Supplementary Materials for

Sponge genomes reveal a pre-metazoan origin of the sex determination toolkit and sex chromosomes

Jose M. Lorente-Sorolla[1], Aida Verdes[1], Irene de Sosa[1], Carles Galià-Camps[1], Vanina Tonzo[1], Cristina Díez-Vives[2,3], Patricia Álvarez-Campos[4], Carlos Leiva[5], María Conejero[1], Marta Turon[1], Sami El Hilali[6], Sergi Taboada[1,3], Maria Eleonora Rossi[7], Marta Álvarez-Presas[8], Juan A. Hermoso[9], Jordi Paps[10], Xavier Grau-Bové[11], Ute Henstchel[12], Vasiliki Koutsouveli[1,12], Alison E Wright[13], Ana Riesgo[1,3*]

Corresponding author. Email: anariesgogil@mncn.csic.es

The PDF file includes:

Materials and Methods

1. Sampling, assessment of reproductive activity and sex assignment: Samples of eight gonochoristic species were collected from the North Atlantic Ocean between 2017-2019, the Mediterranean Sea in the summer of 2021 and 2022, and the Kiel Bay in summer of 2021 (Table S1, Supplementary Fig.1). We collected approximately 3-5 cm^3 of sponge tissue per specimen and divided the sample in 4 pieces that were later preserved in three different preservatives: 1. For traditional histology, in 4% formaldehyde in seawater, 2. For further construction of DNA libraries (for both RADseq and WGS), in absolute ethanol and stored at -20 °C, and 3. For transcriptomic analyses, a tissue piece of about 2 cm^3 was preserved in RNAlater at 4ºC for 24 h and then frozen at -20 °C until further processing.

Since sponge sex is only possible to determine through presence of gametes, we processed the samples for histology. Samples of sponges with siliceous spicule content (*Petrosia ficiformis*, *Geodia hentscheli, Geodia barretti*, *Axinella damicornis*, *Halichondria panicea*, and *Phakellia ventilabrum*) went through a step of desilicification in 5% hydrofluoric acid overnight and then rinsed with distilled water at least twice. Then, their tissues and those of the sponges without spicules (*Chondrosia reniformis* and *Oscarella lobularis*) were processed for light microscopy Tissues were dehydrated through an increasing ethanol series and later embedded in paraffin after a brief rinse in xylene. Then, paraffin blocks were sectioned at 5 μm with an HM 325 rotary microtome (ThermoFisher-Scientific) and sections stained with hematoxylin and eosin, using standard protocols, and mounted in slides with DPX. Slides were observed with an Olympus microscope (BX43) with a UC50 camera at the Museo Nacional de Ciencias Naturales de Madrid (MNCN-CSIC). Samples with oocytes were coded as females and samples with any spermatogenic stage as males. We did not find any case of hermaphroditism among our samples.

2. DNA and RNA extraction, RADseq and RNAseq library preparation: DNA was extracted from all samples using the DNeasy Blood & Tissue kit (Qiagen) following the manufacturer's protocol, except for the cell lysis time which was conducted overnight. Double-stranded DNA was quantified with Qubit dsDNA HS assay (Life Technologies). For RADseq library preparation we followed a protocol from Peterson *et al.* (2012)[1] with modifications described in Taboada *et al.* (2022)[2]. RNA was extracted separately from all samples using the Invitrogen RNA mini kit (Thermofisher) following the manufacturer's protocol using TRIzol for cell lysis and quantified with NanoDrop. Further, mRNA libraries were constructed using the Illumina Stranded mRNA Prep kit, quantified using Qubit dsDNA HS assay (Life Technologies) and checked for size and quality using a TapeStation 2200 (Agilent Technologies, USA). They were then pooled and sequenced using paired-end 150-bp reads on an Illumina NovaSeq6000 at Novogene Europe (Cambridge, UK). Our final dataset of genomic and transcriptomic resources for assessing sex determination in sponges can be found in Table S2.

3. WGS library preparation: DNA extracted as above from tissues of the sponges (see Supplementary Table 2) and prepared to obtain WGS libraries. The genomic DNA was randomly sheared into short fragments using enzymes, and the obtained fragments were end-repaired, A-tailed, and further ligated with an Illumina adapter. The fragments with adapters were PCR amplified, size selected, and purified. The library was checked with Qubit 3.0 and real-time PCR for quantification, and Bioanalyzer for size distribution detection. Quantified libraries were pooled and sequenced on the Illumina platform NovaseqX plus, according to effective library concentration and data amount required (6 Gb/sample), done at Novogene.

4 Genome annotation: The chromosome-level genomes of *Oscarella lobularis*, *Chondrosia reniformis*, *Petrosia ficiformis, Axinella damicornis,* and *Phakellia ventilabrum, Halichondria panicea* were sequenced by the Aquatic Symbiosis Genome Consortium (Table S2). We annotated the genes in the assemblies of *O. lobularis*, *C. reniformis* and *P. ficiformis* using a combination of tools for *de novo* and evidence-based gene prediction (BRAKER2 2.1.6[3], Augustus 3.5.0[4,5], StringTie 2.2.1[6], and GenomeThreader 1.7.1[7]) and optimal gene model selection (Mikado 2.3.4[8]). This procedure is described below. First, we mapped bulk paired RNA-seq libraries to the reference genome using the read aligner STAR 2.7.10b[9] without multi-mapping reads (flag: --outFilterMultimapNmax 1), only considering uniquely mapping reads for splice junctions (--outSJfilterReads Unique), reporting splice junction-supporting reads and keeping only the reads with junctions that passed filtering (--outFilterType BySJout, --alignSJDBoverhangMin 1, --alignSJoverhangMin 8), and reporting the alignment strand based on intron motifs (--outSAMstrandField intronMotif). The resulting coordinate-sorted BAM file (--outSAMtype BAM SortedByCoordinate) was used to produce an initial set of transcript predictions using StringTie in conservative mode (-t -c 1.5 -f 0.0 flags), from which open reading frames (ORFs) were predicted using TransDecoder 5.7.1[10]. Predicted peptides were collapsed by sequence similarity using CD-HIT 4.8.1 (-c 0.95) and complete genes (i.e. with start and end codons) of non-extreme lengths (>600 and <10,000 amino-acids) were retrieved for later use in the *de novo* gene prediction step. Specifically, these were used to train Augustus iteratively within BRAKER2, by aligning them to the reference genome using GenomeThreader (BRAKER2 flag: --prg=gth –trainFromGth), and using the original STAR-produced alignments as further evidence (--bam=<file>). Second, we used Mikado to select the best gene predictions from each locus, selecting from the output of BRAKER2/Augustus (all exons and coding exons [CDS] were used separately), an unguided Stringtie assembly, and the filtered set of GenomeThreader training peptide alignments to the reference genome. To build the Mikado hints file, (i) all sources of evidence were considered as strand-specific; (ii) a score of 1 was associated with the BRAKER2/Augustus predictions and 0 for the others; (iii) the CDS from BRAKER2/Augustus were considered as a reference annotation; (iv) redundant models were excluded from all samples; and (v) CDS sequences with errors were removed from the model set. To build the Mikado configuration file (mikado conFig), we clustered transcripts with a minimum cDNA overlap of 20% (--min-clustering-cdna-overlap 0.2) and any for CDSs (--min-clustering-cds-overlap 0.0), set the programme to permissive mode with regards to ORF splitting policy (--mode permissive). After preparing the transcripts sets with the mikado prepare submodule, we prepared further evidence sources to be considered by Mikado: (i) predicted ORFs for all Mikado models, using TransDecoder; (ii) evidence-based splice junction coordinates from STAR (see above), obtained using the junctools convert 1.2.4 module of Portcullis 1.1.2.; and (iii) homology models obtained with diamond blastx[11] against the 2022-05 release of UniRef50[12], adding the percentage of positive-scoring alignments and the traceback operations fields to the reported output (flag: -f 6 qseqid sseqid pident length mismatch gapopen qstart qend sstart send evalue bitscore ppos btop). These additional sources of evidence were considered by Mikado (mikado serialise submodule, disabling start codon adjustment with –no-start-adjustment). The best gene model for each locus was selected using the mikado pick module, prioritising reference models (--reference-update). Finally, transcript and peptide sequences for each gene model were retrieved using gffread 0.11.7[13]. The completeness of the resulting gene set was evaluated using BUSCO 5.5.0[14] in protein mode (-m proteins) against the Metazoa database of universal orthologs (-l metazoa_odb10).

Annotation of *P. ventilabrum, Axinella damicornis,* and *Halichondria panicea* was performed elsewhere, and the details can be found in the references in Table S2.

5. Repeatome analysis: The density of transposable elements (TEs) per chromosome was calculated in 100 Kbp windows for each species with chromosome-level reference genome (*C. reniformis* and *O. lobularis*). First, the Repeatmodeller[15] output was filtered to only include TEs, removing simple repeats, low-complexity regions, rRNAs, and unknown repeats. Subsequently, bedtools v2.30.0[16] was used to create a bed file with 100 Kbp windows for each reference genome using the options "makewindows -w 100000", and to calculate the density per window using the options "intersect -a genome_windows.bed -b ALL_TEs.bed -c". Finally, TE distribution plots per chromosome were created in R using the ggplot2 package v3.5.1[17]. Additionally, we repeated these analyses exclusively including long interspersed nuclear elements (LINEs), which were extracted from the Repeatmodeller output and plotted as LINE density per Mbp.

6. Sex chromosome identification: To identify putative sex chromosomes and sex determining regions (SDR) we employed several complementary approaches. First, we ran the FindZX[18] pipeline with default parameters for detecting and visualising sex chromosomes based on genome coverage and heterozygosity differences. The strength and nature of sex-specific genomic signals depend on the degree of sex chromosome differentiation and Y/W degeneration. In homomorphic systems, heterozygosity provides the clearest signature, whereas in heteromorphic systems, coverage differences are most informative, expected to be approximately the double on the X/Z chromosome in the homogametic sex compared to the heterogametic sex. Restricting the number of allowed mismatches enhances detection by preventing reads from the sex-limited chromosome from aligning to its gametolog, and heterozygosity can appear higher in either the heterogametic or homogametic sex depending on whether reads from the sex-limited chromosome successfully align or not. The analyses were conducted using WGS reads in 50kb, 100kb and 500kb windows, unfiltered, 0.2 and 0.0 mismatches allowed and their species reference genome in the no-synteny mode. We also calculated differences in coverage with mosdepth [v0.3.2][19], SNP density relative to the reference and F_{ST} values between sexes.SNP density and F_{ST} values are expected to be higher in the heterogametic sex chromosomes due to the differentiation and reduced recombination. k-mer–based analyses provide a complementary and highly sensitive approach, as they are independent of a reference genome and can detect sex-linked differences even when high sequence homology allows equal read mapping across sexes and at the same time can recover signals that may be overlooked by coverage- or SNP-based approaches when read mapping differs between sexes. WGS reads of each individual and also in male and female datasets were aligned to the reference chromosome-level genome with bowtie2[20] and indexed with samtools [v1.14][21] . SNP calling was done with bcftools [v1.12][21] and calls filtered with vcftools [v0.1.16][22] to only keep biallelic sites that pass the bcf_filter. Male/Female mapped read files were used to perform a coverage analysis after normalization. The analysis of the difference in coverage was performed with mosdepth [v0.3.2] in 1 Mb windows. SNP density relative to the reference genome in 100kb windows was calculated with vcftools and differences between male and female with the R script SNPdensity_permutations.R from the SexFindR pipeline[23] that calculates true means and difference and carries out a permutation test that determines p-values for SNP density differences windows. Kernel-smoothed F_{ST} was calculated with the *populations* program of Stacks from the RADseq reads. K-mer analyses were performed following the SexFindR pipeline. kmersGWAS[24] was used for k-mer counting, and PLINK[25] for sex-association testing of k-mers with p-value < 0.00001. Significant k-mers were subsequently assembled into short contigs using ABySS[26],

which were then aligned to the reference genome with BLAST to determine their genomic positions

7. Intraspecific inversion detection in sex-chromosomes: To evaluate the presence of putative inversions in the sex-determining gene regions, we conducted a multi-evidence approach based on (i) mapping alternative haplotypes to the reference assemblies, (ii) genotypes and (iii) mapping quality for those species with male and female specimens. Since the reference genome for both *C. reniformis* and *O. lobularis* derived from the heterogametic sex, we mapped the alternative haplotypes to the reference assemblies using minimap2[27] and visualized the dotplots in D-GENIES[28]. For the genotype approach, we used the R script "individual Detection of linkage by Genotyping (iDlG)"[29], which identifies linked regions across chromosomes at the individual level based on mean genotype calculated in sliding windows. This makes it possible to not only identify inversions but other regions with high linkage such as sex chromosomes. We established windows of 5,000 SNPs with sliding steps of 1,000 SNPs, and graphically visualized the iDlG results with the R package "ggplot2". To further support the presence of inversions in those regions, we kept from each specimen "sam" file those paired-reads mapping colinearly[30] (forward–forward; reverse–reverse) in the same chromosome. From those, we retained only those pairs mapping at a distance above the expected from library preparation (400 bp), and subsequently visualized the relative position of both collinear reads with "ggplot2".

8. Detection of sex-specific loci using RADseq data: RADseq is a crucial tool to study the genetic basis of sex determination in species without heteromorphic sex chromosomes[31]. Although RADseq alone can identify sex chromosome systems, it cannot determine sex chromosome linkage or the size of the non-recombining region. However, when used with a reference genome, sex-specific loci can be mapped to identify linkage groups and sex determining regions. RADseq data also allows us to calculate SNP density, and F_{ST} between males and females across the genome, both predicted to be higher in the sex-determining, non-recombining regions of sex chromosomes. These indicators are expected to converge in the sex determining regions of the sex chromosomes. To this end, we used the RADsex pipeline [v1.2][32]. The quality of raw reads was assessed with FastQC before and after demultiplexing with the process_radtags module of Stacks [v2.61][33]. In addition, the detection and mapping of sex-specific loci was conducted with RADsex, which compares presence/absence of non-polymorphic markers between individuals in two groups, with a minimum marker depth of 1 for the process step, and 5 for the distribution, significance and mapping steps, keeping only the significative loci present in just one sex (i-e-, sex-specific loci). To assess the significance of the male/female distribution of markers we conducted a 1,000 permutations test followed by a chi-square test comparing the observed results against the permutations-based expected frequencies. The sex-specific loci were blasted against a reference transcriptome with local blast[34] and then aligned to their genome with bwa[35] to map them to reference genes.

9. Gene expression patterns: We aligned the RNAseq data from each male and female of each species of those with chromosome-level genomes with HISAT2[36] and the alignments were passed to StringTie for transcript assembly[6]. We filtered transcripts with expression lower than 2 TPM and those expressed in less than 2 samples per condition when possible, and we then tested for differential gene expression (DE) between males and females with edgeR after normalisation[37]. In both *Geodia* spp., given there is no chromosome-level genome, we used *de novo* transcriptomes built previously[38,39], and also we built *de novo* transcriptomes for the rest of the species for further annotation. Then reads that were quality filtered with Prinseq lite v 0.20.4 (94), and

decontaminated from rRNA with SortMeRNA v 4.3.6(95) with the provided SILVA 119 Ref NR 99 database(96) were mapped with Bowtie2 aligner(74) and expression levels estimated with RSEM v 1.2.21(97). Expected counts (number of reads mapped) and the normalised trimmed mean of M values (TMM) were used for testing for differential gene expression (DE) between males and females with edgeR after normalisation and using the glmQLFit function. Expression levels of genes with sex-specific loci were retrieved from the count table for plotting. Then, genes with sex-specific loci identified were locally blasted against the genomes and *de novo* transcriptomes (retrieving all isoforms with a hit) of their corresponding species using BLAST and then annotated using Blast2GO PRO(98). The gene annotations from the sex-loci were retrieved and implemented in ShinyGO(99) against the human database to perform Gene Ontology (GO) enrichments (with Benjamini-Hochberg FDR corrections of 0.05) and then plotted using ggplot in R.

10. Identification of sex-biased Alternative Splicing: We used rMATs 4.3.0(100) to identify Alternative Splicing (AS) events in our transcriptomes. We used the trimmed reads from our RNA-seq data for each species (Table S2). Besides assessing annotated splice junctions in the reference genome, rMATs can also evaluate differential splicing between two groups of samples, which in this case were females and males. rMATs measures splicing at each splice site with the percentage spliced in or PSI, which indicates the proportion of two alternative isoforms at each splice site, being 1 and 0 the extremes at which only one of the two alternatives is expressed, and 0.5 indicating equal expression. To compare splicing between sample groups, rMATS calculates the inclusion level difference (ΔPSI), defined as the average PSI in males minus the average PSI in females. ΔPSI ranges from +1 (the isoform is expressed exclusively in males) to −1 (the alternative isoform is expressed exclusively in females). rMATS uses a likelihood-ratio test to identify significant differences in ΔPSI between males and females, and we used the criterion of an FDR P value <0.05 to call differential splicing events. This was calculated both for all genes present in the genome and then for genes with sex-specific loci.

11. Macrosynteny analyses: We used Orthofinder v.2.5.4(101) to identify and compare orthologous groups shared between the six sponge species with chromosome-level genomes: *Oscarella lobularis*, *Chondrosia reniformis*, *Petrosia ficiformis*, *Axinella damicornis*, *Halicondria panicea*, and *Phakelia ventilabrum*. With the set of orthologous groups we then looked at conservation of syntenic regions across sponge genomes with the R package macrosyntR(102). Then, genome location information for each ortholog was extracted from genome annotation files and used to identify regions of synteny by inferring significantly ancient linkage groups (ALG) which were ordered and displayed on Oxford grids for pairwise species comparisons. To evaluate inter-chromosomal rearrangements and conservation of sex-related genetic regions across sponges, we mapped the genomic location of CLGs across sponge chromosomes and generated ribbon diagrams with macrosyntR.

12. Evolutionary trajectories of genes with sex-specific loci: To analyse the presence and evolution of genes associated with sex determination in sponges, we conducted a comprehensive search of an extensive metazoan taxonomic database, with a special emphasis on early-branching animal lineages such as Ctenophora, Cnidaria, Placozoa, and Porifera. To avoid redundancies arising from the use of transcriptomes and ensure better quality and completeness, we included only complete sponge genomes (Porifera). The analysis was performed using a pipeline developed specifically for this study (EvoDomainSearch). This tool allows a systematic and controlled search for genes based on the presence of conserved domains. First, searches were performed with HMMER (v3.4), using HMM models generated from representative sequences of target genes related to sex

determination in metazoans. The recovered sequences were subsequently filtered to retain only those containing the target domains of each gene. The final sequences were aligned using MAFFT (v7.525)(103), employing the L-INS-i strategy for high-quality alignment. Subsequently, ambiguous alignment regions were removed with BMGE (v1.12)(104) using the parameters -m BLOSUM30 -b 3 -h 0.8 -g 0.8, with the aim of retaining only informative and phylogenetically reliable positions. Phylogenetic inferences were then performed using two maximum likelihood methods: FastTree (v2.1.11)(105) for rapid family tree exploration, and IQ-TREE (v2.3.6)(106) for more accurate analyses, employing automatically selected substitution models and statistical assessment using ultrafast bootstrap (UFBoot). This approach allowed us to identify candidate genes in sponges with high confidence and assess their phylogenetic relationship with homologs from other metazoan groups and, when present, their unicellular relatives.

13. Evolutionary age and diversification of sex-specific loci: To investigate the evolutionary origins and diversification of sex-specific genes in sponges, we implemented a multi-step analytical approach that combined ortholog inference with phylostratigraphic reconstruction. We first looked into the shared genome complements of the six species with chromosome-level genomes with OrthoVenn 3(107). Then, orthologous genes were collected with OrthoFinder 3 as before to infer orthologous groups (OGs) and to establish homologous relationships across species. To assign evolutionary ages to genes, we applied phylostratigraphic analysis using GenEra v1.4.2(108) (Barrera-Redondo et al. 2023). Protein sequences predicted from each sponge species were queried against the NCBI non-redundant (NR) protein database using DIAMOND v2.1.9 (66) in sensitive mode. The search employed an e-value threshold of 1e-5 and used the -a flag to input both the species-specific *fasta* files and corresponding NCBI Taxonomy IDs. GenEra automatically traced phylogenetic relationships using NCBI Taxonomy to determine the most basal taxonomic node associated with each gene hit. Genes were then assigned to discrete evolutionary strata based on their inferred taxonomic age.

Among the annotated gene sets, we extracted the genes with sex-specific loci previously identified with RADSex in each species. These genes were mapped to their respective phylostrata to assess the relative timing of their evolutionary origin. We further analysed whether sex-related genes within the same stratum were associated with shared or distinct OGs, enabling insights into patterns of gene duplication and lineage-specific diversification.

To evaluate the functional characteristics of genes with sex-specific loci across evolutionary strata, we performed Gene Ontology (GO) enrichment analysis using ShinyGO as before. Enriched GO terms were then categorized into broader semantic groups reflecting known ancestral functions (e.g., DNA repair, cell cycle regulation, cytoskeleton organization, apoptosis). This allowed us to assess the extent to which genes originally involved in ancient cellular processes were subsequently co-opted into core sexual reproduction functions such as meiosis, gametogenesis, and fertilization. A full list of enriched GO terms and their classifications by stratum is provided in Table S18F. Lastly, we characterized the evolutionary origin of syntenic genes with sex-specific loci identified following the methodology outlined in the Methods Section 10. These analyses facilitated the exploration of conserved genomic architecture and potential linkage patterns associated with sex-related gene clusters.

Supplementary Text

Sex chromosomes can be identified using an array of methods, including karyotyping, cytogenetics, NGS and bioinformatics to unveil coverage and heterozygosity differences, sex-specific SNPs and association tests, and sex-biased expression. Sponge karyotyping is notoriously difficult given the minute size of sponge cells and it has been performed only in a few species(109–111), and therefore we opted for a genomic approach in this study. To investigate the sex chromosome evolution across Porifera, we conducted a phylogeny-wide genome sampling of six sponge gonochoristic species, and two more gonochoristic sponges for which a chromosome-level genome was not available and transcriptomic data was used instead (Fig. S1; Tables S1–S2). But before processing the tissue samples for sequencing, we determined the sex of each individual sponge by surveying the tissue for gametes (Figs. S2–3), finding female and male gametes always in different individuals, pointing to different sex ratios in the populations surveyed, with female-skewed ratios in *Chondrosia reniformis*, *Geodia hentscheli*, *Petrosia ficiformis* and *Halichondria panicea*, and male-skewed ratios in the rest of species (Table S1).

Section 1. Sex chromosome identification and identity of the sex chromosome in *C. reniformis*:

We conducted FindZX for all species but only 2 showed a pattern consistent with sex-linked regions (Fig. 1; Fig. S4). In *C. reniformis,* we found a region within 2.6 to 3.8 Mb on chromosome 11 that showed significantly higher heterozygosity in females and higher coverage in males (Fig. 1A–B; Fig. S5; Tables S3–S6). This region coincides with an area of significant sequence divergence between males and females, where 12 out of the 25 loci with the highest (over 0.20) sex-specific F_{ST} values were located (Table S3). It also exhibits significantly higher SNP density in females (i.e. number of SNPs per unit of genome length; Table S7) and a significantly greater accumulation of sex-specific loci than in other chromosomes ($\chi^2 = 24.339$, df = 13, *p*-value = 0.02815; Fig. 1C, Fig. S6; Table S8B-C), with three times more female- than male-specific loci. In addition, *k*-mer analysis identified female-specific *k*-mers that assembled into 2,178 contigs located within the SDR but extending slightly beyond its boundaries (Fig. 1D), while only 38 *k*-mers were significantly associated with males (Fig. S7A). Combined with the observation that coverage differences between sexes increase with mapping stringency, these results suggest a young ZW system where Z and W are divergent enough to show mapping coverage differences but still conserve homology to align and detect SNPs from W[18].

In chromosome 2 we found higher male coverage, in the region within 3.6 to 4.7 Mb, but there was no heterozygosity bias (Fig. 1A–C; Fig. S5; Table S3–S6). This region contained only a few sex-specific loci (mostly male-specific; Fig. 1C; Fig. S6) and sex-specific k-mers (Fig. 1D; Supp.Fig. 7A-B), consistent with the genomic signature of an older, more differentiated Z chromosome[18]. Interestingly, within the alternate haplotype of the reference genome, we identified with FindZX a scaffold that displayed a female-biased coverage pattern and aligned (via BLAST) precisely with the SDR of chromosome 2 (Fig. S5C–D). In contrast, another scaffold from the alternate haplotype aligned with the SDR of chromosome 11 but exhibited the same coverage bias as the reference sequence (Fig. S5C–D). Likewise, the sex-specific *k*-mers blast found 2437 contigs in the scaffold aligned with chromosome 2 (Fig. S7C–D). This scaffold may represent the corresponding W chromosome homolog, consistent with a more advanced stage of differentiation, resulting from an inversion (Fig. 1E; Fig. S8A). Taken together, these findings suggest that the reference genome of *C. reniformis* corresponds to a ZW (female) individual, with chromosome 2 representing an older, more differentiated ZW pair and chromosome 11 representing a younger, less diverged zw pair, a pattern reminiscent of multi-sex-chromosome systems observed in several

animals(11–13, 112–114). In summary, our results indicate that the individual from which the reference genome was assembled was probably ZW/zw, suggesting that the system in *Chondrosia reniformis* might involve the combination of two pairs of sex chromosomes that segregate independently.

Section 2: Sex chromosome identification and identity of the sex chromosome in *O. lobularis*:

In *Oscarella lobularis*, FindZX detected significantly higher heterozygosity and coverage in males within the 2-3 Mb region of chromosome 5 (Fig. 1E–F), with an increase proportional to mapping stringency. This pattern is consistent with the presence of a differentiating sex chromosome (*10*). Analyses of RADseq data identified 634 putative male-specific loci and none in females, with a significant accumulation on chromosome 5 ($X^2 = 101.69$, df = 19, p-value = 2.641e-13; Fig. 1H; Fig. S6; Table S8B, S8D) and iDlG identified a clear pattern of divergence in chromosome 5 from positions 2–2.75 Mb (Fig. 1I). Since we also observed a higher coverage on males than females (Fig. 1G), this might be indicating that the reference genome was assembled for a male sponge with chromosome 5 acting as the Y chromosome. This hypothesis is also in line with the results obtained with iDIG, as females reached the "2" value, indicating that almost all polymorphisms found for them are homozygous for an allele different than that of the reference genome. However, coverage data analysed in FindZX appears to support a ZW system (Fig. 1F; Fig. S11). Additionally, we performed a *k*-mer analysis following the sexfindR pipeline (*83*) that recovered male-specific *k*-mers covering the SDR (Fig. 1J; Fig. S11).

Given that the sex of the reference genome is unknown, we further explored how our data could point to either a female or male identity for it. We assembled a draft female genome using only the reads of the females, and then mapped the male reads to it using bowtie. Then, we extracted unmapped male reads and re-mapped them to chromosome 5 from the reference chromosome genome. We observed a peak of reads that successfully mapped to the SDR of chromosome 5 from the reference, indicating that the reference chromosome level genome was done from a male individual (Fig. S12).

Notably, the SDR of chromosome 5 corresponds to one of the clearest transposable elements (TE) islands in the *O. lobularis* genome (Fig. S13), a feature typical of Y chromosomes(115). The observed increase in coverage differences with mapping stringency in FindZX (Fig. S11) further supports the presence of sequence divergence between homologous regions. This suggests that while the X and Y chromosomes remain partially homologous, this similarity likely facilitated the assembly of the Y chromosome or X/Y co-assembly into the reference genome. Altogether, these results indicate that *O. lobularis* likely possesses an XY sex-determination system in an intermediate stage of differentiation, with chromosome 5 functioning as the X/Y chromosome.

None of the other four species with reference genomes showed significant biases in their heterozygosity, coverage or SNP density between females and males (Fig. S4; Tables S4–7) that could be indicative of a large sex-linked region. However, in *A. damicornis*, we detected a notable decrease in the male coverage in chromosome 12 (positions 10–12Mb), but it was not accompanied with an increase in heterozygosity or accumulation of RADsex markers (Fig. S4, Fig. S6; Tables S4–7). In addition, no accumulation of TEs accompanied by biases in other parameters, or accumulation of LINEs was detected for any of the genomes of these four species (Supp. Figs. 9–10).

Section 3. Sex-specific loci and sex-biased expression of sex determination markers:

3.1. Shared loci: RADSEX identified 7,657 sex-specific loci across the different species surveyed, pointing to different sexual systems based on the number of loci associated with each sex (Table S8). In *Chondrosia reniformis* and *Oscarella lobularis*, where previous analyses provided strong evidence for ZW and XY systems respectively, our RADSEX results were consistent with those findings. In the remaining species, where no sex chromosomes were detected, the distribution of sex-specific loci suggests the presence of polygenic XY– or ZW–like systems. Specifically, XY systems were inferred in *Geodia hentscheli*, *Geodia barretti*, *Axinella damicornis*, and *Phakellia ventilabrum*, and ZZ/ZW systems in *Petrosia ficiformis*, and *Halichondria panicea* (Fig. 2A; Table S8).

The sponge sex-specific loci mapped to 1,437 unique genes in the genomes of each species (Table S9A–B), 359 of which were previously described as involved in the process of sex determination in other animals (Fig. 2A) but with only 20 in total properly annotated within the GO term category sex determination (GO:0007530) (Fig. 2C; Table S9C-D). For each species, approximately one third of the genes where the sex-specific loci mapped were previously known to have a role in sex determination or gametogenesis (Fig. 2A–C; Table S9C-D). Of the 1,452 genes, approximately 22% were shared by 2 or more species (Fig. 2B). Of these shared genes, very few had known sex-related functions (Fig. 2B; Tables S9C-D, S10). From now on, only those that were present in all of one sex and none of the other will be discussed.

In our sponge species, sex determination pathways seem to converge on the same cohort of genetic regulators that maintain and repair DNA structure and determine gonad fate, as well as regulate the cell cycle, with minor contributions for those in metabolic and/or structural processes (Fig. 2B, E–G). Besides gametogenesis (or sexual reproduction), the annotations recovered for the sex-specific loci fell primarily into the categories of DNA repair and recombination, cell cycle, and translation regulation, suggesting that these genes govern the progression of meiosis and expression of targets through the genomes, similar to what has been found in mammalian Y chromosome genes[65] and those of other organisms[66]. Interestingly, among the genes that were shared by 4 or more species, we found several transposable elements or transposon-related genes that were male-specific, including RTBS, suggested to be involved in promoting spermatogenesis and sperm motility of the snakeskin fish(116) , POL2, POL3 and POL5, that are transposable elements abundant in sex chromosomes and play a key role in sex chromosome evolution(25, 117). One of the most interesting shared genes is YRD6, an uncharacterized protein described in *C. elegans*, that encodes a reverse transcriptase domain and that is shared by all sponge species in our study except for *G. barretti* and *A. damicornis* (Table S11D). YRD6 was found to be downregulated by hypoxia in testicles of the medaka fish *Oryzias latipes*[70], but no other link to sexual reproduction has ever been described to our knowledge. Other male-specific genes widely shared among sponges were the ATP-dependent helicase PIF1 and the RNA binding protein ZNFX1. DExD/H-box RNA helicases are classically associated with sexual development and sex determination, including PL10, vasa, and many other DDX helicases(118). Here, DDX helicases were identified as sex-specific loci in several species, including female-specific in *Halichondria panicea* or male-specific in *Chondrosia reniformis* (Table S10). The role of DNA helicases in sex determination, in turn, has been much less studied. DNA helicases were originally recognized as enzymes that can unwind double-stranded DNA, and among them, the Pif1p helicase family, which belongs to the SF1 superfamily of helicases, is conserved from yeast to humans(119). In our study, PIF1, which has been previously found to be sex specific in the holothurian *Apostichopus japonicus*[73] and differentially expressed in testicular transcriptome from the fish *Channa punctatus*[74], was identified as male-specific and shared by 3 sponge species (Table S11D).

FEM1 genes, here represented by sex markers associated to males in both *Chondrosia reniformis* and *Axinella damicornis* (Table S9D, S10), have been identified as sex-determining genes in the hermaphrodite nematode *Caenorhabditis elegans*[75]. In *C. elegans*, *Fem* genes were inhibited by *Tra-2* and could repress the *Tra-1* through ubiquitination, which is a suppressor of *Mab-3,* and therefore, the expression of the *Fem* genes in males could induce *Mab-3* expression and determine male development(120). Although *FEM1* was consistently associated to male sex determination in our sponges, we have not identified a homologous sequence to *MAB3*, and therefore the downstream pathway most likely differs from the one acting in nematodes. Also, Structure-specific endonuclease subunit SLX1, whose deficiency produces defects in spermatid differentiation(121), was found shared by *Chondrosia reniformis* and *Halichondria panicea*. Finally, *SPER* or *Egg peptide speract receptor*, that facilitates fertilisation and promotes sperm motility in sea urchins(122, 123), was found to be male-specific in *O. lobularis* and *P. ficiformis* (Table S9D, S10).

3.2. Unique Sex-specific genes in SDR regions: In *Chondrosia reniformis*, we identified 658 sex-specific loci, and approximately 7% of them have been previously found to be involved in sex determination processes in other animals (Table S9C, S10). Fifteen of the sex-specific loci were located in the candidate sex-determining regions of chromosomes 2 and 11 (Table S11A), and included *POL, POL2, POL3, POL4, PO21, PIF1, YRD6* (in Chromosome 2) and *CFDP2, PAR14, CENP-E, PTPRD, ADAS, LORF2, GARS,* and *CDK13* (in Chromosome 11). Several of those sex-specific loci have known sex-determining function, and had either male-specificity (*PTPRD*) or female-specificity (*CDK13, CENP-E*) in chromosome 11 (Table S11). Cell cycle and homeostatic regulators of the SDR in chromosome 11 showed moderate to high gene expression levels similar across sexes, while the transposable elements were low expressed but consistently more in females (Table S12). *PTPRD*, or receptor-type tyrosine-protein phosphatase delta, is a candidate trigger for temperature sex determination in reptilians, with male-specific upregulation of *PTPRD* that could act to stabilize commitment to male fate[80]. Cyclin-dependent kinases (CDKs), which regulate the cell cycle and ultimately normal proliferation of cells, and can be found in strata of sex chromosomes of vertebrates(124). In the absence of CDK in mutant mice, male and female infertility is produced by defective attachment of telomeres to the nuclear envelope, resulting in failed or incomplete synapsis of homologous chromosomes(125). Centromere-associated protein E (CENP-E) is a kinesin motor that mediates kinetochore-microtubule attachment and chromosome alignment during cell division(126), it is consistently more expressed in females with B chromosomes in the mouse *Apodemus flavicollis*(127) and it has sex-specific alternative splicing in zebrafish gonads(128). Interestingly, the centromere-associated protein E sequences in sponges appear in different clades, suggesting independent evolutionary trajectories (Fig. S16).

In *Oscarella lobularis*, RADsex identified 634 putative sex markers in males (Fig. 2A; Table S8) 6 of which blasted against genes located in the candidate sex-determining region of chromosome 5 (*G2E3, PIF1, HMCN1, MRC1, DAPK1, RECQ*), and were not found in the SDR of the main sex chromosome of *Chondrosia reniformis*. These sex-specific genes located in the SDR region (Table S11B) showed higher expression in females than in males, similarly to other genes located in the SDR that did not exhibit sex-specific behaviour (Fig. S19D–E). All these sex-specific genes of the SDR of chromosome 5 have known functions related to sex determination in other organisms: *G2/M phase-specific E3 ubiquitin-protein ligase*, or *G2E3*, is conserved across metazoans and is the ancestor of *Phf7*(129), which controls male sex determination in the *Drosophila* germline[87]. *G2E3* plays a crucial role in the gametogenesis of the flatfish *Cynoglossus semilaevis*(130) and is also located in the candidate sex-determining region of the holothurian *Apostichopus japonicus*

being upregulated in males[73] . Like *G2E3, PIF1* and *RecQ* were sex specific in *Apostichopus japonicus*[73], and *RecQ* is involved in DNA repair mechanisms during recombination, that are fundamental to inactivate sex chromosomes(131, 132). *HMCN1*, which promotes cleavage furrow maturation during cytokinesis, plays an important role in gonad development in *Schistosoma japonicum*(133). Finally, *DAPK1*, which is a calcium/calmodulin-dependent serine/threonine kinase, is in the gene cluster of *DMRT* in humans(134) and fish(135) and a sex marker in ratite birds(136).

3.3. Lineage-specific sex-specific loci: The large number of genes with sex-specific loci involved in a wide array of genetic and cellular processes in sponges indicates a cumulative polygenic system for all species, regardless of the presence of sex chromosomes in *O. lobularis* and *C. reniformis*. Besides the conserved genetic regulators that were shared across several species, we identified species-specific genes involved in a panoply of biological processes, ranging from DNA repair and regulation of cell cycle, to regulation of signaling cascades (WNT, JNK and MAPK, among others) and embryonic morphogenesis (Fig. 2A; Fig. S18; Tables S10, S13). Only those with known sex determination or gametogenesis-related function are discussed below (Fig. 2A). A fundamental aspect of sexual reproduction is the capacity of recombination to exchange genetic material during meiosis. After synapsing and pairing, homologous autosomes faithfully separate into haploid gametes and then shuffle genetic material through meiotic recombination. While ancient sex chromosomes only pair at pseudoautosomal regions (PAR) and rarely form breaks outside PAR, young sex chromosomes of fish fully pair revealing that the degeneration and divergence of sex chromosomes is sometimes not sufficient to abrogate pairing[95]. In this process, several genes have been reported to have a prominent role[96], including many that we found to be sex-specific across several sponge species, such as those recognizing meiotic double-strand breaks (DSBs) involved in their repair (Fig. 2B; Fig. S18), such as *RAD52* in *Phakellia ventilabrum* and *RAD51* in *Halichondria panicea* (Table S9D). These genes are central to the exchange of genetic information between non-sister homologs in eukaryotes as their homologs (RecA proteins) in bacteria, and are thought to be the primary sexual genetic machinery that arose to facilitate sexual cycles in eukaryotes[97].

Signaling pathways have crucial roles for sex determination, especially in mammals[98]. Consistently, the MAPK cascade was found to be enriched when analysing the male sex-specific loci of *Chondrosia reniformis*, *Petrosia ficiformis*, *Phakellia ventilabrum*, and *Halichondria panicea*, albeit with different genes (Fig. S18, Table S13). MAPK cascades are evolutionary conserved, intracellular signal transduction pathways that respond to various extracellular stimuli and control a large number of fundamental cellular processes including growth, proliferation, differentiation, motility, stress response, survival and apoptosis[99]. Several of their downstream genes are growth-arrest specific gene 6 that controls gonadotropin release(137), Tumor Necrosis receptors (*TRAF* genes), and growth factor receptors (*EGFR* and *FGFR1–2*), all identified as sex-specific in different species of our target sponges (Tables S9D, S13).

In *O. lobularis* we also identified male sex markers in other sex-related genes out of the candidate SDR (Tables S9–S10), that are associated to male sex differentiation and are part of the SRY cascade: including *ATRX*, and *INSRR*. Notably *ATRX*, a chromatin remodeler downstream of SRY in mammals[101], is involved in testis development in marsupials(138). Interestingly, although *ATRX* genes are present in all sponge species, except for *P. ficiformis*, this important sex-determining gene is sex-specific only in *O. lobularis* (Fig. S17; Table S13). Also, in the SRY pathway, insulin receptor kinase or INSRR(139) is critical for the upregulation of the pathway and testicular

differentiation in mammals(40, 140) . Other genes with known sex-related functions that were male-specific in *O. lobularis* were *BRCA2/FANCD1,* which are involved in DNA repair(132), and whose mutations produces sex reversal in zebrafish(141), and *WDR48* that could cause male DSD in humans and is involved in seminiferous tubules development(142).

In *C. reniformis*, the male sex marker *SPT48*, which locates in Chromosome 11 but out of the candidate SDR, is testis-specific in humans and mice and lack of expression produces azoospermia(143). Also, male-specific was the marker found in the gene *UBXN* in *C. reniformis* (Table S9D). *UBXN* genes positively regulate spermatogenesis in *C. elegans* via TRA-1A degradation(144). As for female sex-markers, we identified the *Sterol O-acyltransferase 1* (*SOAT1*), which is upregulated in the female gonad of the frog *Hoplobatrachus rugulosus* during sex development(145).

In *P. ficiformis*, we identified a female sex-specific marker on an emblematic gene for mammal sex determination, *WNT4*, involved at several steps of mammalian female development[111,112]. Also, female-specific sex markers were found in the genes *H2AX* that is necessary for chromatin remodeling and the inactivation of sex chromosomes during meiosis in male mice(146), *TBX*, which intriguingly is upregulated during testicular development in the rainbow trout(147), and finally, *UBE3A* which regulates the expression of the sex-determining gene *SOX9* in mice(148).

In *Geodia barretti*, we identified a male sex-specific marker on the gene bromodomain-containing protein *BRWD1* (Table S10), which is essential for both male and female fertility, but promotes haploid spermatid–specific transcription[116], and in *G. hentscheli*, male sex-specific markers were found in *AGO2* (Table S10), which is an argonaute protein, and these regulate entry into meiosis and influence silencing of sex chromosomes in the male mouse germline(149), and *CIRBP* (Table S10), which has a role in sex determination in the snapping turtle, *Chelydra serpentina*(150). In *P. ventilabrum*, male-specific sex markers were associated to *MEIG1*, a Meiosis-expressed gene 1 protein specific to the male germline in mammals(151) and molluscs[120], by binding to meiotic chromosomes in spermatocytes prior to meiosis, and *HAP2* (Table S10). *HAP2/GCS1* proteins mediate membrane fusion between gametes in a broad range of eukaryotes, ranging from algae and higher plants to protozoans and cnidaria, which suggest they are derived from an ancestral gamete fusogen[121]. In cnidarians, *HAP2* localises in the testes, indicating a role regulating the male machinery for gamete production[121].

In *A.damicornis*, the gene containing a male-specific marker, *PKRDC*, is a DNA-dependent protein kinase catalytic subunit that plays an important role in testis function via meiotic regulation(152). In *H. panicea*, we identified a female-specific sex marker in the gene *NOM1* (Table S10), which is a sex determining gene in the fish *Acrossocheilus fasciatus* together with *WNT4* and *DMRT3*(153) and another in the gene *MRCKA* (Table S10), which causes severe male infertility in humans(154).

Section 4. Expression of genes with sex-specific loci and alternative splicing: In organisms without divergent sex chromosomes, and therefore largely identical genomes, intergenomic conflict may arise, which may be resolved by the evolution of sex-specific gene expression. One of the best characterized examples of AS in the pathway of somatic sex determination is that of the genus *Drosophila*[125]. There, *Sex-lethal* (*SXL*), *Transformer* (*TRA*), and *Transformer 2* (*TRA2*), operates in a multilevel splicing cascade that regulates sex-specific splicing of the downstream targets (male-specific lethal-2, fruitless, and doublesex), controlling sexually dimorphic cell differentiation and dosage compensation[125].

To understand gene expression patterns in genes with sex-specific loci, we first looked at differential gene expression using the matrix of transcripts produced by StringTie in the species with reference genomes. We found that in general, very few transcripts had differential expression, from 0.07% to 3.6% of the total transcripts (Table S14). Similarly, for the two *Geodia* spp., we found 0.2–0.4% of the transcriptome differentially expressed (Table S14). The species with more transcripts differentially expressed were *P. ventilabrum*, the *C. reniformis* and then *P. ficiformis* (Table S14). Among these transcripts differentially expressed we found that between 0.3 to 11.5% contained sex-specific loci, but none in *O. lobularis* or *A. damicornis* (Fig. 3A, Table S14). We also observed that for most genes with sex-specific genes, several transcripts were produced, with expression in both sexes (Fig. S18), although generally more in one of the 2 sexes, equalizing the overall gene expression for those loci (Fig. 18B–H; Fig. S19; Tables S14, S15).

Then, we used RNA-seq short read data to understand AS on the genomes of the six species with rMATs. We provide evidence that all sponges had AS, ranging from only 3.5% of their genes with AS events (838 events), as in *P. ficiformis*, to a 60.5% (10626 events) as in *C. reniformis* (Fig. 3A, Table S16A). Across sponges, we detected all five types of AS, skipped exons (SE), mutually exclusive exons (MX), alternative 5′ splice sites (A5), alternative 3′ splice sites (A3) and retained intron (RI), with SE being the most abundant in all species (Fig. 3A, Table S16A). In general a very low percentage of the genes with sex loci presented significant sex-biased AS, and SE was the most common AS type among them (Fig. 3A, Table S16A). In *C. reniformis*, the sponge with more AS, 142 genes showed significantly sex-biased AS, and from those only 11 were genes with sex-specific loci (Table S14A). Interestingly, in *O. lobularis*, there was no sex-biased AS event, either among the autosomal genes or those with sex-specific loci (Fig. 3A, Table S16A), while in *H. panicea*, there were 6 genes with significantly sex-biased events, but none of them were genes with sex-specific loci (Fig. 3A, Table S16A). In the rest of species, less than 6 genes with sex-specific loci contained AS significantly associated to sex (Fig. 3A, Table S14A). Our results indicate that sex-biased AS had only a minor role reducing conflict related to gene expression in sponges.

Only in *Chondrosia reniformis*, whose sexual system was identified as ZZ/ZW, expression of genes with sex-specific loci was significantly higher in females for those genes that produced several isoforms (Table S15). In the rest of species, the expression of genes with sex-specific loci was slightly higher in males (but not significantly), for both female and male sex loci and with and without isoforms (Fig. 3C–J; Fig. S18). Our results indicate that although sex-specific loci are determining the sex in all of our target species, the expression of these sex-specific loci is generally equalized across sexes because of the production of several isoforms expressed in both sexes, and only in a few cases we detect sex-biased AS (Fig. 3C–J; Fig. S18).

Some known genes involved in sex determination, such as *FEM1, WNT4, ATRX, WRN, WDR48*, and *STAR* that presented sex-specific loci produced a single transcript, in most cases with more expression in the sex that had the bias to, like *ATRX* in *Oscarella lobularis* more expressed in males (Fig. S18A) or *WNT4* and *STAR5* in *Petrosia ficiformis* more expressed in females (Fig. S18C). However, in other cases, the expression was biased towards the opposite sex in which we found the SNP fixed in RADSEX, like *WRN* in *Chondrosia reniformis* which is a gene with a male-specific locus that was more expressed in females (Fig. S18B) or *WDR48* in *Phakellia ventilabrum* that is a gene with a female-specific locus with more expression in males (Fig. S18G). In the species where the gene for GATA contained sex-specific loci, *Geodia barretti* and *Geodia*

hentscheli, we detected several sex-specific isoforms, where one was more expressed in females and the other in males (Fig. S18D–E).

Section 5. Conservation of synteny and evolutionary age of sex markers: Orthofinder generated 16953 orthogroups, with 91.5% of total proteins from all species being placed into an orthogroup. There were 4074 orthogroups containing proteins from all species, with 941 single copy orthogroups shared between all species, which is less than the 5% of the genomic content for the species. Using the set of single copy orthologs shared between the six sponge species, we investigated how macrosyntenic regions have been rearranged across sponge genomes (Fig. 4A). Our analysis shows well conserved synteny between the homoscleromorph *O. lobularis* and the five demosponges (Fig. 4A), which is indicative of good conservation of chromosome architecture across sponges and is not always the case in other animal phyla(155). Our analysis shows that all of the assembled chromosomes of the six sponge species have either a 1:1 equivalence or result from fusions and fissions of chromosomes across classes (Fig. 4A). Additionally, we find higher conservation of synteny between the homoscleromorph *O. lobularis*, and the demosponges *C. reniformis* and *P. ficiformis*, while there is a higher degree of rearrangement within *P. ventilabrum*, *A. damicornis*, and *H. panicea*, possibly influenced by the presence of a great number of chromosome fission events between *A. damicornis* and *H. panicea*(155).

Our results indicate only slight conservation of synteny among genes with sex-specific loci, since among the 2895 sex loci detected across all species, only 65 were found in synteny: 18 of *C. reniformis*, 10 of *O. lobularis*, 36 of *P. ficiformis*, and 1 of *H. panicea* (Fig. 4B; Table S17), which represents 8% of the syntenic orthologs. Most of the sex-loci of *C. reniformis* that showed synteny were located in chromosome 11, whereas for the rest of the species they were distributed across many chromosomes (Fig. 4B). Only 3 of the sex-loci that showed synteny were identified as sex-specific in more than 1 species: for instance, *OXSR1*, that was identified as sex specific in *C. reniformis* and located in chromosome 11 was also sex-specific in *O. lobularis* and showed synteny to its chromosome 13 (Fig. 4B; Table S17). Similarly, two sex-loci identified in *P. ficiformis* (*ENOF1* and *ERBB3*) were also sex-specific for *C. reniformis*, but while the first one is located in chromosome 11 of *C. reniformis* and Chromosome 15 in *P. ficiformis*, the second one is in chromosome 1 of *C. reniformis* and Chromosome 11 of *P. ficiformis* (Fig. 4B; Table S17). Perhaps the most striking result is the conservation of synteny detected for the sex loci of *C. reniformis* that fell on the sex chromosome (chromosome 11). While these genes were approximately 80% of the genomic make–up of the syntenic ALG in dark green, they represented only half of chromosome 11 of *C. reniformis*, given that this chromosome appears to be the result of a fusion between 2 ancestral autosomal chromosomes (Fig. 4A–B). The genomic instability of this fusion could have prompted the appearance of an inversion, leading to the restriction of recombination in this pair of chromosomes and thereby the emergence of the sex chromosomes in *C. reniformis*.

When analysing the 2 species with sex chromosomes, the macrosyntenic pattern is stronger, with 4,582 single-copy orthologs in synteny in 10 ALGs (Fig. S20A). The sex-chromosomes of both species were largely syntenic but resulted from fusions of other chromosomes (Fig. S20A–B). We analysed the GO enrichments of the ALGs where we detected the 115 syntenic genes with sex-loci and those in the sex chromosomes, and recovered regulation of mitotic cell cycle, chromosome organisation, or cell cycle DNA replication (Fig. S20C–E).

Our results support the notion that sex chromosomes evolve independently from autosomes(156), with a great turnover because sponge sex chromosomes are still largely undifferentiated, which makes them more prone to reversion to autosomes. In contrast, fully differentiated sex chromosomes are more stable(157), like the X chromosome of insects that is conserved along the entire class(158), the Z chromosome of birds(159), or the Y chromosome of eutherian mammals(160). Such degree of conservation may also be linked to the evolutionary scale analysed, because while the avian W chromosome arose approximately 140 million years ago(161), the divergence of the demosponge and homoscleromorph clades analysed here occurred at least 580 million years ago(162), placing the putative origin of common sex chromosome (if any) much earlier.

Section 6. Evolutionary age of sex markers: We investigated the evolutionary age of protein-coding genes in six sponge species (5 demosponges and 1 homoscleromorph) using GenEra, assigning each gene to its most probable taxonomic origin based on their homology. The results are summarized in the supplementary table (Table S18A–B). A total of 102,213 genes were analyzed across the species, with individual species contributing between 15,210 loci (*Oscarella lobularis*) and 21,424 loci (*Petrosia ficiformis*). A total of 11,354 genes (11.1%) were classified as species-specific, with *Phakellia ventilabrum* and *Chondrosia reniformis* showing the highest numbers (2,532 and 2,540 genes, respectively). The number of genes assigned to the different taxonomic levels varied by species and clade. However, the majority of genes in each species were assigned to ancestral nodes such as Metazoa (9,805 genes in total), Eukaryota (12,656), and Cellular Organism (54,793), reflecting their conserved nature across animals and deeper eukaryotic lineages. Notably, a substantial proportion (5,544 genes) were inferred to have originated at the Heteroscleromorpha level, consistent with shared innovations within this major demosponge clade (Fig. 4C). This burst in genomic innovation in haplosclerid sponges is similar to what was found in other studies looking at genomic content across animals(54, 163).

To investigate the evolutionary history of genes involved in sexual processes, we collected the sex-specific genes (n = 2,705) from the full GenEra annotation dataset and analyzed them based on their inferred evolutionary origin (Table 18C). The majority of sexual-function genes were assigned to deep evolutionary nodes, with cellular organisms accounting for 1,725 loci (63.8%), followed by Eukaryota (297 loci, 11%), Metazoa (233 loci, 8.6%), and Opisthokonta (98 loci, 3.6%). This suggests a conserved core of sex-related genes that predate the emergence of metazoans (Fig. 4C). Only a minority of sex loci were assigned to sponge-specific or lineage-specific levels: species-specific genes represented just 6.3% (170 loci) of the total, and assignments to sponge clades such as Heteroscleromorpha (134 loci), Demospongiae (43 loci), and more derived sponge lineages (e.g., Axinellida + Suberitida, Chondrillida, etc.) were relatively infrequent. Notably, species such as *Petrosia ficiformis* and *Phakellia ventilabrum* showed a broader distribution of gene origins across taxonomic levels, including genes assigned to Haplosclerida, Porifera, and Demospongiae, whereas *Halichondria panicea* and *Oscarella lobularis* displayed a restricted set of origins with few or no lineage-specific assignments. These patterns suggest that while the majority of sex-related genes are highly conserved, there may be some degree of lineage-specific recruitment or innovation of reproductive genes in particular demosponge clades. However, the generally low proportion of such lineage-specific sex loci implies that the molecular toolkit for sexual reproduction in sponges is largely ancient and evolutionarily stable.

We further characterized the distribution of sex-related genes by mapping them to orthogroups (OGs) and summarizing their presence across phylogenetic strata (Table S18C). Overall, we identified 1,764 OGs containing sex-related genes, comprising a total of 2,639 genes. Notably, there was a close correspondence between the number of OGs and the number of genes at each stratum, suggesting that most sex-related OGs contain only one or a few genes per species.

The deepest nodes in the tree harbored the highest number of sexual OGs: Cellular organisms (1,064 OGs / 1,709 genes), Eukaryota (218 OGs / 253 genes) and Metazoa (180 OGs / 233 genes). The high overlap between the number of sex-related OGs and the number of sexual genes at each level suggests that redundancy within orthogroups is limited, and that sex-related genes tend to be evolutionarily conserved and non-duplicated. This could reflect strong functional constraints on core components of the sexual machinery, which may limit gene duplication or diversification. Moreover, the predominance of these OGs in ancient strata supports the idea that the genetic toolkit for sexual reproduction was already well established in early eukaryotic and metazoan ancestors. The modest contribution of more recent, lineage-specific OGs might correspond to functional fine-tuning or clade-specific adaptations, rather than de novo innovation of entirely new sexual pathways.

To explore the putative functions of sex-related genes conserved in ancient evolutionary strata, we performed GO enrichment analyses for each species and focused on genes assigned to deep taxonomic levels (e.g., Cellular Organisms, Eukaryota, Metazoa, Table S18E). Enriched GO terms were subsequently grouped into functional categories based on their semantic similarity and supported associations with sexual functions in animals, as described in the literature (Table S18D). Specifically, enriched terms were classified into seven biologically relevant categories: DNA Repair & Recombination, linked to meiosis(164), Cell-Cycle Control, associated with gametogenesis(165), Membrane Fusion & Adhesion, involved in fertilization processes(166), Transcriptional Regulation, related to sex determination(167), Cytoskeleton Dynamics, underpinning gamete motility(168), Stress Response, connected to mate recognition mechanisms(169), and Apoptosis, implicated in germline quality control(170). This categorization allowed us to functionally characterize the ancient core of sex-related genes in sponges and assess potential evolutionary conservation or lineage-specific specialization across the species analyzed (Fig. 4D). The GO enrichment analysis of sex loci revealed distinct functional categories associated with specific phylogenetic depths. Genes assigned to deep evolutionary strata, particularly Cellular Organisms and Eukaryota, were predominantly enriched in terms related to DNA repair and recombination, cell-cycle regulation, cytoskeleton organization, membrane fusion and adhesion, stress response, and apoptosis. These categories correspond to key cellular processes known to underlie meiotic regulation, gamete motility, fertilization, and germline quality control in metazoans, indicating that the molecular basis of sexual function is largely rooted in ancient cellular mechanisms. The high enrichment of DNA repair and recombination among the functions in which our sex loci were involved are then in line with the hypothesis that sex emerged as an adaptation for DNA repair, in fact, as a mechanism facilitating a repair process known as homologous recombinational repair, known to carry out double-strand damages[97].

Genes assigned to the Metazoa stratum were more frequently associated with transcriptional regulation and reproductive process terms, consistent with the emergence of more specialized, multicellular reproductive systems and sex determination pathways within animals[97]. For instance, there is now evidence of a high degree of conservation of sex chromosomes for more than 300 million years in vertebrates(171), and our results point to an even higher conservation of the sex

complements across animals, probably dating before the Cambrian. Notably, a subset of syntenic genes shared among all six sponge species (highlighted in dark grey in the circular plot) were predominantly associated with non-metazoan strata such as cellular organisms and Eukaryota. These conserved blocks suggest that the ancient core of sex-related genes not only retains functional signatures of ancestral cellular processes but also exhibits conserved genomic organization across diverse demosponge lineages. This pattern reinforces the idea that while regulatory complexity may have expanded in the metazoan ancestor, the foundational components of sexual function were established much earlier and maintained through strong selective constraints, as supported by data showing presence of meiotic sex and its machinery in the Last Common Ancestor of Eukaryota(172). In fact, recent evolutionary theories suggest that meiotic sex could be a reaction to endosymbiont entry, and more precisely in reaction to reactive oxygen species (ROS) stress(172). We also found enrichments of stress responses and cell fusion processes in genes with sex-specific loci shared across sponges, supporting the idea that the stress and DNA damage, likely created by the incorporation of endosymbionts, could be a driver of meiotic sex appearance, as a mechanism facilitating a repair process known as homologous recombinational repair(3, 172).

Section 7. Patterns of Evolution of sex determination systems in animals: A major stumbling block to understanding the evolution of sexual systems across Metazoa is understanding their current distribution among organisms. To this end, here we built two composite trees for Porifera and Metazoa from previously published studies(173–175) and overlapped the presence of sex chromosomes, sexual mode (gonochorism/hermaphroditism), and sexual strategy (XY and ZW) on each of the clades (Fig. 5). For Porifera, we included 44 families from the four classes (Fig. 5A), and for metazoans we provide information for all the 32 phyla currently known (Fig. 5B). Our final data set of sponges includes 21 gonochoristic, 21 hermaphroditic and one uncharacterized family (no information about its sexual mode has been published) (Fig. 5A), showing that the last common ancestor of all poriferans was hermaphroditic (*8*). Although the transition to separate sexes was very common through sponge evolution, this seems to be uncoupled from the sexual strategy that evolved within those lineages (Fig. 5A). Such diversity in molecular tools to produce dioecy is also pervasive along the entire Tree of Animals (Fig. 5B). While in many clades of metazoans the presence of sex chromosomes appears to have coevolved with a ZW strategy (such as in nematodes, planarians, or echinoderms), in many others these two traits are unlinked (in insects, molluscs, and vertebrates), and there is remarkable diversity in the sex chromosome system exhibited at the phylum, class and even order level (Fig. 5B)(2, 176). But even more, dramatic sex chromosome turnover is observed at the family level in some groups of animals, like carpodatylid geckos, where the evolution of ZW sex determination systems occurs several times independently, with evolution of XY and even environmental sex determination systems(177).

Although the sexual mode is known for all phyla, excepting for Micrognatozoa (Fig. 5B), there are only a handful of studies dealing with the existence of sexual chromosomes and the system they presented, most of them focused in vertebrates(178, 179) or some in other closely-related invertebrates, such as echinoderms(176) or the well-known cephalochordate *Amphioxus*(180). Among the rest of invertebrates, it is not surprising that the bits of information about sex chromosomes are focused in some well-established bilaterian model species, especially within arthropods(181), nematodes(182) or platyhelminthes(183), and few in molluscs(184) and annelids(185). More recently, the interest to investigate sex determination across other non-bilaterian groups has peaked(6, 186), especially since the resources, tools and algorithms have

advanced significantly. But, given the mechanisms that determine sex can vary widely among all known animal taxa, the research in a broader diversity of species and phyla, and specially within the most ancestral ones, is essential in order to understand how sex is determined and how it evolves across metazoans. In this sense, our study is considerably expanding the knowledge on the evolution of sexual strategies across Porifera, since we provide for the first time the sexual system of seven of these families, with four of them presenting a XY-like system, while three are ZW-like (Fig. 5A). And even more importantly, we provide evidence of a conserved set of loci that are prone to sex-specification mechanisms since the dawn of the animal lineage, but an independent emergence of sex chromosomes in Porifera, most likely caused by the instability of structural inversions that occurred during chromosomal fusions.

References

1. Bachtrog, D. et al. Sex determination: why so many ways of doing it? **PLoS Biol. 12**, e1001899 (2014). https://doi.org/10.1371/journal.pbio.1001899

2. Zhu, Z., Younas, L. & Zhou, Q. Evolution and regulation of animal sex chromosomes. **Nat. Rev. Genet.** (2025). https://doi.org/10.1038/s41576-024-00757-3 (in the press).

3. Bernstein, H., Bernstein, C. & Michod, R. E. DNA repair as the primary adaptive function of sex in bacteria and eukaryotes. **DNA Repair 18**, 111–146 (2012). https://doi.org/10.1016/j.dnarep.2012.03.001

4. Hörandl, E. & Speijer, D. How oxygen gave rise to eukaryotic sex. **Proc. R. Soc. B 285**, 20172706 (2018). https://doi.org/10.1098/rspb.2017.2706

5. Moore, E. C. & Roberts, R. B. Polygenic sex determination. **Curr. Biol. 23**, R510–R512 (2013). https://doi.org/10.1016/j.cub.2013.04.004

6. Chen, R. et al. XY sex determination in a cnidarian. **BMC Biol. 21**, 32 (2023). https://doi.org/10.1186/s12915-023-01532-2

7. Steenwyk, J. L. & King, N. Integrative phylogenomics positions sponges at the root of the animal tree. **Science 390**, 751–756 (2025). https://doi.org/10.1126/science.adw9456

8. Riesgo, A. et al. Inferring the ancestral sexuality and reproductive condition in sponges (Porifera). **Zool. Scr. 43**, 101–117 (2014). https://doi.org/10.1111/zsc.12031

9. Riesgo, A., Farrar, N., Windsor, P. J., Giribet, G. & Leys, S. P. The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. **Mol. Biol. Evol. 31**, 1102–1120 (2014). https://doi.org/10.1093/molbev/msu057

10. Sigeman, H., Sinclair, B. & Hansson, B. FindZX: an automated pipeline for detecting and visualising sex chromosomes using whole-genome sequencing data. **BMC Genomics 23**, 328 (2022). https://doi.org/10.1186/s12864-022-08432-9

11. Yoshido, A. et al. Evolution of multiple sex-chromosomes associated with dynamic genome reshuffling in Leptidea wood-white butterflies. **Heredity** 125, 138–154 (2020). https://doi.org/10.1038/s41437-020-0325-9

12. Saunders, P. A., Perez, J., Ronce, O. & Veyrunes, F. Multiple sex chromosome drivers in a mammal with three sex chromosomes. **Curr. Biol. 32**, 2001–2010.e3 (2022). https://doi.org/10.1016/j.cub.2022.03.029

13. Schartl, M. Sex determination by multiple sex chromosomes in Xenopus tropicalis. **Proc. Natl Acad. Sci. USA 112**, 10575–10576 (2015). https://doi.org/10.1073/pnas.1513518112

14. Zhou, Y. et al. Chromosome-level echidna genome illuminates evolution of multiple sex chromosome system in monotremes. **Gigascience** 14, giae112 (2025). https://doi.org/10.1093/gigascience/giae112

15. Charlesworth, D., Charlesworth, B. & Marais, G. Steps in the evolution of heteromorphic sex chromosomes. **Heredity 95**, 118–128 (2005). https://doi.org/10.1038/sj.hdy.6800697

16. Stevison, L. S., Hoehn, K. B. & Noor, M. A. F. Effects of inversions on within- and between-species recombination and divergence. **Genome Biol. Evol. 3**, 830–841 (**2011**). https://doi.org/10.1093/gbe/evr081

17. Natri, H. M., Merilä, J. & Shikano, T. The evolution of sex determination associated with a chromosomal inversion. **Nat. Commun. 10**, 882 (**2019**). https://doi.org/10.1038/s41467-018-08014-y

18. Coffing, G. C. et al. Cephalopod sex determination and its ancient evolutionary origin. **Curr. Biol. 35**, 931–939.e4 (**2025**). https://doi.org/10.1016/j.cub.2025.01.005

19. Galià-Camps, C. et al. Genomic richness enables worldwide invasive success. Preprint at https://doi.org/10.21203/rs.3.rs-3902873/v1 (**2024**).

20. Kocher, T. D., Meisel, R. P., Gamble, T., Behrens, K. A. & Gammerdinger, W. J. Yes, polygenic sex determination is a thing! **Trends Genet. 40**, 1001–1017 (**2024**). https://doi.org/10.1016/j.tig.2024.10.003

21. Picard, M. A. L. et al. The roles of Dmrt (doublesex/male-abnormal-3 related transcription factor) genes in sex determination and differentiation mechanisms: ubiquity and diversity across the animal kingdom. **C. R. Biol. 338**, 451–462 (**2015**). https://doi.org/10.1016/j.crvi.2015.04.010

22. Raymond, C. S. et al. Evidence for evolutionary conservation of sex-determining genes. **Nature 391**, 691–695 (**1998**). https://doi.org/10.1038/35618

23. Bellott, D. W. et al. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. **Nature 508**, 494–499 (**2014**). https://doi.org/10.1038/nature13206

24. Tse, A. C. K. et al. Hypoxia alters testicular functions of marine medaka through microRNAs regulation. **Aquat. Toxicol. 180**, 266–273 (**2016**). https://doi.org/10.1016/j.aquatox.2016.10.007

25. Dechaud, C., Volff, J.-N., Schartl, M. & Naville, M. Sex and the TEs: transposable elements in sexual development and function in animals. **Mob. DNA 10**, 48 (**2019**). https://doi.org/10.1186/s13100-019-0185-0

26. Wang, Y., Yang, Y., Li, Y. & Chen, M. Identification of sex determination locus in sea cucumber Apostichopus japonicus using genome-wide association study. **BMC Genomics 23**, 391 (**2022**). https://doi.org/10.1186/s12864-022-08632-3

27. Roy, A., Basak, R. & Rai, U. De novo sequencing and comparative analysis of testicular transcriptome from different reproductive phases in freshwater spotted snakehead Channa punctatus. **PLoS ONE 12**, e0173178 (**2017**). https://doi.org/10.1371/journal.pone.0173178

28. Nakamura, M. The mechanism of sex determination in vertebrates—are sex steroids the key-factor? **J. Exp. Zool. A Ecol. Genet. Physiol. 313A**, 381–398 (**2010**). https://doi.org/10.1002/jez.616

29. Viger, R. S., de Mattos, K. & Tremblay, J. J. Insights into the roles of GATA factors in mammalian testis development and the control of fetal testis gene expression. **Front. Endocrinol. 13**, 902198 (**2022**). https://doi.org/10.3389/fendo.2022.902198

30. Tremblay, J. J. & Viger, R. S. GATA factors differentially activate multiple gonadal promoters through conserved GATA regulatory elements. **Endocrinology 142**, 977–986 (**2001**). https://doi.org/10.1210/endo.142.3.7995

31. Doniach, T. & Hodgkin, J. A sex-determining gene, fem-1, required for both male and hermaphrodite development in Caenorhabditis elegans. **Dev. Biol. 106**, 223–235 (**1984**). https://doi.org/10.1016/0012-1606(84)90077-0

32. Bopp, D., Saccone, G. & Beye, M. Sex determination in insects: variations on a common theme. **Sex. Dev. 8**, 20–28 (**2014**). https://doi.org/10.1159/000356458

33. Yang, S. Y., Baxter, E. M. & Van Doren, M. Phf7 controls male sex determination in the Drosophila germline. **Dev. Cell 22**, 1041–1051 (**2012**). https://doi.org/10.1016/j.devcel.2012.04.013

34. Nath, S., Welch, L. A., Flanagan, M. K. & White, M. A. Meiotic pairing and double-strand break formation along the heteromorphic threespine stickleback sex chromosomes. **Chromosome Res. 30**, 429–442 (**2022**). https://doi.org/10.1007/s10577-022-09699-0

35. Benson, F. E., Baumann, P. & West, S. C. Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. **Nature 391**, 401–404 (**1998**). https://doi.org/10.1038/34937

36. Pattabiraman, S. et al. Mouse BRWD1 is critical for spermatid postmeiotic transcription and female meiotic chromosome stability. **J. Cell Biol. 208**, 53–69 (**2015**). https://doi.org/10.1083/jcb.201404109

37. Huo, Y. et al. Identification and functional analysis of Tex11 and Meig1 in spermatogenesis of Hyriopsis cumingii. **Front. Physiol. 13**, 961773 (**2022**). https://doi.org/10.3389/fphys.2022.961773

38. Steele, R. E. & Dana, C. E. Evolutionary history of the HAP2/GCS1 gene and sexual reproduction in metazoans. **PLoS ONE 4**, e7680 (**2009**). https://doi.org/10.1371/journal.pone.0007680

39. Windley, S. P. & Wilhelm, D. Signaling pathways involved in mammalian sex determination and gonad development. **Sex. Dev. 9**, 297–315 (**2015**). https://doi.org/10.1159/000444065

40. Pitetti, J. L. et al. Insulin and IGF1 receptors are essential for XX and XY gonadal differentiation and adrenal development in mice. **PLoS Genet. 9**, e1003160 (**2013**). https://doi.org/10.1371/journal.pgen.1003160

41. Biason-Lauber, A. Control of sex development. **Best Pract. Res. Clin. Endocrinol. Metab. 24**, 163–186 (**2010**). https://doi.org/10.1016/j.beem.2009.12.002

42. Magaña-Acosta, M. & Valadez-Graham, V. ATRX: from chromatin remodeling to disease. **Genesis 63**, e70031 (**2025**). https://doi.org/10.1002/dvg.70031

43. Bernard, P. & Harley, V. R. Wnt4 action in gonadal development and sex determination. **Int. J. Biochem. Cell Biol. 39**, 31–43 (**2007**). https://doi.org/10.1016/j.biocel.2006.06.007

44. Plotnikov, A., Zehorai, E., Procaccia, S. & Seger, R. The MAPK cascades: signaling components, nuclear roles and mechanisms of nuclear translocation. **Biochim. Biophys. Acta 1813**, 1619–1633 (**2011**). https://doi.org/10.1016/j.bbamcr.2010.12.012

45. Kaiser, V. B. & Ellegren, H. Nonrandom distribution of genes with sex-biased expression in the chicken genome. **Evolution 60**, 1945–1951 (**2006**). https://doi.org/10.1111/j.0014-3820.2006.tb00537.x

46. Xiao, Y., Xia, Y. & Zeng, X. Sex-biased expression of sex-linked transcripts in spiny frogs with homomorphic sex chromosomes. **BMC Genomics 26**, 636 (**2025**). https://doi.org/10.1186/s12864-025-11846-w

47. Planells, B., Gómez-Redondo, I., Pericuesta, E., Lonergan, P. & Gutiérrez-Adán, A. Differential isoform expression and alternative splicing in sex determination in mice. **BMC Genomics 20**, 202 (**2019**). https://doi.org/10.1186/s12864-019-5572-x

48. Guo, L., Yu, H. & Li, Q. Sex-specific mRNA alternative splicing patterns and Dmrt1 isoforms contribute to sex determination and differentiation of oyster. **Int. J. Biol. Macromol. 283**, 134557 (**2024**). https://doi.org/10.1016/j.ijbiomac.2024.137747

49. Telonis-Scott, M., Kopp, A., Wayne, M. L., Nuzhdin, S. V. & McIntyre, L. M. Sex-specific splicing in Drosophila: widespread occurrence, tissue specificity and evolutionary conservation. **Genetics 181**, 421–434 (**2009**). https://doi.org/10.1534/genetics.108.096743

50. Flintham, E. The evolution of sex-specific gene expression in polygenic traits. **J. Evol. Biol. 38**, 939–951 (**2025**). https://doi.org/10.1093/jeb/voaf050

51. Weber, C. & Capel, B. Sex determination without sex chromosomes. **Philos. Trans. R. Soc. Lond. B Biol. Sci. 376**, 20200109 (**2021**). https://doi.org/10.1098/rstb.2020.0109

52. Fernandez-Valverde, S. L., Calcino, A. D. & Degnan, B. M. Deep developmental transcriptome sequencing uncovers numerous new genes and enhances gene annotation in the sponge Amphimedon queenslandica. **BMC Genomics 16**, 314 (**2015**). https://doi.org/10.1186/s12864-015-1588-z

53. Böhne, A., Wilson, C. A., Postlethwait, J. H. & Salzburger, W. Variations on a theme: genomics of sex determination in the cichlid fish Astatotilapia burtoni. **BMC Genomics 17**, 883 (**2016**). https://doi.org/10.1186/s12864-016-3178-0

54. Kenny, N. J. et al. Tracing animal genomic evolution with the chromosomal-level assembly of the freshwater sponge Ephydatia muelleri. **Nat. Commun. 11**, 3676 (**2020**). https://doi.org/10.1038/s41467-020-17397-w

55. Simakov, O. et al. Deeply conserved synteny and the evolution of metazoan chromosomes. **Sci. Adv. 8**, eabi5884 (**2022**). https://doi.org/10.1126/sciadv.abi5884

56. Fernández, R. & Gabaldón, T. Gene gain and loss across the metazoan tree of life. **Nat. Ecol. Evol. 4**, 524–533 (**2020**). https://doi.org/10.1038/s41559-019-1069-x

57. Graves, J. A. M. & Shetty, S. Sex from W to Z: evolution of vertebrate sex chromosomes and sex determining genes. **J. Exp. Zool. 290**, 449–462 (**2001**). https://doi.org/10.1002/jez.1088

58. Speijer, D. What can we infer about the origin of sex in early eukaryotes? **Philos. Trans. R. Soc. Lond. B Biol. Sci. 371**, 20150530 (**2016**). https://doi.org/10.1098/rstb.2015.0530

59. Augstenová, B., Pensabene, E., Veselý, M., Kratochvíl, L. & Rovatsos, M. Are geckos special in sex determination? Independently evolved differentiated ZZ/ZW sex chromosomes in carphodactylid geckos. **Genome Biol. Evol. 13**, evab119 (**2021**). https://doi.org/10.1093/gbe/evab119

60. Morrow, C. & Cárdenas, P. Proposal for a revised classification of the Demospongiae (Porifera). **Front. Zool. 12**, 26 (**2015**). https://doi.org/10.1186/s12983-015-0099-8

61. Giribet, G. & Edgecombe, G. D. The invertebrate tree of life (Princeton Univ. Press, Princeton, 2020).

Methods

62. Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. **PLoS ONE 7**, e37135 (**2012**). https://doi.org/10.1371/journal.pone.0037135

63. Taboada, S. et al. Genetic diversity, gene flow and hybridization in fan-shaped sponges (Phakellia spp.) in the North-East Atlantic deep sea. **Deep Sea Res. Part I Oceanogr. Res. Pap. 181**, 103274 (**2022**). https://doi.org/10.1016/j.dsr.2021.103685

64. Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M. & Borodovsky, M. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. **NAR Genom. Bioinform. 3**, lqaa108 (**2021**). https://doi.org/10.1093/nargab/lqaa108

65. Stanke, M. et al. AUGUSTUS: ab initio prediction of alternative transcripts. **Nucleic Acids Res. 34**, W435–W439 (**2006**). https://doi.org/10.1093/nar/gkl200

66. Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. **Bioinformatics 24**, 637–644 (**2008**). https://doi.org/10.1093/bioinformatics/btn013

67. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. **Nat. Biotechnol. 33**, 290–295 (**2015**). https://doi.org/10.1038/nbt.3122

68. Gremme, G., Brendel, V., Sparks, M. E. & Kurtz, S. Engineering a software tool for gene structure prediction in higher organisms. **Inf. Softw. Technol. 47**, 965–978 (**2005**). https://doi.org/10.1016/j.infsof.2005.09.005

69. Venturini, L., Caim, S., Kaithakottil, G. G., Mapleson, D. L. & Swarbreck, D. Leveraging multiple transcriptome assembly methods for improved gene structure annotation. **Gigascience 7**, giy093 (**2018**). https://doi.org/10.1093/gigascience/giy093

70. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics 29**, 15–21 (**2013**). https://doi.org/10.1093/bioinformatics/bts635

71. Haas, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. **Nat. Protoc. 8**, 1494–1512 (**2013**). https://doi.org/10.1038/nprot.2013.084

72. Buchfink, B., Reuter, K. & Drost, H.-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. **Nat. Methods 18**, 366–368 (**2021**). https://doi.org/10.1038/s41592-021-01101-x

73. Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B. & Wu, C. H. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. **Bioinformatics 31**, 926–932 (**2015**). https://doi.org/10.1093/bioinformatics/btu739

74. Pertea, G. & Pertea, M. GFF utilities: GffRead and GffCompare. **F1000Res. 9**, 304 (**2020**). https://doi.org/10.12688/f1000research.23297.2

75. Waterhouse, R. M. et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. **Mol. Biol. Evol. 35**, 543–548 (**2018**). https://doi.org/10.1093/molbev/msx319

76. Flynn, J. M. et al. RepeatModeler2 for automated genomic discovery of transposable element families. **Proc. Natl Acad. Sci. USA 117**, 9451–9457 (**2020**). https://doi.org/10.1073/pnas.1921046117

77. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. **Bioinformatics 26**, 841–842 (**2010**). https://doi.org/10.1093/bioinformatics/btq033

78. Wickham, H. Ggplot2: elegant graphics for data analysis 2nd edn (Springer, Cham, 2016).

79. Pedersen, B. S. & Quinlan, A. R. Mosdepth: quick coverage calculation for genomes and exomes. **Bioinformatics 34**, 867–868 (**2018**). https://doi.org/10.1093/bioinformatics/btx699

80. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. **Nat. Methods 9**, 357–359 (**2012**). https://doi.org/10.1038/nmeth.1923

81. Danecek, P. et al. Twelve years of SAMtools and BCFtools. **Gigascience 10**, giab008 (**2021**). https://doi.org/10.1093/gigascience/giab008

82. Danecek, P. et al. The variant call format and VCFtools. **Bioinformatics 27**, 2156–2158 (**2011**). https://doi.org/10.1093/bioinformatics/btr330

83. Grayson, P., Wright, A., Garroway, C. J. & Docker, M. F. SexFindR: a computational workflow to identify young and old sex chromosomes. Preprint at https://doi.org/10.1101/2022.02.21.481346 (**2022**).

84. Voichek, Y. & Weigel, D. Identifying genetic variants underlying phenotypic variation in plants without complete genomes. **Nat. Genet. 52**, 534–540 (**2020**). https://doi.org/10.1038/s41588-020-0612-7

85. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. **Am. J. Hum. Genet. 81**, 559–575 (**2007**). https://doi.org/10.1086/519795

86. Jackman, S. D. et al. ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. **Genome Res. 27**, 768–777 (**2017**). https://doi.org/10.1101/gr.214346.116

87. Li, H. Minimap2: pairwise alignment for nucleotide sequences. **Bioinformatics 34**, 3094–3100 (**2018**). https://doi.org/10.1093/bioinformatics/bty191

88. Cabanettes, F. & Klopp, C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. **PeerJ 6**, e4958 (**2018**). https://doi.org/10.7717/peerj.4958

89. Galià-Camps, C. et al. De novo genome assembly, inversion detection, and worldwide adaptation on the invasive species *Styela plicata*. **Sci. Rep**. 15, 40328 (2025). https://doi.org/10.1038/s41598-025-24574-8

90. Corbett-Detig, R. B., Cardeno, C. & Langley, C. H. Sequence-based detection and breakpoint assembly of polymorphic inversions. **Genetics 192**, 131–137 (**2012**). https://doi.org/10.1534/genetics.112.141622

91. Gamble, T. Using RAD-seq to recognize sex-specific markers and sex chromosome systems. **Mol. Ecol. 25**, 2114–2116 (**2016**). https://doi.org/10.1111/mec.13648

92. Feron, R. et al. RADsex: a computational workflow to study sex determination using restriction site-associated DNA sequencing data. **Mol. Ecol. Resour. 21**, 1715–1731 (**2021**). https://doi.org/10.1111/1755-0998.13360

93. Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A. & Cresko, W. A. Stacks: an analysis tool set for population genomics. **Mol. Ecol. 22**, 3124–3140 (**2013**). https://doi.org/10.1111/mec.12354

94. Camacho, C. et al. BLAST+: architecture and applications. **BMC Bioinformatics 10**, 421 (**2009**). https://doi.org/10.1186/1471-2105-10-421

95. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. **Bioinformatics 26**, 589–595 (**2010**). https://doi.org/10.1093/bioinformatics/btp698

96. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. **Nat. Biotechnol. 37**, 907–915 (**2019**). https://doi.org/10.1038/s41587-019-0201-4

97. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics 26**, 139–140 (**2010**). https://doi.org/10.1093/bioinformatics/btp616

98. Koutsouveli, V. Sex, molecules, and gene control. Ecophysiological and evolutionary aspects of key sponge species from Antarctic shallow waters and the deep sea. PhD thesis, Hellenic Centre for Marine Research (2020).

99. Taboada, S. et al. Connectivity and adaptation patterns of the deep-sea ground-forming sponge Geodia hentscheli across its entire distribution. **Mol. Biol. Evol. 42**, msaf145 (**2025**). https://doi.org/10.1093/molbev/msaf145

100. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. **Bioinformatics 27**, 863–864 (**2011**). https://doi.org/10.1093/bioinformatics/btr026

101. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. **Bioinformatics 28**, 3211–3217 (**2012**). https://doi.org/10.1093/bioinformatics/bts611

102. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. **Nucleic Acids Res. 41**, D590–D596 (**2013**). https://doi.org/10.1093/nar/gks1219

103. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. **BMC Bioinformatics 12**, 323 (**2011**). https://doi.org/10.1186/1471-2105-12-323

104. Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics 21**, 3674–3676 (**2005**). https://doi.org/10.1093/bioinformatics/bti610

105. Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. **Bioinformatics 36**, 2628–2629 (**2020**). https://doi.org/10.1093/bioinformatics/btz931

106. Shen, S. et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. **Proc. Natl Acad. Sci. USA 111**, E5593–E5601 (**2014**). https://doi.org/10.1073/pnas.1419161111

107. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. **Genome Biol. 20**, 238 (**2019**). https://doi.org/10.1186/s13059-019-1832-y

108. El Hilali, S. & Copley, R. R. macrosyntR: drawing automatically ordered Oxford grids from standard genomic files in R. Preprint at https://doi.org/10.1101/2023.01.26.525673 (**2023**).

109. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. **Brief. Bioinform. 20**, 1160–1166 (**2019**). https://doi.org/10.1093/bib/bbx108

110. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. **BMC Evol. Biol. 10**, 210 (**2010**). https://doi.org/10.1186/1471-2148-10-210

111. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. **Mol. Biol. Evol. 26**, 1641–1650 (**2009**). https://doi.org/10.1093/molbev/msp077

112. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. **Mol. Biol. Evol. 37**, 1530–1534 (**2020**). https://doi.org/10.1093/molbev/msaa015

113. Sun, J. et al. OrthoVenn3: an integrated platform for exploring and visualizing orthologous data across genomes. **Nucleic Acids Res. 51**, W397–W403 (**2023**). https://doi.org/10.1093/nar/gkad313

Supplementary Material

114. Barrera-Redondo, J., Lotharukpong, J. S., Drost, H.-G. & Coelho, S. M. Uncovering gene-family founder events during major evolutionary transitions in animals, plants and fungi using GenEra. **Genome Biol. 24**, 54 (**2023**). https://doi.org/10.1186/s13059-023-02895-z

115. Imsiecke, G., Custodio, M., Borojevic, R., Steffen, R., Moustafa, M. A. & Müller, W. E. G. Genome size and chromosomes in marine sponges (Suberites domuncula, Geodia cydonium). **Cell Biol. Int. 19**, 995–1000 (**1995**). https://doi.org/10.1006/cbir.1995.1041

116. Ishijima, J., Iwabe, N., Masuda, Y., Watanabe, Y. & Matsuda, Y. Sponge cytogenetics—mitotic chromosomes of ten species of freshwater sponge. **Zoolog. Sci. 25**, 480–486 (**2008**). https://doi.org/10.2108/zsj.25.480

117. Imsiecke, G., Pascheberg, U. & Müller, W. E. G. Preparation and karyotype analysis of mitotic chromosomes of the freshwater sponge Spongilla lacustris. **Chromosoma 102**, 724–727 (**1993**). https://doi.org/10.1007/s10709-016-9895-0

118. Sember, A., Nguyen, P., Perez, M. F., Altmanová, M., Ráb, P. & Cioffi, M. B. Multiple sex chromosomes in teleost fishes from a cytogenetic perspective: state of the art and future challenges. **Philos. Trans. R. Soc. Lond. B Biol. Sci. 376**, 20200098 (**2021**). https://doi.org/10.1098/rstb.2020.0098

119. Patton, J. L. & Gardner, A. L. Parallel evolution of multiple sex-chromosome systems in the phyllostomatid bats, Carollia and Choeroniscus. **Experientia 27**, 105–106 (**1971**). https://doi.org/10.1007/BF02137768

120. Maculatus, X. Molecular analysis of the sex-determining region of the Y chromosome in the platyfish Xiphophorus maculatus. PhD thesis, Univ. Würzburg (**2005**).

121. Śliwińska, E. B., Martyka, R. & Tryjanowski, P. Evolutionary interaction between W/Y chromosome and transposable elements. **Genetica 144**, 267–278 (**2016**). https://doi.org/10.1007/s10709-016-9895-0

122. Boonanuntanasarn, S., Jangprai, A. & Na-Nakorn, U. Transcriptomic analysis of female and male gonads in juvenile snakeskin gourami (Trichopodus pectoralis). **Sci. Rep. 10**, 5240 (**2020**). https://doi.org/10.1038/s41598-020-61738-0

123. Chiang, V. S. C., DeRosa, H., Park, J. H. & Hunter, R. G. The role of transposable elements in sexual development. **Front. Behav. Neurosci. 16**, 923732 (**2022**). https://doi.org/10.3389/fnbeh.2022.923732

124. Tian, C. et al. DExD/H-box RNA helicase genes are differentially expressed between males and females during the critical period of male sex differentiation in channel catfish. **Comp. Biochem. Physiol. Part D Genomics Proteomics 22**, 109–119 (**2017**). https://doi.org/10.1016/j.cbd.2017.02.008

125. Boulé, J.-B. & Zakian, V. A. Roles of Pif1-like helicases in the maintenance of genomic stability. **Nucleic Acids Res. 34**, 4147–4153 (**2006**). https://doi.org/10.1093/nar/gkl561

126. Wolff, J. R. & Zarkower, D. Somatic sexual differentiation in Caenorhabditis elegans. **Curr. Top. Dev. Biol. 83**, 1–39 (**2008**). https://doi.org/10.1016/S0070-2153(08)00401-8

127. Cocquet, J. et al. Deficiency in the multicopy Sycp3-like X-linked genes Slx and Slxl1 causes major defects in spermatid differentiation. **Mol. Biol. Cell 21**, 3497–3505 (**2010**). https://doi.org/10.1091/mbc.E10-07-0601

128. García-Rincón, J., Darszon, A. & Beltrán, C. Speract, a sea urchin egg peptide that regulates sperm motility, also stimulates sperm mitochondrial metabolism. **Biochim. Biophys. Acta 1857**, 415–426 (**2016**). https://doi.org/10.1016/j.bbabio.2016.01.003

129. Hansbrough, J. R. & Garbers, D. L. Speract. Purification and characterization of a peptide associated with eggs that activates spermatozoa. **J. Biol. Chem. 256**, 1447–1452 (**1981**).

130. Bellott, D. W. & Page, D. C. Dosage-sensitive functions in embryonic development drove the survival of genes on sex-specific chromosomes in snakes, birds, and mammals. **Genome Res. 31**, 198–210 (**2021**). https://doi.org/10.1101/gr.268516.120

131. Singh, P. et al. CDK2 kinase activity is a regulator of male germ cell fate. **Development 146**, dev180273 (**2019**). https://doi.org/10.1242/dev.180273

132. Craske, B. & Welburn, J. P. I. Leaving no-one behind: how CENP-E facilitates chromosome alignment. **Essays Biochem. 64**, 313–324 (**2020**). https://doi.org/10.1042/EBC20190073

133. Rajičić, M. et al. B chromosomes' sequences in yellow-necked mice Apodemus flavicollis—exploring the transcription. **Life 12**, 50 (**2022**). https://doi.org/10.3390/life12010050

134. Lin, X. et al. Comprehensive transcriptome analysis reveals sex-specific alternative splicing events in zebrafish gonads. **Life 12**, 1441 (**2022**). https://doi.org/10.3390/life12091441

135. Wang, X. R. et al. Evidence for parallel evolution of a gene involved in the regulation of spermatogenesis. **Proc. R. Soc. B 284**, 20170324 (**2017**). https://doi.org/10.1098/rspb.2017.0324

136. Cui, Z. et al. Identification and functional analysis of E3 ubiquitin ligase g2e3 in Chinese tongue sole, Cynoglossus semilaevis. **Animals 14**, 2579 (**2024**). https://doi.org/10.3390/ani14172579

137. Holloway, J. K., Morelli, M. A., Borst, P. L. & Cohen, P. E. Mammalian BLM helicase is critical for integrating multiple pathways of meiotic recombination. **J. Cell Biol. 188**, 779–789 (**2010**). https://doi.org/10.1083/jcb.200909048

138. Ichijima, Y., Sin, H. S. & Namekawa, S. H. Sex chromosome inactivation in germ cells: emerging roles of DNA damage response pathways. **Cell. Mol. Life Sci. 69**, 2559–2572 (**2012**). https://doi.org/10.1007/s00018-012-0941-5

139. Zhou, A. et al. Characterizing genetic variation on the Z chromosome in Schistosoma japonicum reveals host-parasite co-evolution. **Parasit. Vectors 17**, 207 (**2024**). https://doi.org/10.1186/s13071-024-06250-4

140. Brunner, B. et al. Genomic organization and expression of the doublesex-related gene cluster in vertebrates and detection of putative regulatory regions for DMRT1. **Genomics 77**, 8–17 (**2001**). https://doi.org/10.1006/geno.2001.6615

141. Dong, J. et al. Comparative genomics studies on the dmrt gene family in fish. **Front. Genet. 11**, 563947 (**2020**). https://doi.org/10.3389/fgene.2020.563947

142. Morinha, F. et al. (R)evolution in the molecular sexing of ratite birds: identification and analysis of new candidate sex-linked markers. **Avian Biol. Res. 8**, 145–159 (**2015**). https://doi.org/10.3184/175815515X14380769320720

143. Allen, M. P. et al. Growth arrest-specific gene 6 (Gas6)/adhesion related kinase (Ark) signaling promotes gonadotropin-releasing hormone neuronal survival via extracellular signal-regulated kinase (ERK) and Akt. **Mol. Endocrinol. 13**, 191–201 (**1999**). https://doi.org/10.1210/mend.13.2.0230

144. Pask, A., Renfree, M. B. & Graves, J. A. M. The human sex-reversing ATRX gene has a homologue on the marsupial Y chromosome, ATRY: implications for the evolution of mammalian sex determination. **Proc. Natl Acad. Sci. USA 97**, 13198–13202 (**2000**). https://doi.org/10.1073/pnas.230424497

145. Nagahama, Y., Chakraborty, T., Paul-Prasanth, B., Ohta, K. & Nakamura, M. Sex determination, gonadal sex differentiation, and plasticity in vertebrate species. **Physiol. Rev. 101**, 1237–1308 (**2021**). https://doi.org/10.1152/physrev.00044.2019

146. Nef, S. et al. Testis determination requires insulin receptor family function in mice. **Nature 426**, 291–295 (**2003**). https://doi.org/10.1038/nature02059

147. Rodriguez-Mari, A. & Postlethwait, J. H. Zebrafish provides a novel model for lymphatic vascular research. In **Methods Cell Biol. 105**, 461–490 (Academic Press, 2011). https://doi.org/10.1016/B978-0-12-381320-6.00020-5

148. Naumova, O. Y. et al. Male pseudohermaphroditism: a case study of 46,XY disorder of sexual development using whole-exome sequencing. **Clin. Case Rep. 8**, 2889–2894 (**2020**). https://doi.org/10.1002/ccr3.3286

149. Zhang, J. et al. Spermatogenesis-associated 48 is essential for spermatogenesis in mice. **Andrologia 50**, e13027 (**2018**). https://doi.org/10.1111/and.13027

150. Sasagawa, Y., Yamanaka, K., Saito-Sasagawa, Y. & Ogura, T. Caenorhabditis elegans UBX cofactors for CDC-48/p97 control spermatogenesis. **Genes Cells 15**, 1201–1215 (**2010**). https://doi.org/10.1111/j.1365-2443.2010.01454.x

151. Tang, Y., Chen, J. Y., Ding, G. H. & Lin, Z. H. Analyzing the gonadal transcriptome of the frog Hoplobatrachus rugulosus to identify genes involved in sex development. **BMC Genomics 22**, 539 (**2021**). https://doi.org/10.1186/s12864-021-07879-6

152. Fernandez-Capetillo, O. et al. H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. **Dev. Cell 4**, 497–508 (**2003**). https://doi.org/10.1016/s1534-5807(03)00093-5

153. Yano, A., Nicol, B., Guerin, A. & Guiguen, Y. The duplicated rainbow trout (Oncorhynchus mykiss) T-box transcription factors 1, tbx1a and tbx1b, are up-regulated during testicular development. **Mol. Reprod. Dev. 78**, 172–180 (**2011**). https://doi.org/10.1002/mrd.21279

154. Hattori, T. et al. E6-AP/UBE3A protein acts as a ubiquitin ligase toward SOX9 protein. **J. Biol. Chem. 288**, 35138–35148 (**2013**). https://doi.org/10.1074/jbc.M113.486795

155. Modzelewski, A. J., Holmes, R. J., Hilz, S., Grimson, A. & Cohen, P. E. AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline. **Dev. Cell 23**, 251–264 (**2012**). https://doi.org/10.1016/j.devcel.2012.07.003

156. Schroeder, A. L., Metzger, K. J., Miller, A. & Rhen, T. A novel candidate gene for temperature-dependent sex determination in the common snapping turtle. **Genetics 203**, 557–571 (**2016**). https://doi.org/10.1534/genetics.115.182840

157. Steiner, R., Ever, L. & Don, J. MEIG1 localizes to the nucleus and binds to meiotic chromosomes of spermatocytes as they initiate meiosis. **Dev. Biol. 216**, 635–645 (**1999**). https://doi.org/10.1006/dbio.1999.9520

158. Oliveira, C. F. et al. Foxn1 and Prkdc genes are important for testis function: evidence from nude and scid adult mice. **Cell Tissue Res. 380**, 615–625 (**2020**). https://doi.org/10.1007/s00441-019-03165-w

159. Ren, Y. et al. dmrt3, nom1, abce1, and pkmyt1 play key roles in gonadal sex determination in Acrossocheilus fasciatus. **Aquac. Int. 31**, 317–332 (**2023**). https://doi.org/10.1007/s10499-022-00976-7

160. Cerván-Martín, M. et al. Effect and in silico characterization of genetic variants associated with severe spermatogenic disorders in a large Iberian cohort. **Andrology 9**, 1151–1165 (**2021**). https://doi.org/10.1111/andr.13009

161. Wang, P. J., McCarrey, J. R., Yang, F. & Page, D. C. An abundance of X-linked genes expressed in spermatogonia. **Nat. Genet. 27**, 422–426 (**2001**). https://doi.org/10.1038/86927

162. Lewin, T. D., Liao, I. J. Y. & Luo, Y. J. Annelid comparative genomics and the evolution of massive lineage-specific genome rearrangement in bilaterians. **Mol. Biol. Evol. 41**, msae172 (**2024**). https://doi.org/10.1093/molbev/msae172

163. Wright, A. E., Dean, R., Zimmer, F. & Mank, J. E. How to make a sex chromosome. **Nat. Commun. 7**, 12087 (**2016**). https://doi.org/10.1038/ncomms12087

164. Pokorná, M. & Kratochvíl, L. Phylogeny of sex-determining mechanisms in squamate reptiles: are sex chromosomes an evolutionary trap? **Zool. J. Linn. Soc. 156**, 168–183 (**2009**). https://doi.org/10.1111/j.1096-3642.2008.00481.x

165. Toups, M. A. & Vicoso, B. The X chromosome of insects likely predates the origin of class Insecta. **Evolution 77**, 2504–2511 (**2023**). https://doi.org/10.1093/evolut/qpad169

166. Zhou, Q. et al. Complex evolutionary trajectories of sex chromosomes across bird taxa. **Science 346**, 1246338 (**2014**). https://doi.org/10.1126/science.1246338

167. Cortez, D. et al. Origins and functional evolution of Y chromosomes across mammals. **Nature 508**, 488–493 (**2014**). https://doi.org/10.1038/nature13151

168. Handley, L.-J. L., Ceplitis, H. & Ellegren, H. Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. **Genetics 167**, 367–376 (**2004**). https://doi.org/10.1534/genetics.167.1.367

169. Rossi, E. et al. Independent origins of spicules reconcile palaeontological and molecular evidence of sponge evolutionary history. **Science Advances (2025)**.

170. Chintapalli, S. V. et al. Reevaluation of the evolutionary events within recA/RAD51 phylogeny. **BMC Genomics 14**, 240 (**2013**). https://doi.org/10.1186/1471-2164-14-240

171. Xu, E. Y., Moore, F. L. & Reijo Pera, R. A. A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in metazoans. **Proc. Natl Acad. Sci. USA 98**, 7414–7419 (**2001**). https://doi.org/10.1073/pnas.131090498

172. Fedry, J. et al. Evolutionary diversification of the HAP2 membrane insertion motifs to drive gamete fusion across eukaryotes. **PLoS Biol. 16**, e2006357 (**2018**). https://doi.org/10.1371/journal.pbio.2006357

173. Matson, C. K. & Zarkower, D. Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. **Nat. Rev. Genet. 13**, 163–174 (**2012**). https://doi.org/10.1038/nrg3161

174. Carvalho-Santos, Z., Azimzadeh, J., Pereira-Leal, J. B. & Bettencourt-Dias, M. Tracing the origins of centrioles, cilia, and flagella. **J. Cell Biol. 194**, 165–175 (**2011**). https://doi.org/10.1083/jcb.201011152

175. Bererhi, B. et al. Effect of MHC and inbreeding on disassortative reproduction: a data revisit, extension and inclusion of fertilization in sand lizards. **Ecol. Evol. 13**, e9934 (**2023**). https://doi.org/10.1002/ece3.9934

176. Tilly, J. L. Commuting the death sentence: how oocytes strive to survive. **Nat. Rev. Mol. Cell Biol. 2**, 838–848 (**2001**). https://doi.org/10.1038/35099086

177. Paps, J., Rossi, M. E., Bowles, A. M. C. & Álvarez-Presas, M. Assembling animals: trees, genomes, cells, and contrast to plants. **Front. Ecol. Evol. 11**, 1185566 (**2023**). https://doi.org/10.3389/fevo.2023.1185566

178. Wang, Y. et al. Identification of a candidate sex determination region and sex-specific molecular markers based on whole-genome re-sequencing in the sea star Asterias amurensis. **DNA Res. 32**, dsaf026 (**2025**). https://doi.org/10.1093/dnares/dsaf003

179. Kratochvíl, L. et al. Expanding the classical paradigm: what we have learnt from vertebrates about sex chromosome evolution. **Philos. Trans. R. Soc. B 376**, 20200098 (**2021**). doi: 10.1098/rstb.2020.0097

180. Pšenička, T. et al. Sex chromosome turnovers and stability in snakes. **Mol. Biol. Evol. 42**, msae255 (**2025**). https://doi.org/10.1093/molbev/msae255

181. Shi, C. et al. A ZZ/ZW sex chromosome system in cephalochordate amphioxus. Genetics 214, 617–622 (2020). https://doi.org/10.1534/genetics.120.303051

182. Traut, W. The evolution of sex chromosomes in insects: differentiation of sex chromosomes in flies and moths. **Eur. J. Entomol. 96**, 227–235 (**1999**).

183. Rödelsperger, C. Comparative genomics of sex, chromosomes, and sex-linked gene transcription in Caenorhabditis elegans and other nematodes. In **Comparative Genomics** (eds Agier, N. et al.) 455–472 (Humana, New York, 2018). https://doi.org/10.1007/978-1-4939-7463-4_16

184. Wang, Y., Gasser, R. B., Charlesworth, D. & Zhou, Q. Evolution of sexual systems, sex chromosomes and sex-linked gene transcription in flatworms and roundworms. **Nat. Commun. 13**, 7107 (**2022**). https://doi.org/10.1038/s41467-022-30578-z

185. Han, W. et al. Ancient homomorphy of molluscan sex chromosomes sustained by reversible sex-biased genes and sex determiner translocation. Nat. Ecol. Evol. 6, 1891–1906 (2022). https://doi.org/10.1038/s41559-022-01898-6

186. Korablev, V. P., Radashevsky, V. I. & Manchenko, G. P. The XX-XY (male-heterogametic) sex chromosome system in Polydora curiosa (Polychaeta: Spionidae). **Ophelia 51**, 193–201 (**1999**). https://doi.org/10.1080/00785326.1999.10409409

187. Pratlong, M. et al. Evidence for a genetic sex determination in Cnidaria, the Mediterranean red coral (Corallium rubrum). **R. Soc. Open Sci. 4**, 160880 (**2017**). https://doi.org/10.1098/rsos.160880

188. Pita, L. et al. The chromosomal genome sequence of the kidney sponge Chondrosia reniformis (Nardo, 1847) and its associated microbial metagenome sequences. **Wellcome Open Res. 10**, 72 (**2025**). https://doi.org/10.12688/wellcomeopenres.24166.1

189. Riesgo, A. et al. The chromosomal genome sequence of the sponge Oscarella lobularis (Schmidt, 1862) (Porifera, Homoscleromorpha) and its associated microbial metagenome sequences. **Wellcome Open Res. 10**, 71 (**2025**). https://doi.org/10.12688/wellcomeopenres.24187.1

190. Steindler, L. et al. The chromosomal genome sequence of the stone sponge Petrosia ficiformis (Poiret, 1789) and its associated microbial metagenome sequences. **Wellcome Open Res. 10**, 70 (**2025**). https://doi.org/10.12688/wellcomeopenres.24743.1

Supplementary Figures and Tables

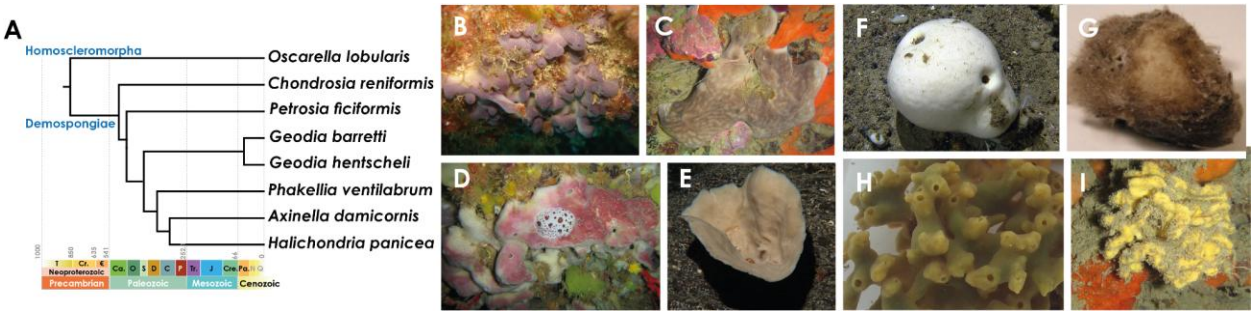

Figure S1. Sponge species in the study. A. Phylogenetic framework for the selected sponge species. B. *Oscarella lobularis* (picture by Ana Riesgo). C. *Chondrosia reniformis* (picture by Marta Turon). D. *Petrosia ficiformis*. E. *Phakellia ventilabrum* (picture by Bernard Picton). F. *Geodia barretti*. G. *Geodia hentscheli* (picture by Paco Cárdenas). H. *Halichondria panicea* (picture by Lars Kumala). I. *Axinella damicornis* (picture by Marta Turon).

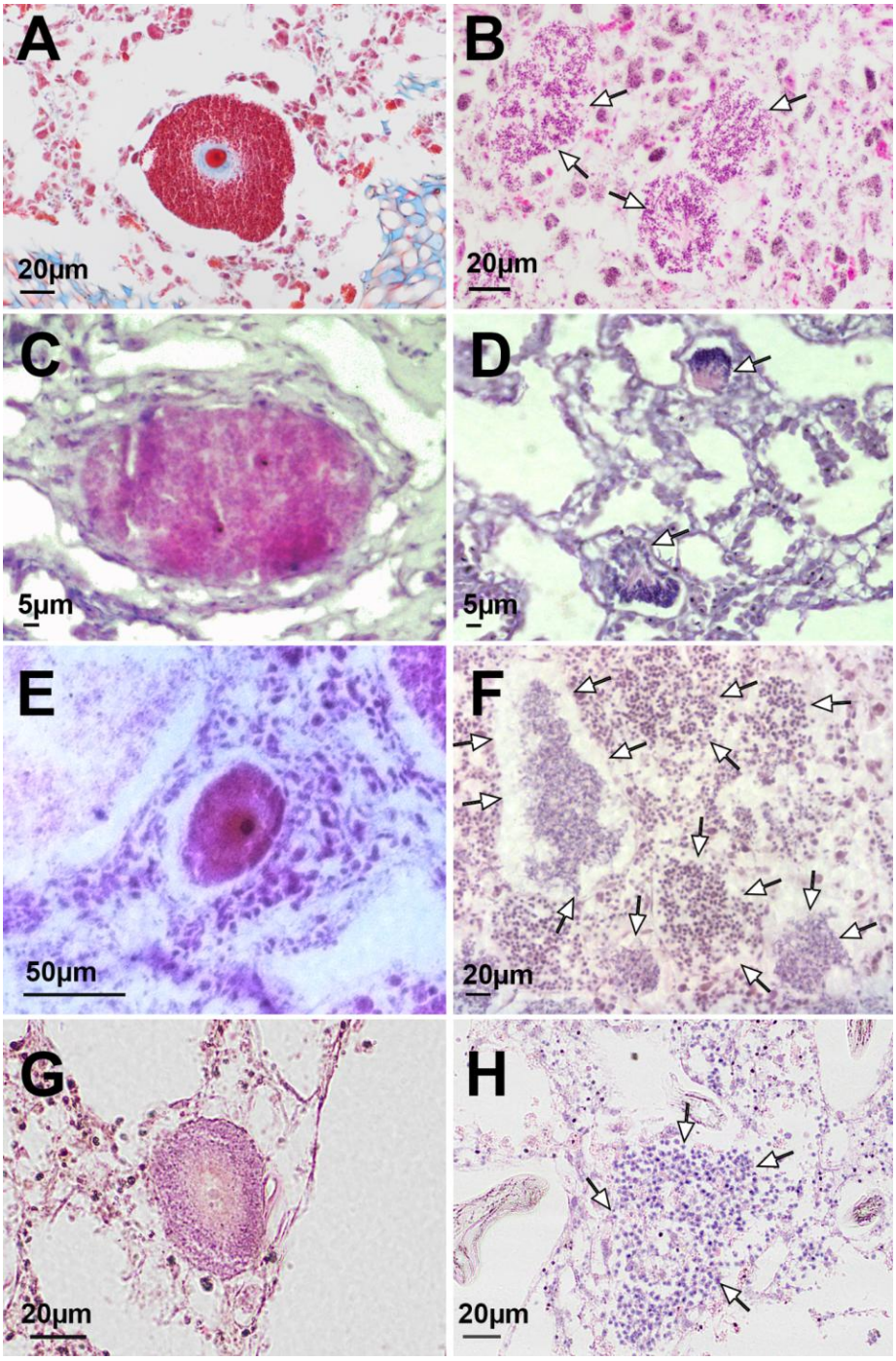

Figure S2. Histological sections of gametes within the tissues of sponges. **A–B.** Oocytes and spermatic cysts (arrows) of *Petrosia ficiformis*. **C–D.** Oocytes and spermatic cysts (arrows) of *Oscarella lobularis*. **E–F.** Oocytes and spermatic cysts (arrows) of *Halichondria panicea*. **G–H.** Oocytes and spermatic cysts (arrows) of *Axinella damicornis*.

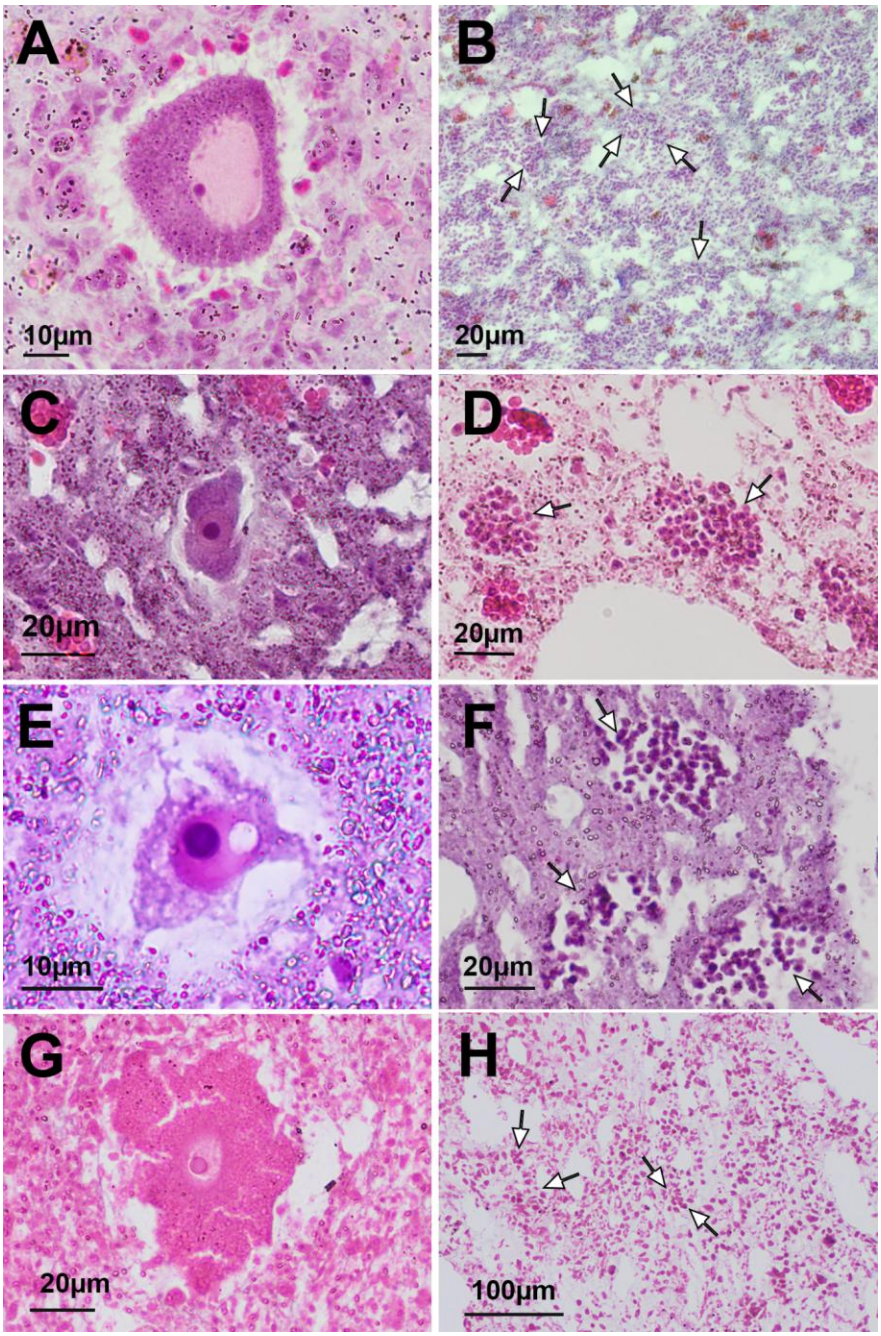

Figure S3. Histological sections of gametes within the tissues of sponges. **A–B.** Oocytes and spermatic cysts (arrows) of *Chondrosia reniformis*. **C–D.** Oocytes and spermatic cysts (arrows) of *Geodia barretti*. **E–F.** Oocytes and spermatic cysts (arrows) of *Geodia hentscheli*. **G–H.** Oocytes and spermatic cysts (arrows) of *Phakellia ventilabrum*.

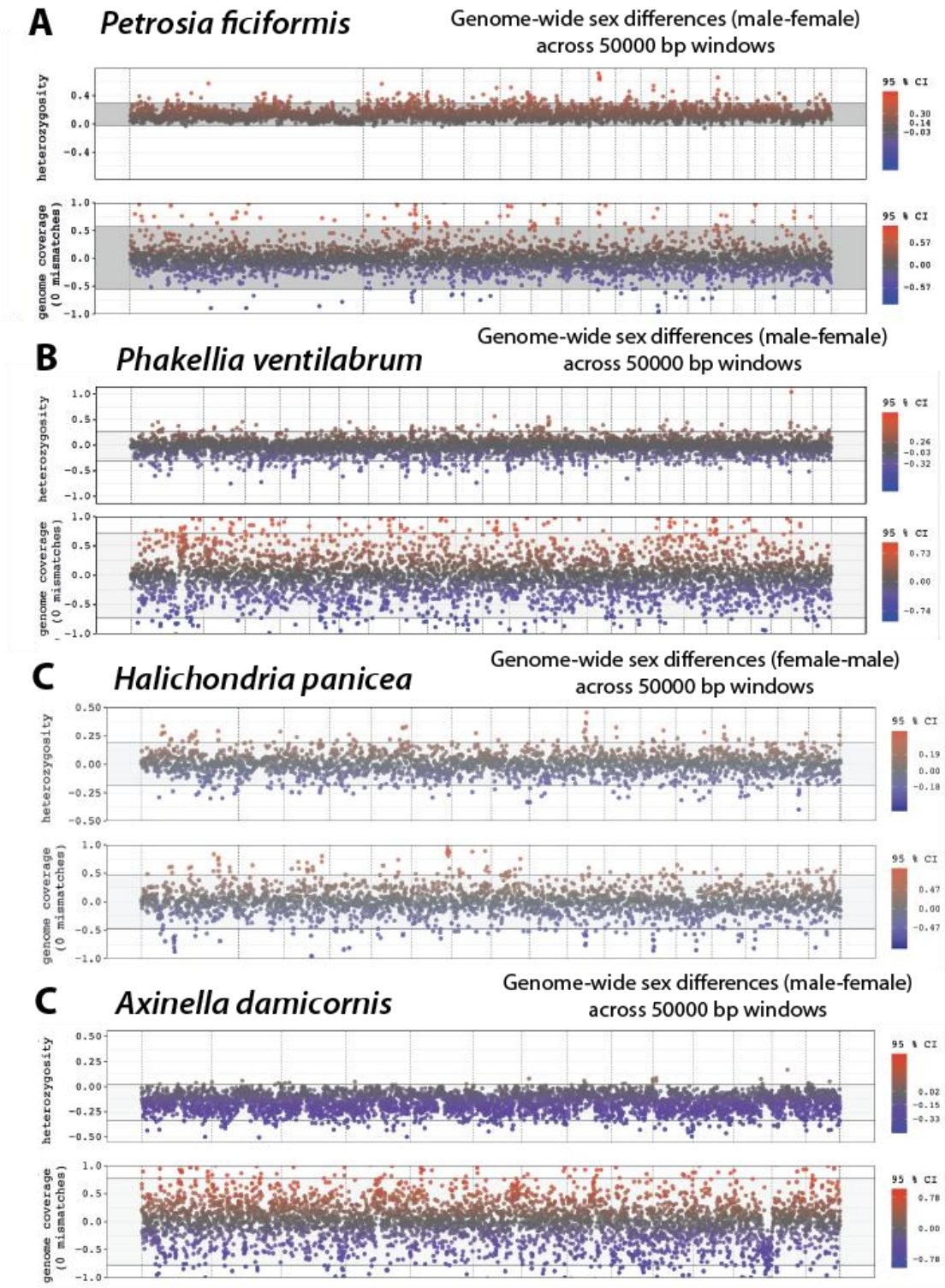

Figure S4. Data from FindZX pipeline (Sigeman et al. 2022) for species for which we failed to identify sex chromosomes. Sex differences (female–male) in heterozygosity and genome coverage (50,000 bp windows). The grey background marks the 95% confidence interval (CI), calculated with bootstrapping from the genome-wide mean value for each metric expected under no sex differences, and the colour scale indicates significant bias to females (red) and to males (blue).

Genome coverage was calculated with 0 mismatches when mapping. **A.** *Petrosia ficiformis*. **B.** *Phakellia ventilabrum*. **C.** *Halichondria panicea*. **D.** *Axinella damicornis*.

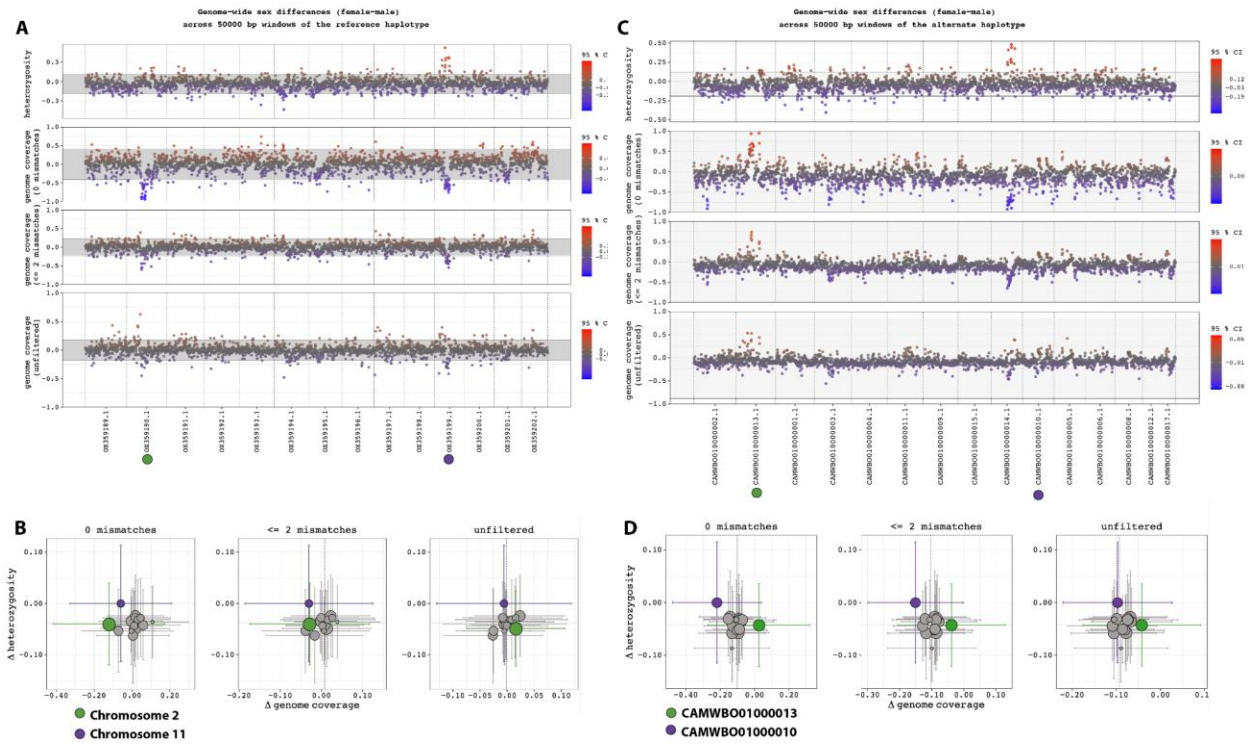

Figure S5. Data from FindZX pipeline (Sigeman et al. 2022) for the reference haplotype (A–B) and the alternate haplotype (C–D) in *Chondrosia reniformis*. Sex differences (female–male) in heterozygosity and genome coverage on 50,000 bp windows calculated with 0, 2 and unfiltered (A–C) mismatches when mapping. The grey background marks the 95% confidence interval (CI), calculated with bootstrapping from the genome-wide mean value for each metric expected under no sex differences, and the colour scale indicates significant bias to females (red) and to males (blue). (B–D). Mean sex differences (female–male) per chromosome on 50,000 bp windows.

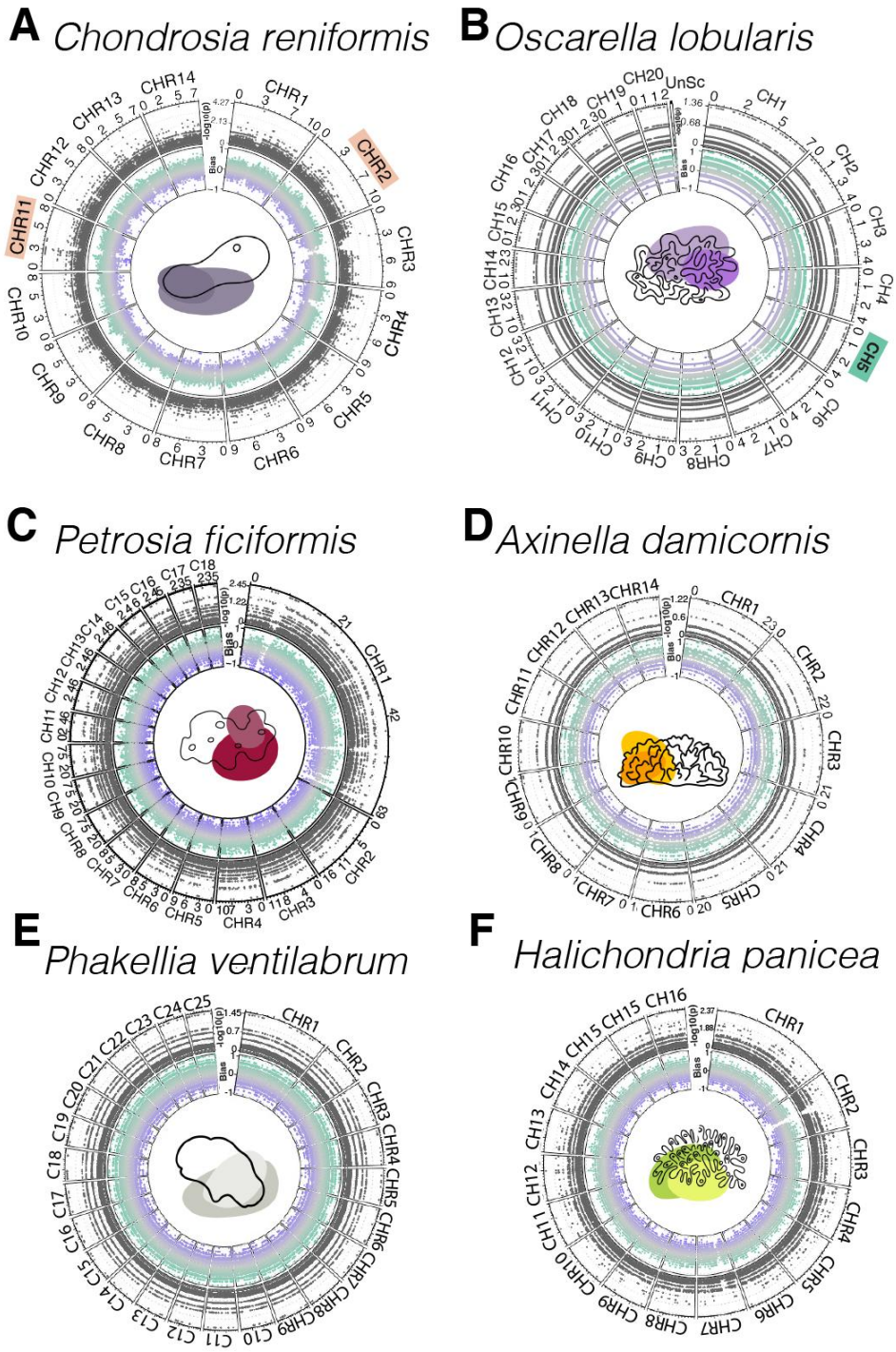

Figure S6. Circos plots showing separated chromosomes (and not unplaced scaffolds) showing in the top track the sex-bias of markers obtained with RADSEX, with a value of 1 if the marker is present in all males (green) and no females, and -1 if the marker is present in all females (purple) and no males. The bottom track shows the probability of association with sex (chi-squared test, using Bonferroni correction).

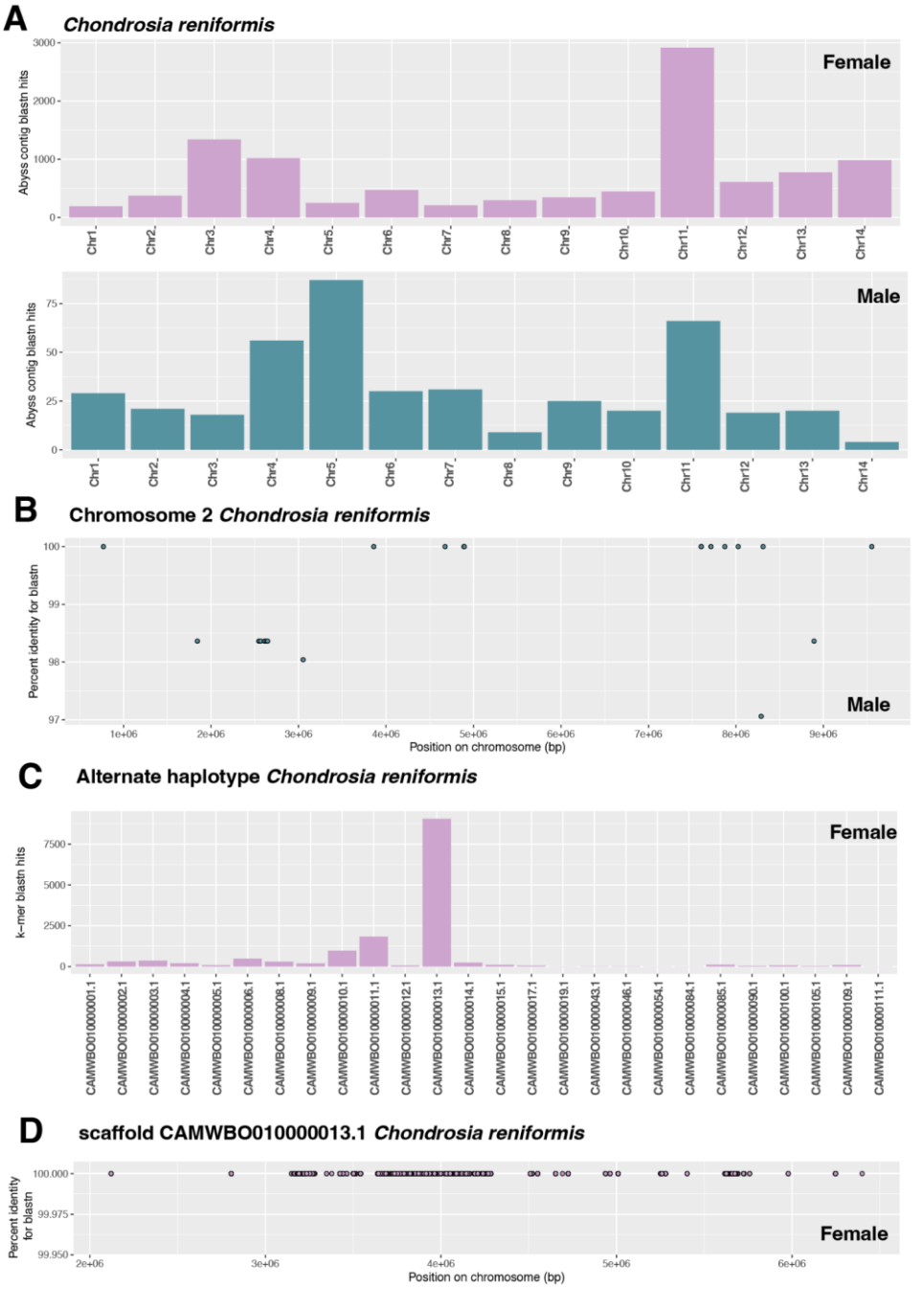

Figure S7. Sex-specific k-mer distribution on *Chondrosia reniformis*. **A.** Number of sex−specific k−mers from kmerGWAS in *Chondrosia reniformis* assembled in contigs with abyss and linked to the reference genome with blastn. **B.** Distribution of male−specific k−mers from kmerGWAS across *Chondrosia reniformis chromosome* 2 from blastn. **C.** Number of female−specific k−mers from kmerGWAS in Chondrosia reniformis assembled in contigs with abyss and linked to the alternate haplotype genome with blastn. **D.** Distribution of female−specific k−mers from kmerGWAS across *Chondrosia reniformis chromosome* scaffold CAMWBBO010000013.1, homologous to the reference chromosome 2 from blastn.

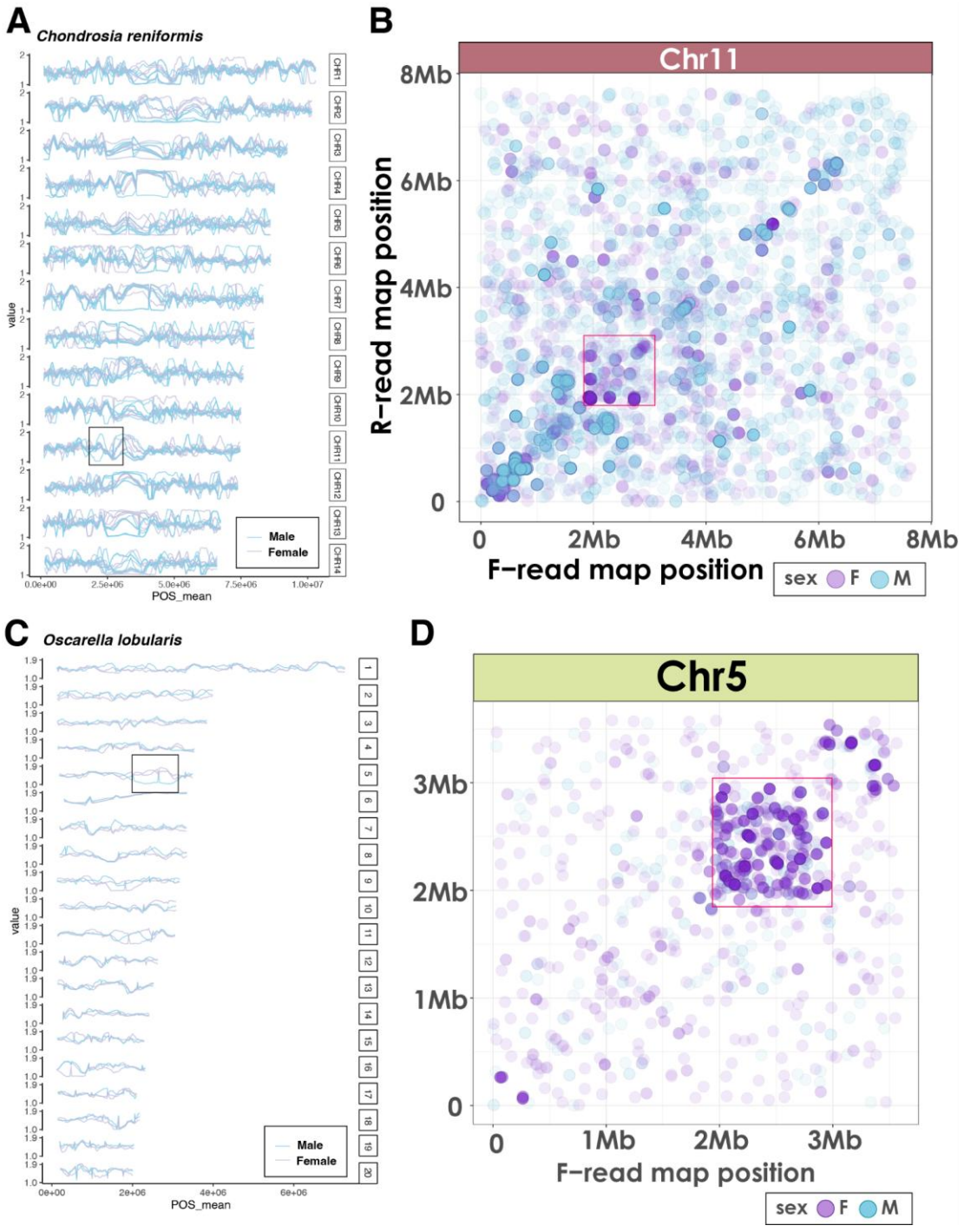

Figure S8. Average genotype for each specimen (5,000 bp windows). Values are as follows: 0 = fixation of reference SNPs, 1 = non-fixation, 2 = fixation of alternative SNPs in *Chondrosia reniformis* chromosomes (A) and *Oscarella lobularis* chromosomes (C). POS_mean= mean position in Mb. (B, D). Position of paired-end reads mapping collinearly (F-F / R-R) in chromosome 11 of *Chondrosia reniformis* (B) and Chromosome 5 in *Oscarella lobularis* (D).

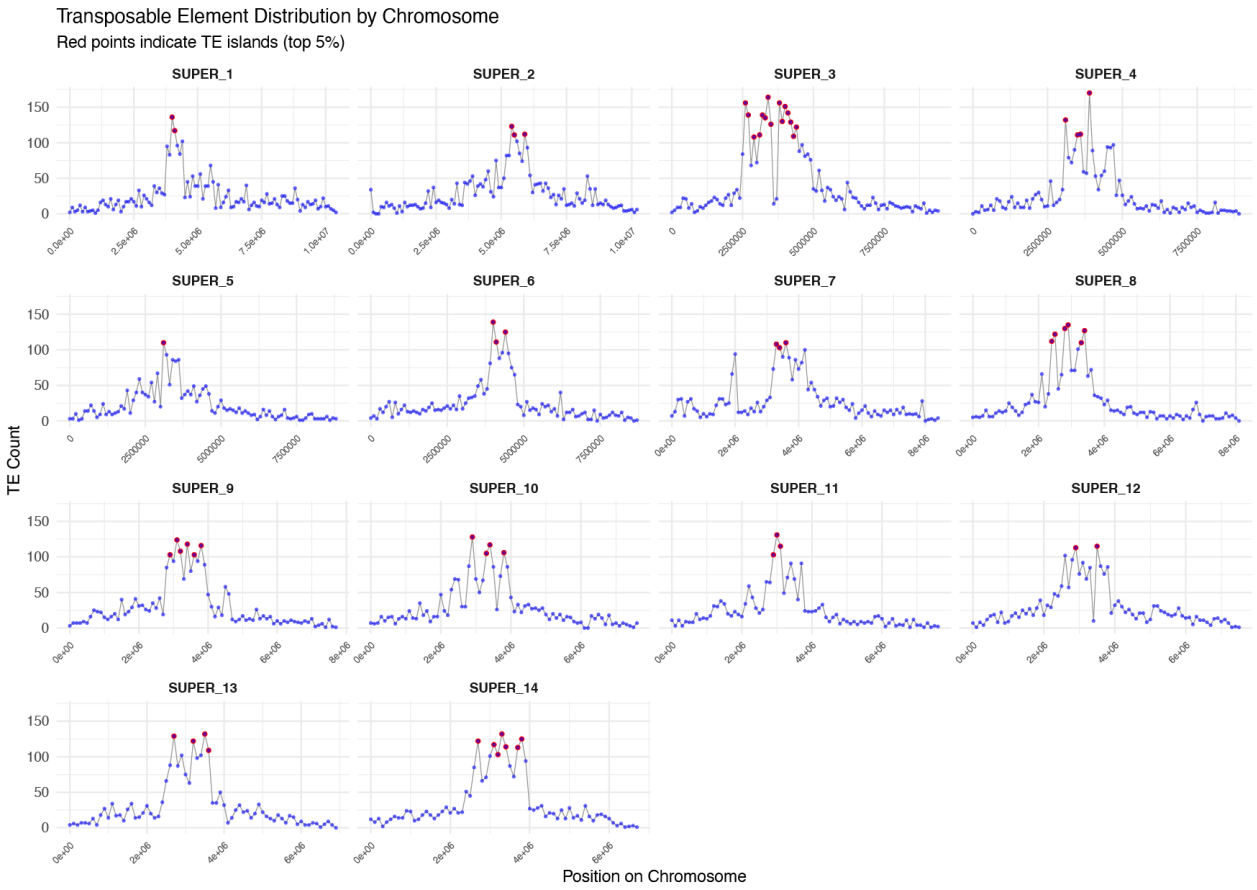

Figure S9. Number of counts of transposable elements (TEs) per Mb in the different chromosomes of *Chondrosia reniformis*. Red dots indicate TE islands (top 5%).

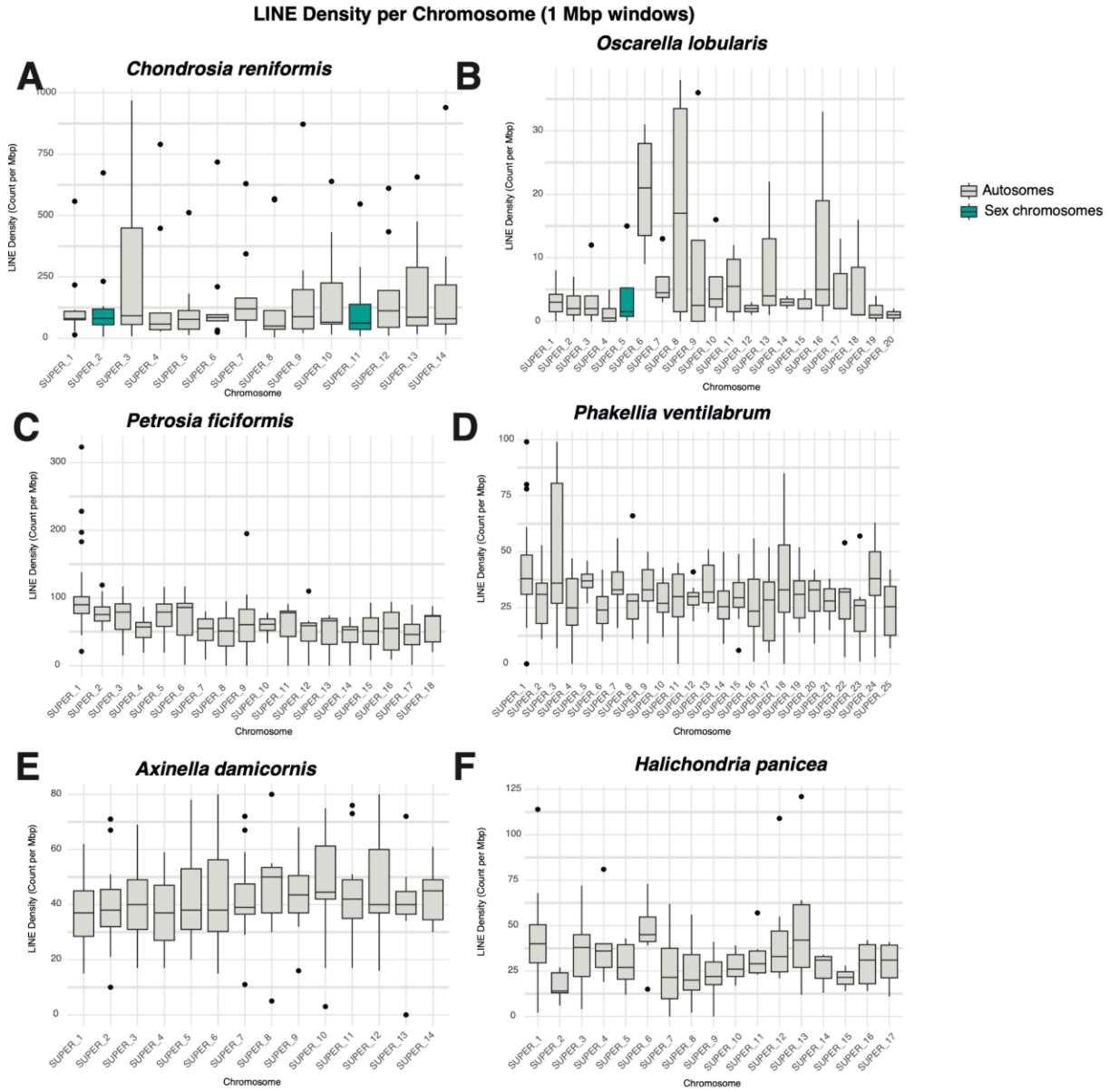

Figure S10. Average (and standard deviation) of counts of LINEs per Mb in the different chromosomes of the sponges.

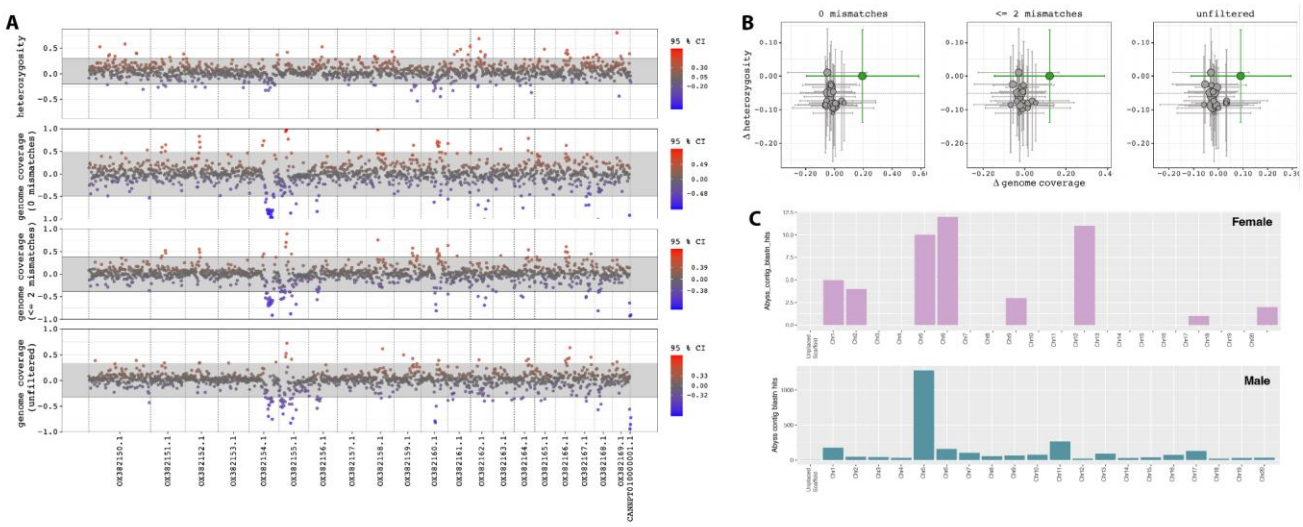

Figure S11. Data from FindZX pipeline (Sigeman et al. 2022) for the reference haplotype (A–B) in *Oscarella lobularis*. Sex differences (female–male) in heterozygosity and genome coverage on 50,000 bp windows calculated with 0, 2 and unfiltered (A) mismatches when mapping. The grey background marks the 95% confidence interval (CI), calculated with bootstrapping from the genome-wide mean value for each metric expected under no sex differences, and the colour scale indicates significant bias to females (red) and to males (blue). (B). Mean sex differences (female–male) per chromosome on 50,000 bp windows. (C). Sex-specific k-mer distribution on *Oscarella lobularis*. `Number of sex-specific k-mers from kmerGWAS in` *Oscarella lobularis* assembled in contigs with abyss and linked to the reference genome with blastn.

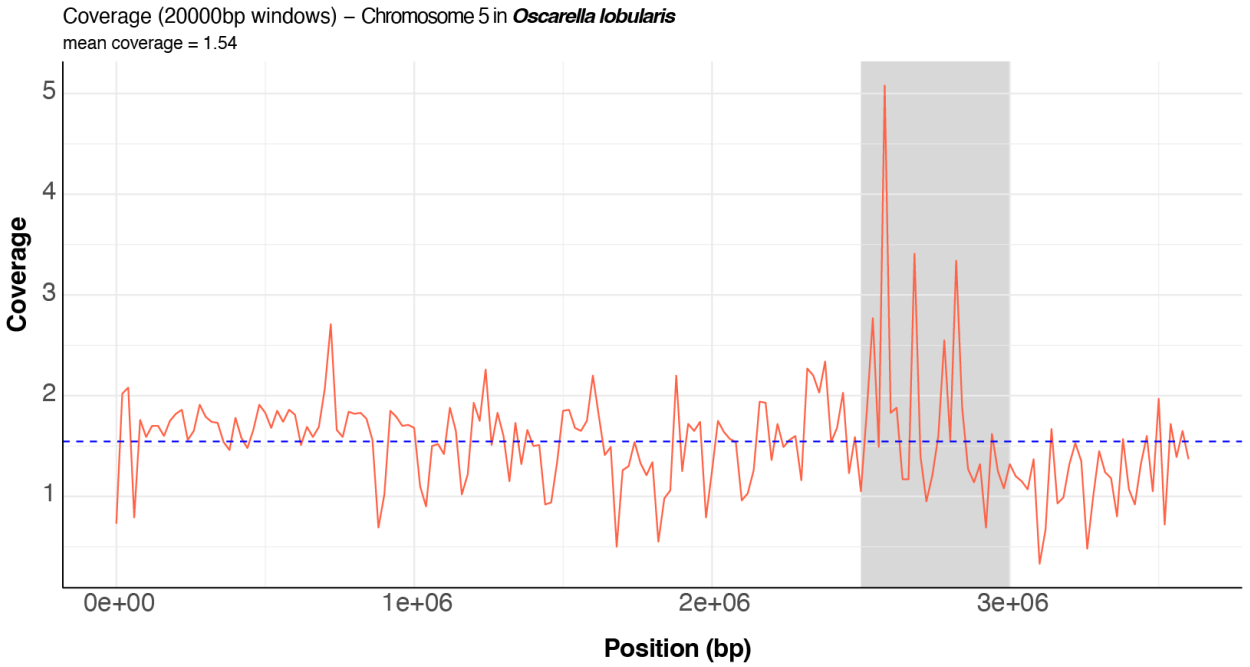

Figure S12. Coverage of male reads on sex chromosome 5 of *Oscarella lobularis*. The reads used for the mapping were those from male individual Olob13 that did not map to the draft female genome assembly done with only reads from female individuals (Supplementary Text Section 2). Note the higher coverage around the sex-determining region identified by FindZX.

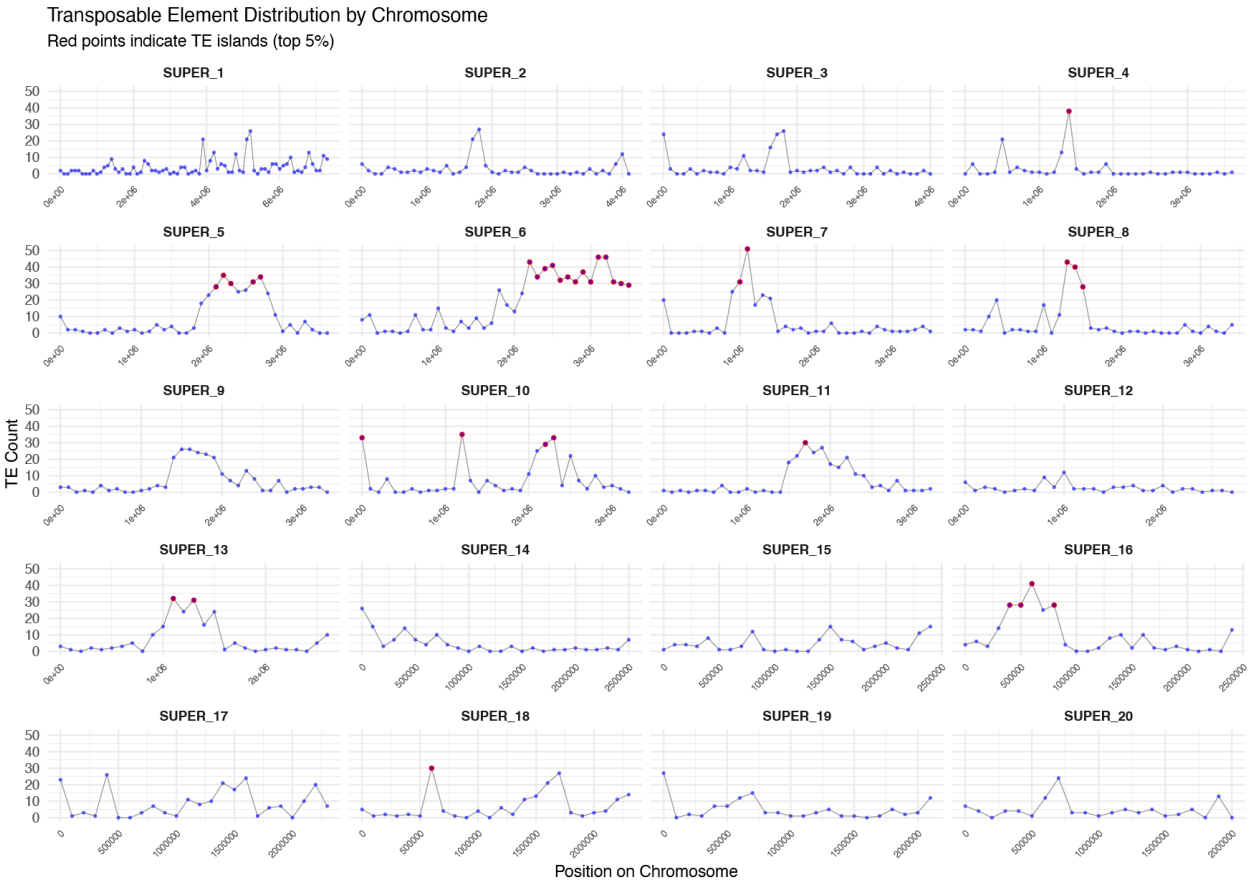

Figure S13. Number of counts of transposable elements (TEs) per Mb in the different chromosomes of *Oscarella lobularis*. Red dots indicate TE islands (top 5%).

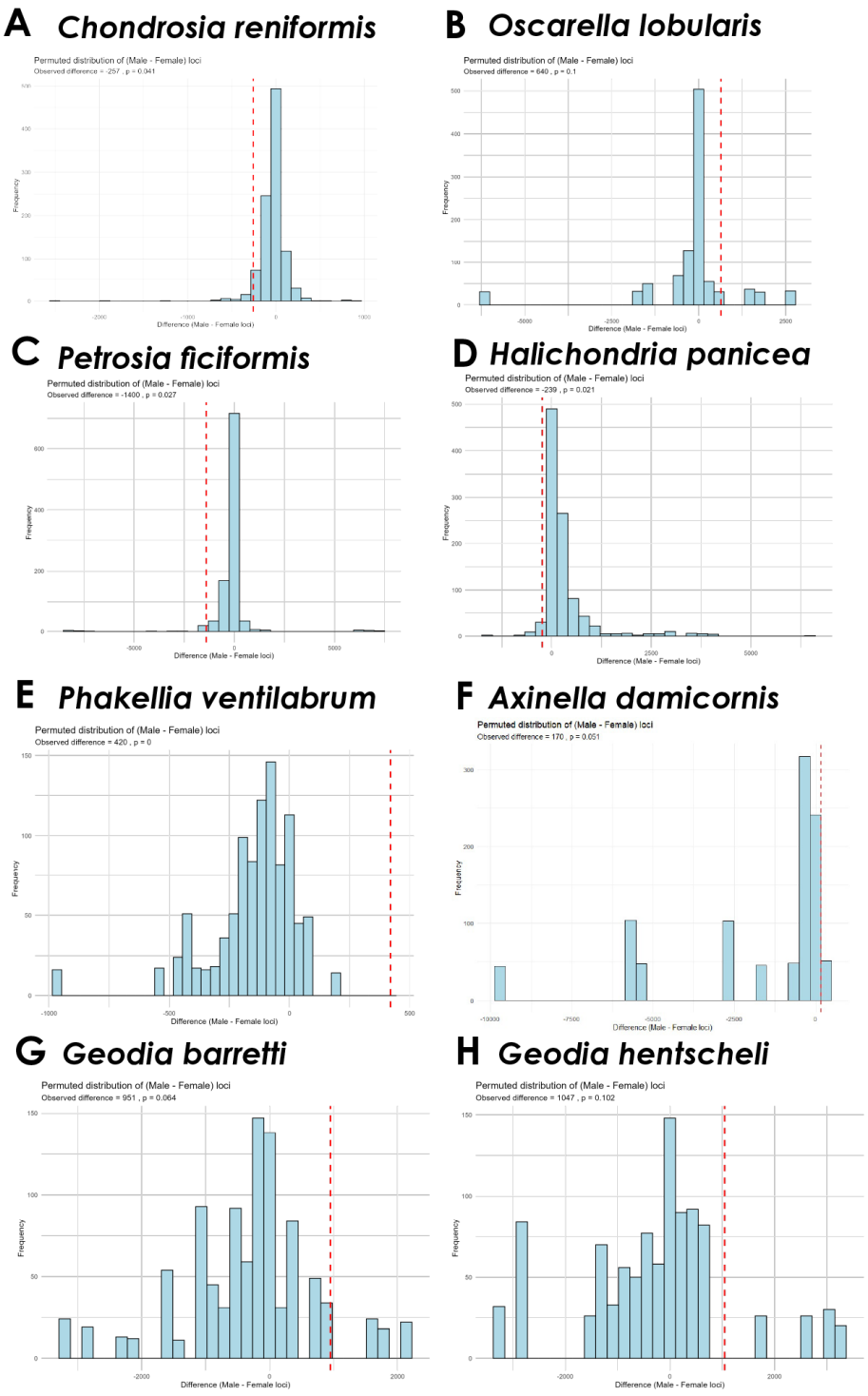

Figure S14. Histogram showing the bias towards female or male loci when randomising the sex assigned to the samples to run RADSEX using 100 permutations. p-values obtained performing a chi-square test comparing the observed result against the permutation-based expected frequencies. The red line points to the observed result, obtained when samples were assigned sex based on the presence of female or male gametes.

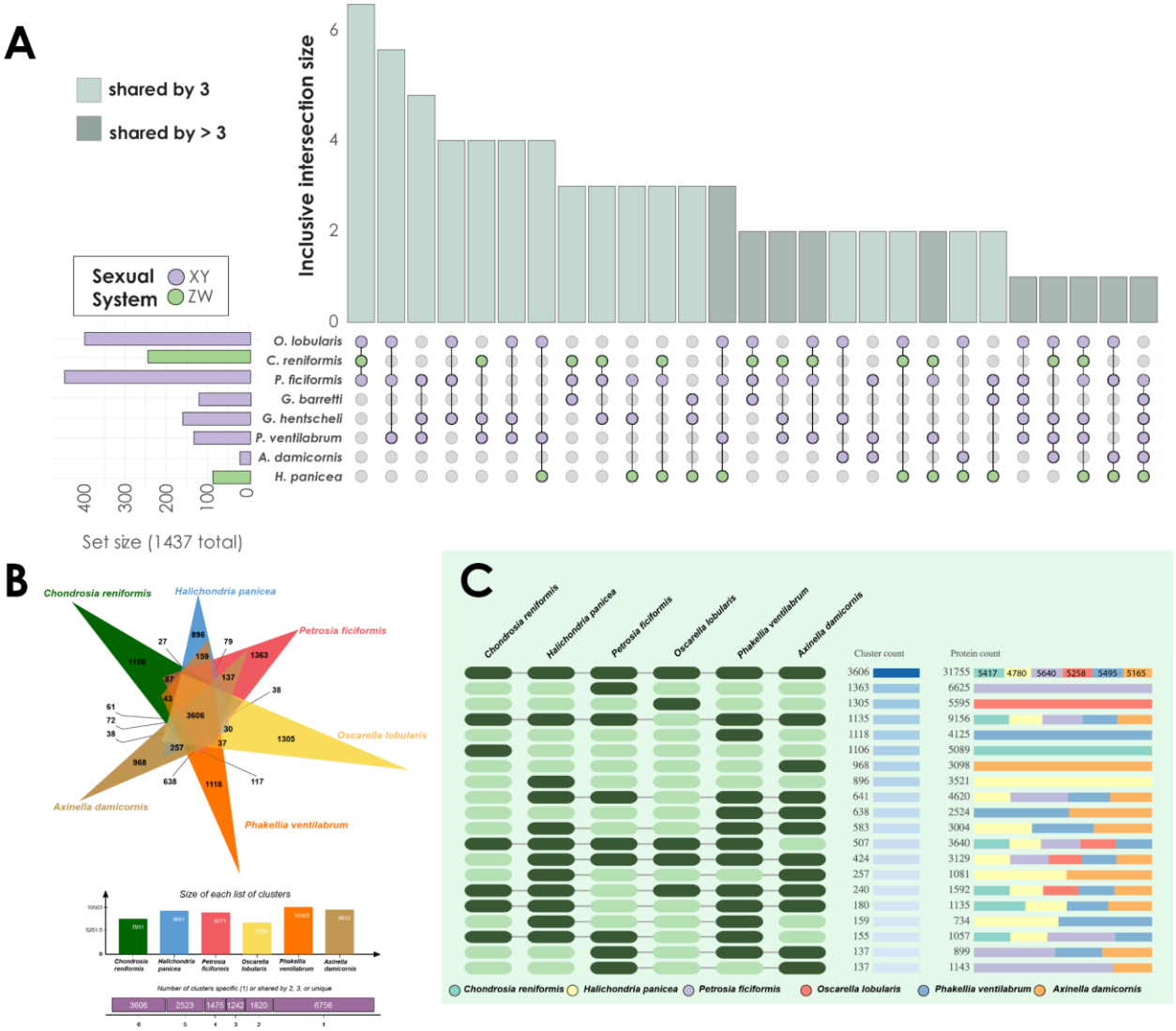

Figure S15. A. Upset plot showing the number of shared genes with sex-specific loci among the different species (only showing those shared by at least 3 species). **B.** Venn diagram from Orthovenn3 showing the number of orthogroups (clusters of proteins) shared among the 6 genomes, and **C.** Diagram of the number of protein clusters (and proteins within) shared among the different genomes.

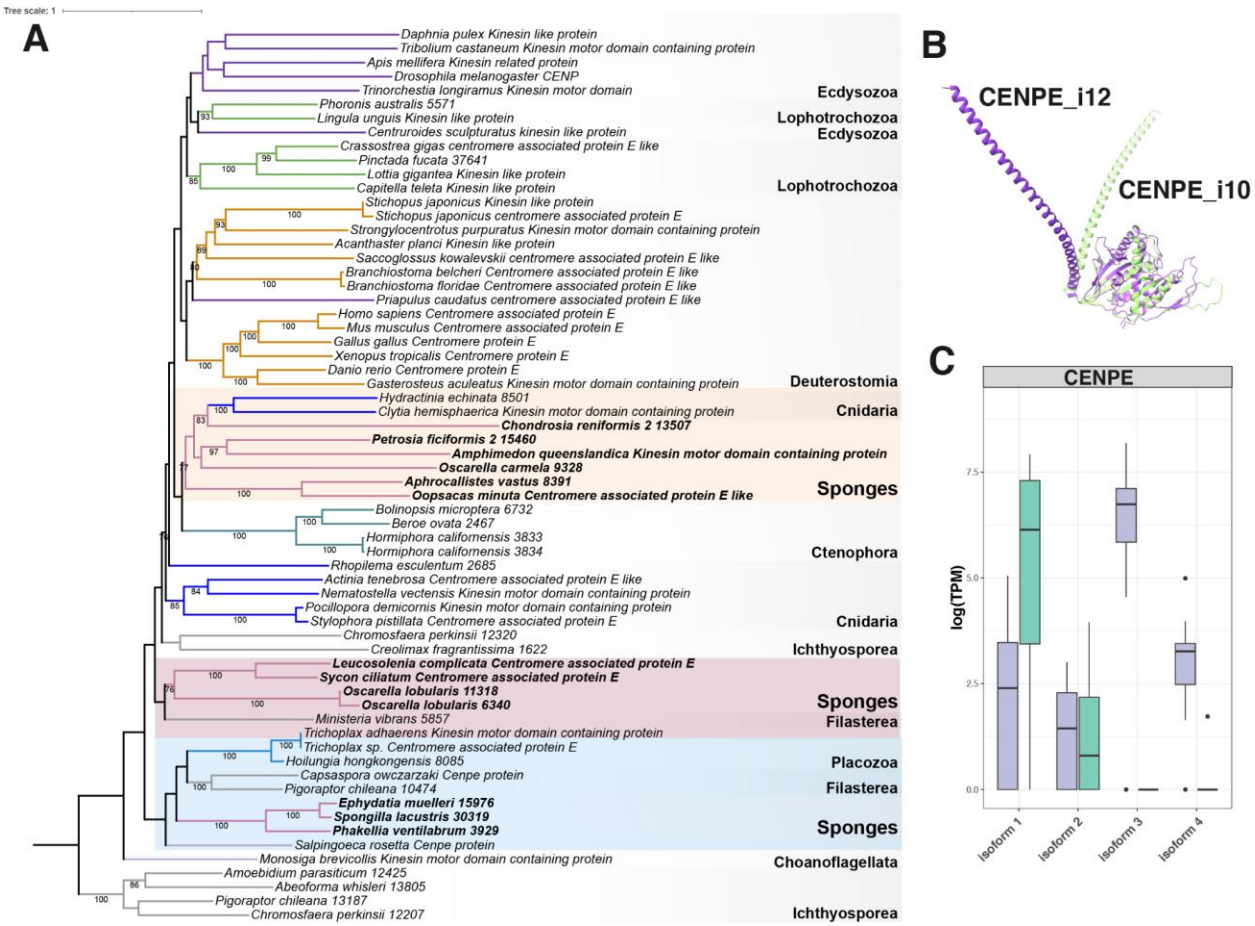

Figure S16. Phylogenetic hypothesis for the evolution of CENP-E proteins across metazoans obtained with IQ-TREE. Bootstrap values are shown below the nodes. **B.** Alphafold structural prediction for CENP-E isoforms that are more expressed in one sex (CENPE_i10, males) or the other (CENPE_i12, females) in *Chondrosia reniformis*. **C.** Expression levels in TPMs for the CENP-E isoforms in *Chondrosia reniformis.* In B and C, green denotes males and purple denotes females.

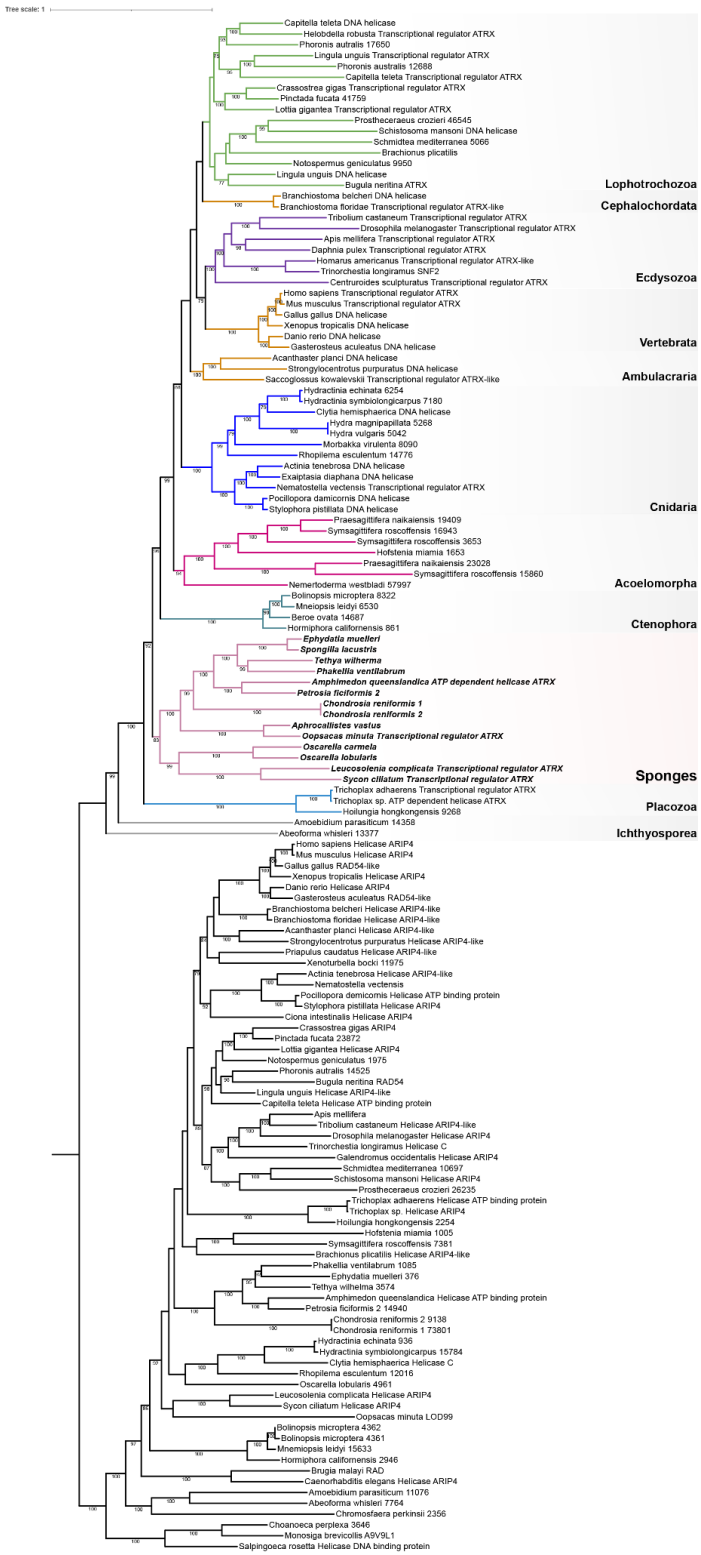

Figure S17. Phylogenetic hypothesis for the evolution of ATRX across metazoans obtained with IQ-TREE. Bootstrap values are shown below the nodes.

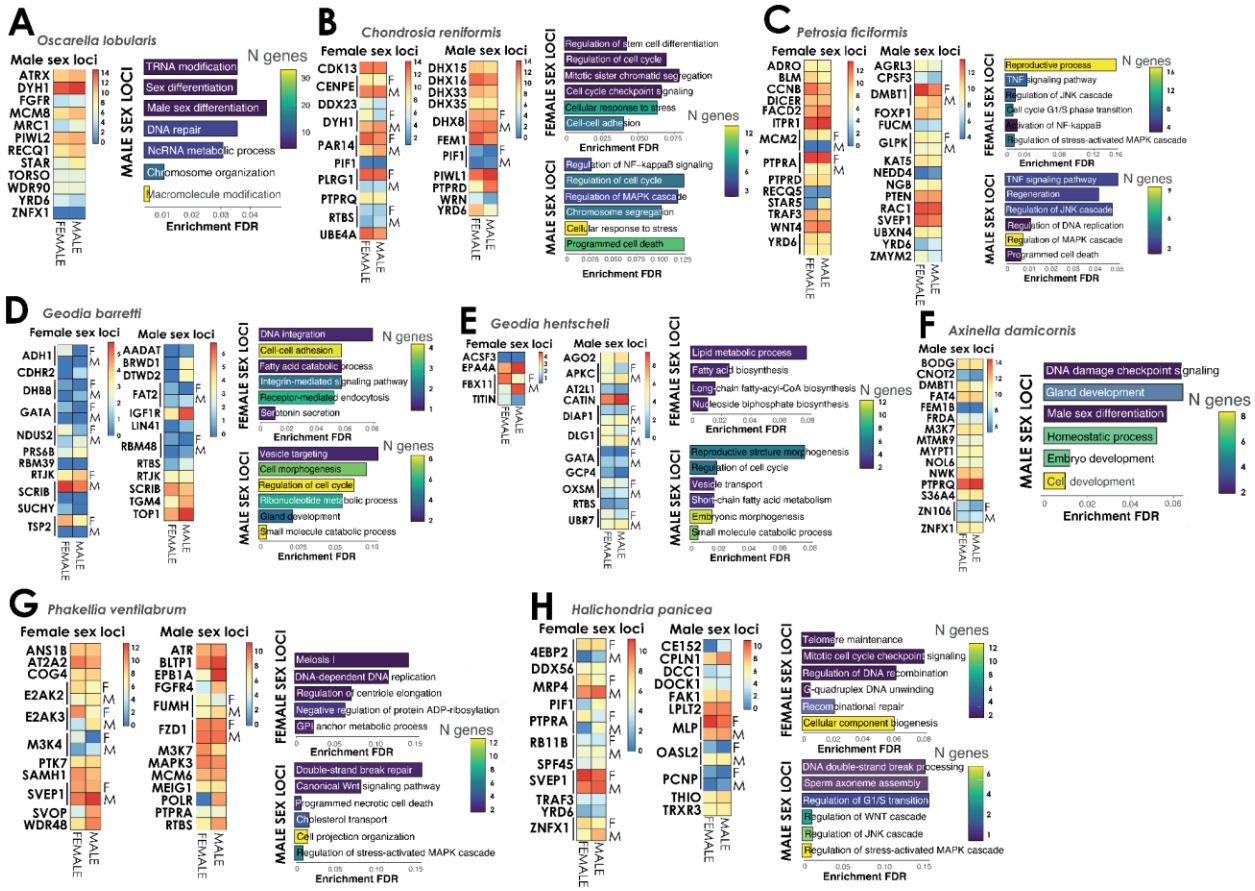

Figure S18. A–H. Relative expression of selected genes with male and female-specific loci in male and female specimens, along with the GO enrichment of both genes with male and female-specific loci in all species. Note that genes with isoforms that are more expressed in one sex than the other are indicated with F and M at the right side of the heatmaps.

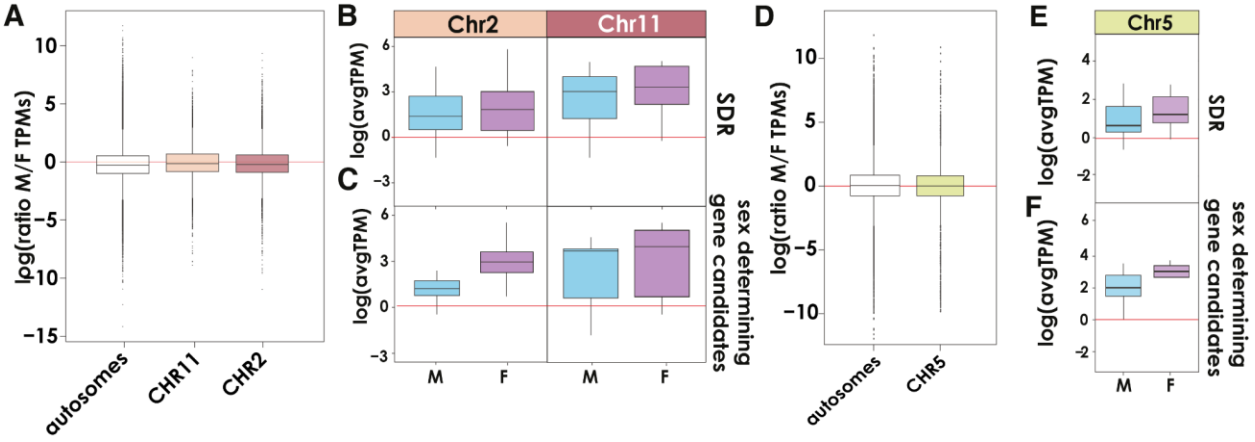

Figure S19. A. Average ratio of male/female expression along the autosomes and sex-linked chromosomes (2 and 11). **B–C.** Average gene expression (TPMs) in females and males in the SDRs and the genes with known sex determination function of the SDRs. **D.** Average ratio of male/female expression (TPMs) along the autosomes and the sex-linked chromosome 5. **E–F.**

Average gene expression (TPMs) in females and males in the SDR and the genes with known sex determination function of the SDR.

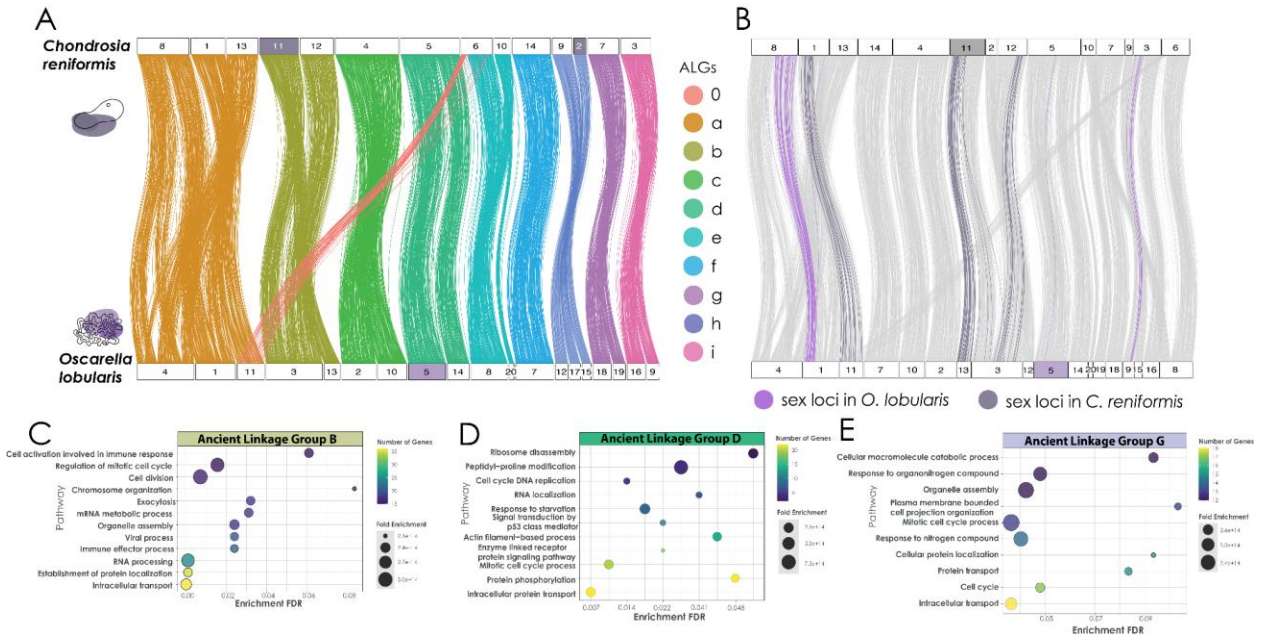

Figure S20. A. Conserved synteny across the species with sex chromosomes: *Chondrosia reniformis* and *Oscarella lobularis*. Colored vertical lines connect orthologous genes (n=4,582) across the two species with significantly enriched conservation of synteny. Each color represents a distinct ALG. Two or more colors converging on a chromosome indicate fusion events. **B.** Conserved synteny of genes with sex-specific loci (n=115) between *Chondrosia reniformis* and *Oscarella lobularis*. Colored vertical lines indicate genes with sex-specific loci, and grey vertical lines are all ALGs shown in Figure S14A. C–E. Gene Ontology enrichments of ALGs on sex chromosomes (ALGB in sex chromosome 11 of *C. reniformis*, ALGD in sex chromosome 5 of *O. lobularis* and ALGG in potential sex chromosome 2 of *C. reniformis*).

Table captions

Table S1. Sampling details for all species considered in the study and basic reproductive measurements.

Table S2. Genomic and transcriptomic resources of each species. References for the genomes: Pita et al. 2025[160]; Riesgo et al. 2025[161]; Steindler et al. 2025[162]).

Table S3. Haplotype-level measures of FST from the Stacks populations program (Catchen et al., 2013) for *Chondrosia reniformis* and *Oscarella lobularis*.

Table S4. **Normalised coverage .bed files** of *Chondrosia reniformis*, *Oscarella lobularis, Axinella damicornis, Petrosia ficiformis,* and *Phakellia ventilabrum* from mosdepth (Pedersen and Quinlan, 2018)

Table S5. Coverage with 0 mismatches allowed in 50kb windows of *Chondrosia reniformis*, *Oscarella lobularis, Axinella damicornis, Petrosia ficiformis,* and *Phakellia ventilabrum* calculated by findZX **(**Sigeman *et al*., 2022).

Table S6. Heterozygosity in 50kb windows of *Chondrosia reniformis*, *Oscarella lobularis, Axinella damicornis, Petrosia ficiformis,* and *Phakellia ventilabrum* calculated by findZX **(**Sigeman *et al*., 2022).

Table S7. SNP density in 100000kb windows of *Chondrosia reniformis*, *Oscarella lobularis, Axinella damicornis, Petrosia ficiformis,* and *Phakellia ventilabrum*. SNPs relative to the reference genome detected by bcftools [v1.12] (Li, 2011) and calls filtered with vcftools [v0.1.16] (Danecek et al., 2011). P-value of the differences between male and female calculated after a 1000 permutations test from SNPdensity_permutations.R from the SexFindR pipeline (*83*).

Table S8. A. **Summary of RADSex** markers identified as sex-specific loci (p>0.05) in each of the species that were present in all or most individuals of one sex and none of the other. Cells in green are male-specific loci, and cells in purple are female-specific loci. B. Results for the chi-square tests of the RADSex markers accumulation. Chi-square test (with Monte Carlo simulation because of low counts) was used to test whether marker frequencies differed from a uniform distribution across chromosomes. C–I. RADsex loci information and annotation on the genome and transcriptomes. C. *Chondrosia reniformis*. D. *Oscarella lobularis*. E. *Petrosia ficiformis*. F. *Phakellia ventilabrum*. G. *Halichondria panicea*. H. *Axinella damicornis*. I. *Geodia barretti*. J. *Geodia hentscheli*.

Table S9. Genes with sex-specific loci. A. Mapping correspondence of sex-specific loci to the reference genomes of *Chondrosia reniformis*, *Oscarella lobularis*, *Petrosia ficiformis*, *Phakellia ventilabrum*, *Axinella damicornis* and *Halichondria panicea*. B. Abbreviations for the genes with sex-specific loci in all species surveyed, which were annotated through blast against SwissProt proteins. C. Genes with sex-specific loci annotated with GO category sex determination in the different species. D. Genes with sex-specific loci shared among the different species and unique to each of them.

Table S10. Functional annotations for the genes with sex-specific loci in all species surveyed, obtained with RADSEX and then annotated through blast against SwissProt proteins. A. *Chondrosia reniformis*. B. *Oscarella lobularis*. C. *Petrosia ficiformis*. D. *Geodia barretti*. E. *Geodia hentscheli*. F. *Phakellia ventilabrum*. G. *Axinella damicornis*. H. *Halichondria panicea*.

Table S11. Functional annotations for the genes with sex-specific loci located in the SDRs of sex chromosomes of *Chondrosia reniformis* (A) and *Oscarella lobularis* (B), obtained by blast hits against swissprot proteins. Their function related to sexual reproduction or sex determination is indicated with a reference to the scientific article where it is described.

Table S12. Gene expression levels (TPMs) for the genes with sex-specific loci located in the SDRs of sex chromosomes of *Chondrosia reniformis* (A, B) and *Oscarella lobularis* (C), obtained with StringTie.

Table S13. Gene ontology enrichments for the genes with sex-specific loci in all species surveyed, obtained with ShinyGO. A. *Chondrosia reniformis*. B. *Oscarella lobularis*. C. *Petrosia ficiformis*. D. *Geodia barretti*. E. *Geodia hentscheli*. F. *Phakellia ventilabrum*. G. *Axinella damicornis*. H. *Halichondria panicea*.

Table S14. Differential gene expression (at transcript level) for all species, performed with Hisat2/StringTie/edgeR in those with a reference genome and trinity/RSEM/edgeR in those without. A. Summary statistics for the analysis in all species. Abbreviations: N transcripts=number of transcripts, N DEG=number of differentially expressed genes/transcripts, N DEG

sexloci=number of differentially expressed genes/transcripts with a sex-specific loci, Up in M=upregulated in males, Up in F=upregulated in females, %DEG=proportion of transcripts differentially expressed, N genes w/sex loci=number of genes with a sex-specific loci, %DEG sexloci=proportion of transcripts with a sex-specific loci that are differentially expressed. B. *Chondrosia reniformis*. C. *Oscarella lobularis*. D. *Petrosia ficiformis*. E. *Axinella damicornis*. F. *Phakellia ventilabrum*. G. *Halichondria panicea*. H. *Geodia barretti*. I. *Geodia hentscheli*.

Table S15. Transcript-level expression for genes with sex-specific loci in all species surveyed. A. *Oscarella lobularis*. B. *Chondrosia reniformis*. C. *Petrosia ficiformis*. D. *Axinella damicornis*. E. *Phakellia ventilabrum*. F. *Halichondria panicea*. G. *Geodia barretti*. H. *Geodia hentscheli*.

Table S16. Alternative splicing events (AS) and association with sex. A. Summary of overall AS in the different species and AS in genes with sex-specific loci. Abbreviations: skipped exons (SE), mutually exclusive exons (MX), alternative 5′ splice sites (A5), alternative 3′ splice sites (A3) and retained intron (RI). B. *Chondrosia reniformis*. C. *Oscarella lobularis*. D. *Petrosia ficiformis*. E. *Phakellia ventilabrum*. F. *Axinella damicornis*. G. *Halichondria panicea*.

Table S17. Table of orthologs in all six species: Sp1=*Oscarella lobularis*, Sp2=*Chondrosia reniformis*, Sp3=*Petrosia ficiformis*, Sp4=*Phakellia ventilabrum*, Sp5=*Axinella damicornis*, Sp6=*Halichondria panicea*. Labels: sex_gene_sp$=indicates genes with sex-specific loci in the given species, Sex_all=indicates genes with sex-specific loci in a certain species.

Table S18. Phylostratigraphic distribution of protein-coding and sex-related genes in six sponge species. The table summarizes results from GenEra analyses across the six species (*Chondrosia reniformis*, *Halichondria panicea*, *Petrosia ficiformis*, *Phakellia ventilabrum*, *Amphimedon queenslandica*, *Oscarella lobularis*), including (A–B) the assignment of 102,213 protein-coding genes to evolutionary strata, (C) the subset of 2,705 sex-related genes and their orthogroup distribution, (D) functional GO enrichment categories inferred for sex loci, and (E) enrichment profiles of sex-related genes across species.