, in which double stranded breaks are preferentially repaired with GC alleles (Arbeithuber et al., 2015). As such, our results point to EBRs happening in open chromatin but not in the same areas as meiotic cross-overs. This trend agrees with our most recent findings in rodent species; where EBRs do not co-localize with meiotic DSBs but with post-meiotic DSBs in open chromatin regions during male spermatogenesis (Álvarez-González, Burden, et al., 2022; Burden et al., 2023).

By using transcriptomic data from gene expression atlas in two ruminants and two outgroups, we defined mammalian ancestral housekeeping genes as those orthologous genes labelled as housekeeping in all species, ruminant ancestral housekeeping genes when only present in the two ruminant species while those not shared with any other species were considered cattle-specific housekeeping genes (Figure 3a). We found that most housekeeping genes are not located in EBRs (Figure 3b,c, Table S5); although not statistically significant. The absence of association could represent a



real depletion of housekeeping genes within EBRs or expected by chance given EBRs account for a small proportion of the genome. However, data on carnivores and other mammals have shown that EBRs tend to occur in the boundaries of topologically associated domains suggesting that SFs could be considered regulatory blocks (Álvarez-González, Burden, et al., 2022; Corbo et al., 2022; Damas et al., 2022). Our findings, combined with previous publications showing changes in expression in genes nearby EBRs (Farré et al., 2019), suggest that only CRs and EBRs not disrupting functional and essential genes became fixed in evolution.

To conclude, we reconstructed the ancestral karyotypes of five ruminants, traced CRs and identified EBRs leading to cattle. Using this dataset, we found that EBRs are not associated with genes, nor housekeeping genes; EBRs are enriched in ruminant-specific transposable elements but not occurring in high GC-rich regions. Overall, the emerging picture from our analysis further supports the Integrative Breakage model (Farré et al., 2015) and starts to uncover the evolutionary forces leading to ruminant evolution. Our results open new avenues to investigate CRs in ruminants and their potential role in speciation.

#### AUTHOR CONTRIBUTIONS

M.F. conceived and devised the study. C.A.-S., S. Q. and M.F. designed and performed experiments and analysis. C.A.-S. and M.F. wrote the initial first draft of the manuscript with input from S. Q. All authors read and approved the final version of the manuscript.

#### FUNDING INFORMATION

We thank Harris Lewin for early access to some of the genome assemblies used herein, and Denis Larkin for early discussions on the data. C.A.-S. and S. Q. are funded by the GTA fellowship programme from the University of Kent. We thank the specialist High Performance Computing system provided by Information Services at the University of Kent.

#### CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

#### DATA AVAILABILITY STATEMENT

Data generated in this manuscript are included as supplementary files. Scripts used to analyse the data can be found at [https://github.com/Farre-lab/AriasSarda\\_ruminants](https://github.com/Farre-lab/AriasSarda_ruminants).

#### ORCID

Cristina Arias-Sardá  <https://orcid.org/0000-0002-4124-1463>

Sarah Quigley  <https://orcid.org/0000-0003-0482-3779>

Marta Farré  <https://orcid.org/0000-0001-9170-5767>

#### REFERENCES

- Álvarez-González, L., Arias-Sardá, C., Montes-España, L., Marín-Gual, L., Vara, C., Lister, N. C., Cuartero, Y., García, F., Deakin, J., Renfree, M. B., Robinson, T. J., Martí-Renom, M. A., Waters, P. D., Farré, M., & Ruiz-Herrera, A. (2022). Principles of 3D chromosome folding and evolutionary genome reshuffling in mammals. *Cell Reports*, 41(12), 111839. <https://doi.org/10.1016/j.celrep.2022.111839>
- Álvarez-González, L., Burden, F., Doddamani, D., Malinverni, R., Leach, E., Marín-García, C., Marín-Gual, L., Gubern, A., Vara, C., Paytuví-Gallart, A., Buschbeck, M., Ellis, P. J. I., Farré, M., & Ruiz-Herrera, A. (2022). 3D chromatin remodelling in the germ line modulates genome evolutionary plasticity. *Nature Communications*, 13(1), 2608. <https://doi.org/10.1038/s41467-022-30296-6>
- Arbeithuber, B., Betancourt, A. J., Ebner, T., & Tiemann-Boege, I. (2015). Crossovers are associated with mutation and biased gene conversion at recombination hotspots. *Proceedings of the National Academy of Sciences*, 112(7), 2109–2114. <https://doi.org/10.1073/pnas.1416622112>
- Berdan, E. L., Mérot, C., Pavia, H., Johannesson, K., Wellenreuther, M., & Butlin, R. K. (2021). A large chromosomal inversion shapes gene expression in seaweed flies (*Coelopa frigida*). *Evolution Letters*, 5(6), 607–624. <https://doi.org/10.1002/evl3.260>
- Burden, F., Ellis, P. J. I., & Farré, M. (2023). A shared 'vulnerability code' underpins varying sources of DNA damage throughout paternal germline transmission in mouse. *Nucleic Acids Research*, 51(5), 2319–2332. <https://doi.org/10.1093/nar/gkad089>
- Burgin, C. J., Colella, J. P., Kahn, P. L., & Upham, N. S. (2018). How many species of mammals are there? *Journal of Mammalogy*, 99(1), 1–14. <https://doi.org/10.1093/jmammal/gyx147>
- Bursed, B., Zamariolli, M., Bellucco, F. T., & Melaragno, M. I. (2022). Mechanisms of structural chromosomal rearrangement formation. *Molecular Cytogenetics*, 15(1), 23. <https://doi.org/10.1186/s13039-022-00600-6>
- Corbo, M., Damas, J., Bursell, M. G., & Lewin, H. A. (2022). Conservation of chromatin conformation in carnivores. *Proceedings of the National Academy of Sciences*, 119(9), e2120555119. <https://doi.org/10.1073/pnas.2120555119>
- Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., da Rin Fioretto, L., ... Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research*, 50(D1), D988–D995. <https://doi.org/10.1093/nar/gkab1049>
- Damas, J., Corbo, M., Kim, J., Turner-Maier, J., Farré, M., Larkin, D. M., Ryder, O. A., Steiner, C., Houck, M. L., Hall, S., Shiue, L., Thomas, S., Swale, T., Daly, M., Korlach, J., Uliano-Silva, M., Mazzoni, C. J., Birren, B. W., Genereux, D. P., ... Zhang, X. (2022). Evolution of the ancestral mammalian karyotype and syntenic regions. *Proceedings of the National Academy of Sciences*, 119(40), e2209139119. <https://doi.org/10.1073/pnas.2209139119>
- Damas, J., Corbo, M., & Lewin, H. A. (2021). Vertebrate chromosome evolution. *Annual Review of Animal Biosciences*, 9(1), 1–27. <https://doi.org/10.1146/annurev-animal-020518-114924>
- Davis, E. B., Brakora, K. A., & Lee, A. H. (2011). Evolution of ruminant headgear: A review. *Proceedings of the Royal Society B: Biological Sciences*, 278(1720), 2857–2865. <https://doi.org/10.1098/rspb.2011.0938>
- 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., & Aiden, E. L. (2017). De novo assembly of the *Aedes aegypti* genome using hi-C yields chromosome-length scaffolds. *Science*, 356(6333), 92–95. <https://doi.org/10.1126/science.aal3327>
- Eisenberg, E., & Levanon, E. Y. (2013). Human housekeeping genes, revisited. *Trends in Genetics*, 29(10), 569–574. <https://doi.org/10.1016/j.tig.2013.05.010>
- Eory, L. (2017). *Pig transcriptome and gene expression atlas*. BioStudies E-MTAB-5895. <https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-5895>
- Farré, M., Bosch, M., López-Giráldez, F., Ponsà, M., & Ruiz-Herrera, A. (2011). Assessing the role of tandem repeats in shaping the genomic

- architecture of great apes. *PLoS One*, 6(11), e27239. <https://doi.org/10.1371/journal.pone.0027239>
- Farré, M., Kim, J., Proskuryakova, A. A., Zhang, Y., Kulemzina, A. I., Li, Q., Zhou, Y., Xiong, Y., Johnson, J. L., Perelman, P. L., Johnson, W. E., Warren, W. C., Kukekova, A. V., Zhang, G., O'Brien, S. J., Ryder, O. A., Graphodatsky, A. S., Ma, J., Lewin, H. A., & Larkin, D. M. (2019). Evolution of gene regulation in ruminants differs between evolutionary breakpoint regions and homologous synteny blocks. *Genome Research*, 29(4), 576–589. <https://doi.org/10.1101/gr.239863.118>
- Farré, M., Micheletti, D., & Ruiz-Herrera, A. (2013). Recombination rates and genomic shuffling in human and chimpanzee—A new twist in the chromosomal speciation theory. *Molecular Biology and Evolution*, 30(4), 853–864. <https://doi.org/10.1093/molbev/mss272>
- Farré, M., Narayan, J., Slavov, G. T., Damas, J., Auvil, L., Li, C., Jarvis, E. D., Burt, D. W., Griffin, D. K., & Larkin, D. M. (2016). Novel insights into chromosome evolution in birds, archosaurs, and reptiles. *Genome Biology and Evolution*, 8(8), 2442–2451. <https://doi.org/10.1093/gbe/evw166>
- Farré, M., Robinson, T. J., & 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(5), 479–488. <https://doi.org/10.1002/bies.201400174>
- Fernández, M. H., & Vrba, E. S. (2005). A complete estimate of the phylogenetic relationships in Ruminantia: A dated species-level super-tree of the extant ruminants. *Biological Reviews*, 80(2), 269–302. <https://doi.org/10.1017/S1464793104006670>
- Foley, N. M., Mason, V. C., Harris, A. J., Bredemeyer, K. R., Damas, J., Lewin, H. A., Eizirik, E., Gates, J., Karlsson, E. K., Lindblad-Toh, K., Zoonomia Consortium, Springer, M. S., Murphy, W. J., Andrews, G., Armstrong, J. C., Bianchi, M., Birren, B. W., Bredemeyer, K. R., Breit, A. M., ... Zhang, X. (2023). A genomic timescale for placental mammal evolution. *Science*, 380(6643), eabl8189. <https://doi.org/10.1126/science.abl8189>
- Graphodatsky, A. S., Perelman, P. L., & O'Brien, S. J. (2020). *Atlas of mammalian chromosomes*. John Wiley & Sons. <https://doi.org/10.1002/9781119418061>
- Harris, R. S. (2007). *Improved pairwise alignment of genomic DNA*. The Pennsylvania State University.
- Heger, A., Webber, C., Goodson, M., Ponting, C. P., & Lunter, G. (2013). GAT: A simulation framework for testing the association of genomic intervals. *Bioinformatics*, 29(16), 2046–2048. <https://doi.org/10.1093/bioinformatics/btt343>
- Hoencamp, C., Dudchenko, O., Elbatsh, A. M. O., Brahmachari, S., & Raaijmakers, J. A. (2021). 3D genomics across the tree of life identifies condensin II as a determinant of architecture type. *Science*, 372(6545), 984–989.
- Huang, Z., De, O., Furo, I., Liu, J., Peona, V., Gomes, A. J. B., Cen, W., Huang, H., Zhang, Y., Chen, D., Xue, T., Zhang, Q., Yue, Z., Wang, Q., Yu, L., Chen, Y., Suh, A., de Oliveira, E. H. C., & Xu, L. (2022). Recurrent chromosome reshuffling and the evolution of neo-sex chromosomes in parrots. *Nature Communications*, 13(1), 944. <https://doi.org/10.1038/s41467-022-28585-1>
- IUCN. (2022). *The IUCN red list of threatened species*.
- Jiang, L., Tsoucas, D., & Yuan, G.-C. (2018). Assessing inequality in transcriptomic data. *Cell Systems*, 6(2), 149–150. <https://doi.org/10.1016/j.cels.2018.02.007>
- Jiang, Y., Xie, M., Chen, W., Talbot, R., Maddox, J. F., Faraut, T., Wu, C., Muzny, D. M., Li, Y., Zhang, W., Stanton, J.-A., Brauning, R., Barris, W. C., Hourlier, T., Aken, B. L., Searle, S. M. J., Adelson, D. L., Bian, C., Cam, G. R., ... Dalrymple, B. P. (2014). The sheep genome illuminates biology of the rumen and lipid metabolism. *Science*, 344(6188), 1168–1173. <https://doi.org/10.1126/science.1252806>
- Joron, M., Frezal, L., Jones, R. T., Chamberlain, N. L., Lee, S. F., Haag, C. R., Whibley, A., Becuwe, M., Baxter, S. W., Ferguson, L., Wilkinson, P. A., Salazar, C., Davidson, C., Clark, R., Quail, M. A., Beasley, H., Glithero, R., Lloyd, C., Sims, S., ... Richard, H. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature*, 477(7363), 203–206. <https://doi.org/10.1038/nature10341>
- Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S., & Karolchik, D. (2010). BigWig and BigBed: Enabling browsing of large distributed datasets. *Bioinformatics*, 26(17), 2204–2207. <https://doi.org/10.1093/bioinformatics/btq351>
- Kim, J., Farré, M., Auvil, L., Capitanu, B., Larkin, D. M., Ma, J., & Lewin, H. A. (2017). Reconstruction and evolutionary history of eutherian chromosomes. *Proceedings of the National Academy of Sciences*, 114(27), E5379–E5388. <https://doi.org/10.1073/pnas.1702012114>
- Kuhn, R. M., Haussler, D., & Kent, W. J. (2013). The UCSC genome browser and associated tools. *Briefings in Bioinformatics*, 14(2), 144–161. <https://doi.org/10.1093/bib/bbs038>
- Kulemzina, A. I., Perelman, P. L., Grafodatskaya, D. A., Nguyen, T. T., Thompson, M., Roelke-Parker, M. E., & Graphodatsky, A. S. (2014). Comparative chromosome painting of pronghorn (*Antilocapra americana*) and saola (*Pseudoryx nghetinhensis*) karyotypes with human and dromedary camel probes. *BMC Genetics*, 15(1), 68. <https://doi.org/10.1186/1471-2156-15-68>
- Kulemzina, A. I., Trifonov, V. A., Perelman, P. L., Rubtsova, N. V., Volobuev, V., Ferguson-Smith, M. A., Stanyon, R., Yang, F., & Graphodatsky, A. S. (2009). Cross-species chromosome painting in Cetartiodactyla: Reconstructing the karyotype evolution in key phylogenetic lineages. *Chromosome Research*, 17(3), 419–436. <https://doi.org/10.1007/s10577-009-9032-3>
- Kulemzina, A. I., Yang, F., Trifonov, V. A., Ryder, O. A., Ferguson-Smith, M. A., & Graphodatsky, A. S. (2011). Chromosome painting in Tragulidae facilitates the reconstruction of Ruminantia ancestral karyotype. *Chromosome Research*, 19(4), 531–539. <https://doi.org/10.1007/s10577-011-9201-z>
- Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A resource for timelines, Timetrees, and divergence times. *Molecular Biology and Evolution*, 34(7), 1812–1819. <https://doi.org/10.1093/molbev/msx116>
- Larkin, D. M., Pape, G., Donthu, R., Auvil, L., Welge, M., & Lewin, H. A. (2009). Breakpoint regions and homologous synteny blocks in chromosomes have different evolutionary histories. *Genome Research*, 19(5), 770–777. <https://doi.org/10.1101/gr.086546.108>
- Lazar, N. H., Nevenon, K. A., O'Connell, B., McCann, C., O'Neill, R. J., Green, R. E., Meyer, T. J., Okhovat, M., & Carbone, L. (2018). Epigenetic maintenance of topological domains in the highly rearranged gibbon genome. *Genome Research*, 28(7), 983–997. <https://doi.org/10.1101/gr.233874.117>
- Mammal Diversity Database. (2022). *Mammal diversity database (1.9)*.
- Merkin, J., Russell, C., Chen, P., & Burge, C. B. (2012). Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science*, 338(6114), 1593–1599. <https://doi.org/10.1126/science.1228186>
- Mérot, C., Berdan, E. L., Cayuela, H., Djambazian, H., Ferchaud, A.-L., Laporte, M., Normandeau, E., Ragoussis, J., Wellenreuther, M., & Bernatchez, L. (2021). Locally adaptive inversions modulate genetic variation at different geographic scales in a seaweed fly. *Molecular Biology and Evolution*, 38(9), 3953–3971. <https://doi.org/10.1093/molbev/msab143>
- Moreno, P., Fexova, S., George, N., Manning, J. R., Miao, Z., Mohammed, S., Muñoz-Pomer, A., Fullgrabe, A., Bi, Y., Bush, N., Iqbal, H., Kumbham, U., Solovyev, A., Zhao, L., Prakash, A., García-Seisdedos, D., Kundu, D. J., Wang, S., Walzer, M., ... Papatheodorou, I. (2022). Expression atlas update: Gene and protein expression in multiple species. *Nucleic Acids Research*, 50(D1), D129–D140. <https://doi.org/10.1093/NAR/GKAB1030>
- Murphy, W. J., Larkin, D. M., der Wind, A. E., Bourque, G., Tesler, G., Auvil, L., Beever, J. E., Chowdhary, B. P., Galibert, F., Gatzke, L.,

- Hitte, C., Meyers, S. N., Milan, D., Ostrander, E. A., Pape, G., Parker, H. G., Raudsepp, T., Rogatcheva, M. B., Schook, L. B., ... Lewin, H. A. (2005). Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. *Science*, *309*(5734), 613–617. <https://doi.org/10.1126/science.1111387>
- Navarro, A., & Barton, N. H. (2003). Chromosomal speciation and molecular divergence—Accelerated evolution in rearranged chromosomes. *Science*, *300*(5617), 321–324. <https://doi.org/10.1126/science.1080600>
- Papatheodorou, I., Moreno, P., Manning, J., Fuentes, A. M.-P., George, N., Fexova, S., Fonseca, N. A., Füllgrabe, A., Green, M., Huang, N., Huerta, L., Iqbal, H., Jianu, M., Mohammed, S., Zhao, L., Jarnuczak, A. F., Jupp, S., Marioni, J., Meyer, K., ... Brazma, A. (2019). Expression atlas update: From tissues to single cells. *Nucleic Acids Research*, *48*, gkz947. <https://doi.org/10.1093/nar/gkz947>
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- Rambaut, A. (2009). *FigTree, a graphical viewer of phylogenetic trees*. University of Edinburgh.
- Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., Lee, C., Ko, B. J., Chaisson, M., Gedman, G. L., Cantin, L. J., Thibaud-Nissen, F., Haggerty, L., Bista, I., Smith, M., ... Jarvis, E. D. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, *592*(7856), 737–746. <https://doi.org/10.1038/s41586-021-03451-0>
- Signorell, A. (2023). *Tools for Descriptive Statistics*. [R package DescTools version 0.99.48]. Comprehensive R Archive Network (CRAN). <https://CRAN.R-Project.Org/Package=DescTools>
- Slate, J., van Stijn, T. C., Anderson, R. M., McEwan, K. M., Maqbool, N. J., Mathias, H. C., Bixley, M. J., Stevens, D. R., Molenaar, A. J., Beever, J. E., Galloway, S. M., & Tate, M. L. (2002). A deer (subfamily Cervinae) genetic linkage map and the evolution of ruminant genomes. *Genetics*, *160*(4), 1587–1597. <https://doi.org/10.1093/genetics/160.4.1587>
- Tempel, S. (2012). Using and understanding RepeatMasker. In Y. Bigot (Ed.), *Mobile Genetic Elements*, 859 (pp. 29–51). Humana Press. [https://doi.org/10.1007/978-1-61779-603-6\\_2](https://doi.org/10.1007/978-1-61779-603-6_2)
- Wang, B., Chen, L., & Wang, W. (2019). Genomic insights into ruminant evolution: From past to future prospects. *Zoological Research*, *40*(6), 476–487. <https://doi.org/10.24272/j.issn.2095-8137.2019.061>
- Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K. (2001). Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, *292*(5517), 686–693. <https://doi.org/10.1126/science.1059412>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Arias-Sardá, C., Quigley, S., & Farré, M. (2023). Patterns of chromosome evolution in ruminants. *Molecular Ecology*, *00*, 1–11. <https://doi.org/10.1111/mec.17197>