Sexual antagonism and sex determination in three syngnathid species alongside male pregnancy gradient and varying sex roles.

Abstract

Teleost fish show an enormous diversity of sex determination systems, varying from environmental sex determination to full sex chromosomes. Traditionally sex chromosomes are considered within a framework of sexually antagonistic relationships and are viewed not only as sex determination systems but also as a means to resolve sexual conflict by confining conflicting genes to one sex. However, the relationship between emergence sexually-antagonistic loci, resolution of sexual conflict and sex determination is not fully understood. In this study we take a look at genomic and transcriptomic signatures of sexual antagonism, and relate them to putative sex determination systems in three species from the family Syngnathidae: *Nerophis ophidion*, *Syngnathus typhle* and *Hippocampus erectus*. The family is famous for their extreme form of parental care - male pregnancy. Male pregnancy forms a parental investment gradient among species of the family. We selected species alongside this gradient to investigate the effects male pregnancy has on sexual conflict in syngnathids. We found that both S.typhle and N.ophidion seem to have environmental sex determination and that sexual conflict is primarily resolved via differential gene and allele expression. On the other hand *H.erectus* possess an XY sex chromosome system. Strikingly, the identified homomorphic Y chromosome is not homologous to the Y chromosome of *H.erectus* identified in an independent study. We hypothesize that both sex-linked chromosomes evolved due to captive breeding and represent transition states between sex determination systems. Furthermore, we propose that the master sex determination gene is located outside of these sex-linked chromosomes.

Introduction

A phenotype that increases reproductive success of an individual, i.e., its fitness, will be favored by selection. This includes phenotypes directly influencing survival, e.g., the famous dark melanic and pale morphs of the peppered moth (Walton and Stevens 2018), and phenotypes maintained by mate choice and sexual conflict, e.g., increased phenotypic diversity in female damselflies due to male harassment (Takahashi et al. 2014). The imposed selection on a phenotype stems from a dynamic interaction between the organism and its environment e.g. niche construction. As such, the environment influences the development of an organism by affecting its gene expression, embryogenesis and survival, while the organism itself might choose and change its environment, allowing to adjust its survival and reproduction probability. This makes the organism both subject and object of its own evolution (Levins and Lewontin 1985; Svensson 2018; Erik I Svensson 2023). Such bi-directional cause and effect relationships exist not only between the organism and its abiotic environment but also between the sexes of

the same species - where one sex selects for certain phenotypes of the opposite sex and vice versa (Svensson 2018; Erik I Svensson 2023; Erik I. Svensson 2023).

Sexual reproduction selected for the emergence of male and female gametes in most multicellular organisms, and, as reflected in most metazoans, the separation of male and female sexes (Bachtrog et al. 2014; Furman et al. 2020). The mechanisms of sex determination (SD), however, vary greatly among clades and even among closely related species (Bachtrog et al. 2014; Furman et al. 2020). In mammals and birds, the SD system is genetic and relies on generally conserved sex chromosomes (XY and ZW, respectively), while reptiles and teleost fish show an enormous diversity of SD strategies - ranging from environmental SD (ESD), monogenic SD, with one master sex determination gene, including SD systems with homomorphic and heteromorphic sex chromosomes to polygenic SD (Judith E Mank et al. 2006; Judith E. Mank et al. 2006; Bachtrog et al. 2014; Furman et al. 2020). The variation in SD loci among closely related species suggests that a transition or a “turnover” event from one SD locus or chromosome to another has occurred between them. In some teleost clades, like cichlids, such turnover events occur frequently (Lichilín et al. 2021(Lichilín et al. 2021). Though evolved independently, novel SD loci often seem to be co-opted from a limited number of genes that are involved in a classical sex differentiation network (Bertho et al. 2018) Seemingly, an exception - the de novo evolved salmonid master SD gene (*sdY*) which represents a duplication of IFN regulatory factor 9 (*irf9*), has been shown to interact with a female-determining transcription factor f*oxl2*, disrupting the female differentiation pathway allowing testis development (Bertho et al. 2018). This confirms the notion that the selection for SD loci is limited to genes within the classical sex differentiation cascade or the genes that interact with it.

Selection acts differently on males and females (Parsch and Ellegren 2013; Bachtrog et al. 2014; Mank 2017; Rowe et al. 2018; Furman et al. 2020; Erik I Svensson 2023). The amount of energy invested in gamete production, i.e. "costly eggs” vs. “cheap” and abundant sperm, parental care, mate choice, sex roles and sexual dimorphism all contribute to this differential in natural and sexual selection between the sexes. In some circumstances this can affect the distribution range of a species (Dudaniec et al. 2022) and ultimately increase or decrease species extinction risks by influencing effective population size, genetic drift, inbreeding, and rates of maladaptive allele purging (Rankin et al. 2011; Lumley et al. 2015; Martínez-Ruiz and Knell 2017; Erik I. Svensson 2023).

Conflicting male and female optimal fitness strategies can turn into a “tug of war” - a bi-directional cause and effect dynamics maintained by sexual antagonism (SA) (Rice and Holland 1997; Svensson 2018; Iversen et al. 2019; Svensson 2019). SA often emerges in relation to SD and the appearance of genetic SD regions (Bachtrog et al. 2014; Smith et al. 2023). The connection between these phenomena is not fully understood, however, different models of co-evolution of SD and SA loci have been suggested (Bachtrog et al. 2014; Cheng and Kirkpatrick 2016; Furman et al. 2020; Smith et al. 2023). Though sex chromosomes can be enriched for SA genes, antagonism is often resolved via sex-specific gene expression (Parsch and Ellegren 2013; Bachtrog et al. 2014; Grath and Parsch 2016) and the causal relationship between SD and SA isn’t always understood (i.e. location of SA loci or emergence of SA loci in genomic regions before or after sex-linkage (Smith et al. 2023). The resolution of sexually antagonistic relationships is complicated by the fact that males and females possess almost the same genome. When certain alleles are advantageous to one sex and neutral or disadvantageous to the other, an intralocus sexual conflict occurs which is maintained by

sexually antagonistic selection (SAS) (Cheng and Kirkpatrick 2016). Such conflict supposedly leaves a distinct signature within a genome and can be detected by investigating intersex allele frequency variations or changes in genetic diversity at particular loci (Cheng and Kirkpatrick 2016; Wright et al. 2018; Lichilín et al. 2021) and probably sex-biased gene expression (Mank 2017).

Distinguishing different forms of selection according to their genomic signatures can be challenging. In order to unravel complex relationships between the sexes and detect traces of SS and SAS within the genome, it is essential to have a system with clear roles of particular phenotypes in SA relationships.

The family Syngnathidae is famous for their unusual morphology and extreme form of parental care - their unique evolution of male pregnancy. Within the family, sex roles vary from traditional with female choice, to reversed where females compete for male mates. Syngnathids also exhibit a diverse set of mating strategies and sexual dimorphism in size and ornamentation (Roth et al. 2020). Their phenotypic diversity expands also towards a diversity in SD systems: Currently, we know about monogenic SD in *Phyllopteryx taeniolatus* (Qu et al. 2021) (*amhr2y* on Chr4), potentially ESD in *Syngnathus scovelii,* and sex chromosomes in *Hippocampus abdominalis* and *Hippocampus erectus* with no identified master SD gene (Long et al. 2023). Some members of this family also possess a modified adaptive immune system - a loss of the MHC II pathway in *Syngnathus* and modifications in *Hippocampus* (Roth et al. 2020). Coupled with differential immune needs due to male pregnancy, this adaptation offers interesting possibilities to study the effects of SS and SAS over viability and reproduction.

For this study, we selected three syngnathid species with different sex roles, mating strategies, and sexual dimorphism. *Hippocampus erectus* here represents a species with a traditional sex role, monogamy, and larger males. *Syngnathus typhle* and *Nerophis ophidion* represent species with inverted sex roles, polygamy, and large decorated females. *N. ophidion* has a much stronger sexual dimorphism than *S. typhle* where males and females are often difficult to differentiate when they are not in their mating season. This dimorphism is much more pronounced in N. ophidion whereas in S. typhle it is limited to be displayed in the mating season only. We relied on a previously published chromosome-level *H. erectus* genome assembly and two pipefish long-read assemblies that were newly generated for this study (PRJNA947442). To detect signatures of SS and SAS and identify sex-linked regions within the genome, we used RNA-seq data obtained from head kidney and gill of males and females of each species. The chosen tissues are not involved in reproduction nor in the establishment of secondary sexual signals. Gills represent a first physical barrier to pathogens and are in direct interaction with the environment. Head kidneys are major lymphoid organs in teleosts that are supposed to have a similar function as the thymus in other vertebrate groups (Press and Evensen 1999). The choice of the organs has followed a conservative approach to assure that they are not subject to hormonal influence or sex-limited/heavily specialized, as e.g. reproductive organs as gonads, testes or brood pouch and any parts of the body that are directly related to sexual dimorphism. Hence, sex bias in expression is more likely to be directly related to the effects of SS and SAS over viability and survival. Indirectly, due to male pregnancy, reproduction would also be affected as eggs require oxygen supply and interact with the father's mucosal immune system (Whittington et al. 2015; Nygård et al. 2019; Parker et al. 2023). Besides genome wide Fst and genes with sex-biased expression (SBGs), we looked at differential allele expression. In other species SBGs were previously shown to also display high intersex Fst values, and the

emergence of sex-specific genetic architecture was suggested as a way to resolve sexual conflict (Cheng and Kirkpatrick 2016; Wright et al. 2018). This complicates the separation of ongoing intralocus conflict signatures and the effects of sex-specific genetic architecture. As a resolution, we here propose the investigation of intersex bias in allele expression to increase the sensitivity in detecting intralocus conflict signatures.

Since many sex chromosomes show divergence in allele frequencies and enrichment for SBGs, we used these metrics to identify the putative sex-linked regions. In addition, we employed kmerGWAS analysis to identify sex-linked loci even if only represented by a single-base difference (Cheng and Kirkpatrick 2016).

This study sheds light on sex-linked genomic regions in three syngnathid species, the underlying processes involved in sexual conflict and the relationships between the loci under SA and sex-linked regions.

Methods

Sample collection, nucleic acid isolation and sequencing.

DNA

An adult *Syngnathus typhle* male and an adult *Nerophis ophidion* female, were taken from our animal facility at GEOMAR, Germany. Large animals were chosen to achieve sufficient amounts of high molecular weight (HMW) DNA for Pacific Biosciences library preparation. *N. ophidion* displays a strong sexual size dimorphism with large females and small males. We thus sampled a female *N. ophidion* for genome sequencing.
DNA extraction, library preparation, and PacBIO sequencing was performed by the Norwegian Sequencing Centre, Oslo, Norway (sequencing.uio.no).
The DNA samples were processed according to the Pacific Biosciences Express library preparation protocol without prior fragmentation. The final libraries were size-selected with a BluePippin with 11 kb cut-off and sequenced on a Sequel II instrument using a Sequel II Binding kit v2.0 with Sequencing chemistry v2.0. Circular Consensus Sequences (CCS) were generated using the CCS pipeline (SMRT Link v10.1.0.119588) with default settings.

RNA

Hippocampus erectus individuals were bred at the GEOMAR aquaria facilities, Kiel, for several generations prior to the experiment. *Syngnathus typhle* and *Nerophis ophidion* were wild-caught in the Baltic Sea (54°23′N; 10°11′E, Germany) and subsequently kept in aquaria facilities at GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany. Ten males of each species were used. Five females were used in the case of *N. ophidion* and *S. typhle*, and seven in the case of *H. erectus*. All fish were euthanized with an overdose of MS-222 (Tricaine, 500 mg L−1; Sigma-Aldrich). After dissection, head kidney and gill samples were collected from each individual and placed in an RNAlater solution (Invitrogen) and incubated at 4°C for one week. Samples were then stored at −20°C until extraction. Total RNA was extracted using an RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. Libraries were prepared using Illumina TruSeq stranded mRNA kit (2×150bp) and sequenced on an Illumina NovaSeq 6000 instrument at the Competence Centre for Genomic Analysis (CCGA), Kiel.

Genome and transcriptome assembly and annotation.

Genome assembly.

The *S. typhle* and *N. ophidion* HiFi reads (CCS reads with >99% accuracy) were assembled using the Genome Assembly pipeline (SMRT Link 10.1.0.119588) with default settings
Gene space completeness of the assemblies was assessed with BUSCO (v5.2.2) (Manni et al. 2021) using the Actinopterygii reference protein set. Continuity of the assemblies was assessed with Quast (v5.0.2) (Mikheenko et al. 2018). The chromosome-level assembly of *H.erectus* and supporting annotation used in this study was previously published by Li et.al (PRJNA613176) (Li et al. 2021).

Transcriptome assembly and gene prediction.

The RNA reads generated in this study were quality trimmed with Fastp (v0.20.1) (Chen et al. 2018) with default settings. Reads from both organs were pooled and assembled with Trinity (v2.8.1) (Grabherr et al. 2011), to generate a single transcriptome assembly per species. In order to reduce redundancy of transcripts per gene, the resulting assemblies were subjected to several filtering steps. First, assemblies were filtered with Transrate (v1.0.3) (Smith-Unna et al. 2016) using raw reads as evidence. The software assigns a quality score for the entire assembly and per contig on the basis of various metrics, like the base quality scores, number of reads mapped per contig etc., determines an optimal threshold and filters the assembly. Transrate-filtered assemblies were subjected to CD-Hit (v4.8.1) (Li and Godzik 2006) clustering transcripts with 98% similarity threshold into one. These contigs were then used to identify putative protein sequences encoded within the assemblies with TransDecoder (v5.5.0). Obtained protein sequences were clustered with CD-Hit (v4.8.1) (Li and Godzik 2006) with 95% similarity threshold. Sequence identifiers of these filtered proteins were used to filter the Transdecoder CDS fasta file to reduce redundancy by only selecting the CDS sequences whose protein product passed a similarity filter. We used these filtered cDNA sequences as our transcriptome assemblies in all subsequent analyses.

Repeat masking and genome annotation.

Prior to genome annotation, repetitive elements were identified and masked. EDTA (v2.0) (Xu and Wang 2007; Gremme et al. 2013; Xiong et al. 2014; Ou and Jiang 2018; Ou et al. 2019; Shi and Liang 2019; Su et al. 2019; Zhang et al. 2022) was used to produce a *de novo* repeat library for *S.typhle* and *N.ophidion* assemblies. The resulting repeat libraries were used to mask the corresponding assemblies with RepeatMasker (v4.1.2-p1) using the settings “-a -gff -pa 32 -u -xsmall -nolow”.

In order to use RNA-sequencing data as evidence for gene prediction, we had to align them to the genome. For both *S. typhle* and *N. ophidion,* RNA reads were thus aligned to their respective species' genome assemblies with STAR (v2.7.9a) (Dobin et al. 2013) in two-pass mode. Only uniquely mapped reads were considered. Filtered transcriptome assemblies were aligned to the genomes with GMAP (version 2017-11-15) (Wu and Watanabe 2005) with the following options "-B 5 -t 32 --input-buffer-size=1000000 --output-buffer-size=1000000 -f samse".
The *S. typhle* and *N. ophidion* genome assemblies were annotated using BRAKER (v2.1.6) (Lomsadze 2005; Stanke et al. 2006; Gotoh 2008; Stanke et al. 2008; Li et al. 2009; Barnett et al. 2011; Iwata and Gotoh 2012; Lomsadze et al. 2014; Buchfink et al. 2015; Hoff et al. 2016; Hoff et al. 2019; Brůna et al. 2020; Brůna et al. 2021). We ran BRAKER twice. First with RNA, both the unassembled reads and the transcriptome assemblies, then with OrthoDB proteins (odb10_vertebrata_fasta) (Kriventseva et al. 2019) as evidence. Afterwards, we combined the annotations with TSEBRA (v1.0.3) (Gabriel et al. 2021), giving the RNA evidence more weight (pref_braker1.cfg). The cDNA and amino acid sequences were obtained from the final BRAKER GTF with AGAT (v0.8.0) (Dainat et al. 2023) toolkit using agat_sp_extract_sequences.pl script with "--aa --clean_final_stop –clean_internal_stop" options to get amino acid sequences, and "-t exon –merge" to get cDNA. The putative function of predicted protein sequences was determined with Orthofinder (v2.4.0) (Emms and Kelly 2019) using sequences from the following fish species:
Danio rerio (GRCz11), Gasterosteus aculeatus (BROADS1), Hippocampus comes (H_comes_QL1_v1, GCA_001891065.1), Hippocampus erectus (PRJNA613176), Lepisosteus oculatus (LepOcu1), Oryzias latipes (ASM223467v1), Syngnathus acus (GCF_901709675.1_fSynAcu1.2), Takifugu rubripes (rubripes.fTakRub1.2), Syngnathus rostellatus (GCF_901709675.1).

Transcriptome functional annotation.

The transcriptomes were functionally annotated in a similar manner to the genomes. We used predicted proteins from *H. erectus*, *S. typhle* and *N. ophidion* transcriptome assemblies together with the previously mentioned 8 fish species to run orthology analysis with Orthofinder (v2.4.0) (Emms and Kelly 2019).

Variant calling.

Genome-independent approach.

Following a similar approach as described in El Taher et al. (El Taher et al. 2021), KisSplice (v2.4.0) (Sacomoto et al. 2012; Lopez-Maestre et al. 2016; Benoit-Pilven et al. 2018) was used to call SNPs directly from raw RNA reads without prior genome mapping. The SNP data was then processed with the kissDE R package (v1.14.0) (Sacomoto et al. 2012; Lopez-Maestre et al. 2016; Benoit-Pilven et al. 2018) to identify SNPs that showed differential expression between

sexes. To predict their functional impact and identify the genes that exhibited differential allele expression, the SNPs were placed on the corresponding transcriptome assemblies with BLAT (v36) (Kent 2002) using "—minIdentity=80" option and KisSplice2refTranscriptome (v1.3.3) (Sacomoto et al. 2012; Lopez-Maestre et al. 2016; Benoit-Pilven et al. 2018). At this step, we utilized KisSplice2refTranscriptome flags to filter the data, which excludes SNPs that occurred in multiple assembled genes (SNP_in_mutliple_assembled_genes flag) and could simply represent a sequencing error (Possible_sequencing_error flag). Then, we filtered the dataset to include only bi-allelic SNPs.
As we wanted to investigate both, sex-biased allele expression and identify sex-linked regions within the genome, the aforementioned SNPs were divided into two categories: i) differentially-expressed alleles (DEA) - alleles that show a bias in expression, i.e. differences in read counts for pairs of variants, between the sexes. ii) Sex-specific alleles (SSA) (male or female) - were defined as SNPs where all individuals of one sex are homozygous at this locus, whereas at least 75% of individuals of the opposite sex are heterozygous at this locus. We consider these as fixed differences but point out that due to the usage of RNA sequencing data those could also represent extremes of differential allele expression. However, the KisSplice developers previously showed that expressed allele frequencies very closely reflect true allele frequency differences as would be derived from DNA sequencing (Sacomoto et al. 2012; Lopez-Maestre et al. 2016). In the following, with SNPs of category ii) we consider two potential sex determination systems - an XY system with a male-specific Y and ZW system with a female-specific W chromosome.

Genome-based approach.

The RNA reads were aligned to the annotated genomes with STAR (v2.7.9a) (Dobin et al. 2013) in two-pass mode with the options "--outFilterIntronMotifs RemoveNoncanonical --outSAMunmapped None --outFilterMultimapNmax 1". The alignment data was then processed with the Genome Analysis Toolkit (GATK) (v4.2.2.0-17) (Poplin et al. 2017) and Picard (v2.25.4) (Broad Institute 2019). First, duplicate reads were marked (MarkDuplicates) and reads with Ns in the cigar strings were split into multiple alignments (SplitNCigarReads). Then, potential variants were called on each sample in gVCF mode (–ERC GVCF option). gVCF files for each sample were combined into a single gVCF per species and the variants were called with GATK's joint genotyping method.
As GATK guidelines do not recommend to use standard hard filtering criteria when working with SNPs called from RNA-seq data, the final VCFs were filtered with "QD < 2.0" and "FS > 30.0" thresholds applied, and using "--select-type-to-include SNP --restrict-alleles-to BIALLELIC" options to keep only high quality bi-allelic SNPs.

Fst calculation.

Using SNPs resulting from the genome-based approach, we calculated intersex Weir and Cockerham Fst values for each of the genomes in windows of three sizes - 1Kb, 10Kb, 100Kb with VCFtools (v0.1.14) (Danecek et al. 2011). To minimize bias introduced by uneven numbers

of males and females, we used data from only 5 individuals and 2 organs of each sex and set a missing data threshold to 20 percent

Overlap between approaches.

In order to minimize software-related bias, we investigated the overlap between SNPs called by KisSplice and GATK. To place KisSplice-called variants on a genome, we aligned KisSplice SNPs to the genome using STAR (v2.7.9a) (Dobin et al. 2013) and processed these alignments with KisSplice2refgenome (v2.0.5) (Sacomoto et al. 2012; Lopez-Maestre et al. 2016; Benoit-Pilven et al. 2018). Alleles and their locations were compared using a custom python script.

Differential gene expression analysis.

Transcript abundance was calculated with Kallisto (v0.46.1) (Bray et al. 2016) and imported into DESeq2 (v1.22.2) (Love et al. 2014) with tximport (v1.22.0) (Soneson et al. 2015). Genes with total counts less than 10 across all samples were removed from the dataset. Expression between sexes was analyzed using DESeq2 per species and organ with sex representing the model contrast. The difference in expression was considered significant if an adjusted p-value was lower than 0.01. The resulting gene lists were used in conjunction with Fst scans to identify genes that are differentially expressed and located within a region of elevated Fst.

Intersex SNP density analysis - kmerGWAS analysis.

For both SNP density calculations and kmerGWAS analysis, we followed the SexFindR workflow suggested by Grayson and colleagues (Grayson et al. 2022).
The SNP density per sample was calculated with vcftools in 10kb windows on bi-allelic SNPs called by GATK from previous steps. The results were analyzed using SNPdensity_permutations_fugu.R script, provided within the SexFindR workflow.
The script calculates the mean number of SNPs in males and females per window and then calculates the difference between the male and female means.
The KmerGWAS analysis was performed according to the SexFindR recommendations with no modifications, using kmersGWAS library (Voichek and Weigel 2020), PLINK (v1.07) (Purcell et al. 2007), ABySS (v2.0.2.) (Jackman et al. 2017), and BLAST (v2.5.0) (Altschul et al. 1990; Camacho et al. 2009; Boratyn et al. 2012). Results were visualized with R.

Results.

Genome assemblies and annotation.

To identify sex-linked genomic regions and provide additional support for genome-independent SNP calling, we generated draft genome assemblies and their annotations for two of the studied species. While the here used *H. erectus* dataset was already published (Li et al. 2021), the *N. ophidion* and *S. typhle* genomes were newly generated for this study. From here, the species will be referred to as HE (*Hippocampus erectus*), NO (*Nerophis ophidion*), and ST (*Syngnathus typhle*).

The NO genome was sequenced to ~22x coverage with PacBio HiFi reads of 14.9 Kb mean read length. The resulting assembly length was 1.7 Gb with an N50 of 1.2 Mb. For the ST genome, the coverage with PacBio HiFi was ~40x with a mean read length of 14.7 Kb. The total length of the assembly was 405 Mb with an N50 of 2.3 Mb. While we did not reach chromosomal resolution with these assemblies, the N50 of 1.2 and 2.3 Mb allows us to detect large sex-linked regions, blocks of genes with sex-biassed allele expression patterns and their synteny (i.e. corresponding linked locations).

The BUSCO gene space completeness scores revealed high levels of completeness with the *Actinopterygii* dataset for HE, NO, and ST assemblies were:

HE: Complete:96.6%/Missing:2.3%, NO: Complete:93.9%/Missing:4.0%, and ST: Complete:94.4%/Missing:4.1% respectively. The BRAKER2 workflow resulted in 27’499 genes annotated in ST (BUSCO C:93.7%/M:4.8%), and 38’598 genes in NO (BUSCO C:93.2%/M:4.3%). For comparison, the current HE annotation contains 20’137 genes (BUSCO C:97.4%/M:1.7%). The inflated gene number in NO could be attributed to the extremely high repeat content of the genome which makes gene prediction difficult even when the genome is masked (Roth et al. 2020), resulting in erroneous predictions. This inflated gene number bias was to some extent corrected by the orthology assignment. 19’762 genes (98.1%) of HE were attached to orthogroups, 29’875 (77.44%) of NO, and 23’907 (86.9%) of ST. To facilitate comparisons across species and their genomes, in the subsequent analyses we only used genes that had orthologues to the zebrafish *Danio rerio*. This reduced the bias in gene numbers across our species set.

Alignment of variant calling approaches.

We usedSNP data to assess allele frequency differences between the sexes and identify putative sex-linked regions. For each part, however, a different variant calling approach was applied. We wanted to avoid that assembly and annotation quality influence differential allele expression results and thus opted for a genome-independent variant calling approach in this part of the study. The KisSplice software allows to identify SNPs and alternative splicing events from raw RNA-seq reads without prior genome information (Lopez-Maestre et al. 2016). The SNPs are then placed on transcripts assembled from the same RNA-seq datasets. KisSplice was previously demonstrated to derive true allele frequency differences from expression differences (Sacomoto et al. 2012; Lopez-Maestre et al. 2016).

The GATK RNA variant call workflow is a well established, rather conservative protocol that relies on genome information. GATK-derived reference genome based data was used to search

for sex-linked regions within the genome and provided additional support for SNPs called with the genome-independent approach. To compare the two approaches, we placed pre-filtered bi-allelic SNPs called with KisSplice on the genome. We found that around 79 to 83 % of KisSplice SNPs were confirmed by the GATK bi-allelic SNPs suggesting high congruence.

Differential allele expression and sex-linked SNPs.

Differential allele expression between the sexes.

We here considered a scenario where an intralocus sexual conflict would cause a difference in allele expression. Sex differences in allelic expression point towards an emerging or ongoing sexual conflict. The accumulation of SSAs at certain locations in the genome supports actual sex-linkage of the underlying SNPs/alleles linked to the evolution of a sex chromosome. Sex chromosomes can ultimately resolve sexual conflicts by confining SA genes/alleles to the sex they benefit. Species with XY sex determination systems are expected to have considerably more or only male-specific SNPs due to their male Y chromosome, while ZW species are expected to have more or only female-specific SNPs.

Here, allele level differential expression between the sexes was looked at specifically. We categorized bi-allelic SNPs into SSAs male/female, and DEAs.

All three examined syngnathid species showed signs of sex-specific allele expression. The number of DEAs was considerably higher in NO than in the two other species assessed (Ratio: 3:1). Further, we looked at species-specific distributions of male and female SSAs. While male and female SSAs were identified in NO and ST, HE only harbored male SSAs (Figure 1. Panels A in each species). Overall, in our analysis ST exhibits the weakest divergence between the sexes in both DEA and SSA numbers. On the other side of a spectrum, HE had ~100 times more male SSAs than ST and ~17 times more than NO. To examine the genomic locations of these male SSAs identified using transcriptome assemblies, genome and transcriptome annotations were matched. The HE *chr14* appeared to harbor ~2.4 times more genes with male-specific alleles than the next best hit (124 vs. 52). With the exception of *chr14*, other notable hits came from unplaced scaffolds. Upon further examination, it appeared that both the *chr14* and the aforementioned scaffolds share all their 1-1 orthologues with *chr1* of *Hippocampus abdominalis* (HA) suggesting that they all belong to the same chromosome. With that in mind, the expanded *chr14* would be the size of ~34.3Mb (vs 13.1Mb original length), which is similar to ~36 Mb of its supposed homologue - the *chr1* of HA.

The expanded *chr14* would contain 327 genes that bear male-specific alleles (~9.5 per Mb) compared to the now next best 15 hits from *chr4* (~0.7 per Mb) This pattern marks HE as a candidate XY species.

To assess biological processes and pathways that might be affected by differential allele expression, we applied an over-representation analysis (ORA) to gene lists that contained DEAs (Figure 1. Panels B, for each species). The analysis suggested “Cell Cycle”, “ncRNA metabolic process”, and “Signaling by Rho GTPases” to be overrepresented in gene lists with DEA in all three species. The broad KEGG term “Cell Cycle” includes 469 genes involved in cell replication and nuclear division. The pathways involved in RNA processing and metabolism also appeared overrepresented in all species. The ncRNA metabolic pathway is affected by differential allele expression in all studied species, and ncRNA processing only in NO and ST. Some more broad

GO-terms "RNA degradation", "Metabolism of RNA", "tRNA metabolic process", however, appear only in some species. The third overrepresented pathway - "Signaling by Rho GTPases" - is involved in many cellular processes and helps cells to respond to external and internal stimuli via Rho GTPases - a family of molecular-switch proteins (G proteins) (Bar-Sagi and Hall 2000; Jaffe and Hall 2005; Hodge and Ridley 2016; Perry and Maddox 2019; Boueid et al. 2020). The overrepresentation of this pathway could be related to reproduction as several members of the family were found to be crucial for spermatogenesis (as examined in mammals) (Lui et al. 2003; Zhang et al. 2010; Zhu et al. 2023).

Processes related to noncoding RNAs were found to be influenced by DEA between sexes in all species. Noncoding RNAs have an important role in pregnancy progression. miRNAs regulate the physiological processes in a sexually dimorphic manner, ensuring normal fetal development, successful pregnancy, and susceptibility to disease (Varì et al. 2021). ncRNAs are implicated in promoting sex-differences, e.g. X inactivation-specific transcript (XIST) is a long ncRNA that mediates X chromosome inactivation (Brown et al. 1991). Epigenetic mechanisms are also responsible for governing gene-hormonal-environmental interactions. There is a considerable amount of data from teleost fish showing that miRNAs and lncRNAs are involved in sex determination processes and gonadal differentiation (Bizuayehu et al. 2012; Juanchich et al. 2013; Jing et al. 2014; Lau et al. 2014; Xiao et al. 2014; Juanchich et al. 2016; Juanchich et al. 2016; Tao et al. 2016; Wang et al. 2017; Pinhal et al. 2018; Qiu et al. 2018; Yan et al. 2021).

HE and NO share a higher number of pathways affected by sex-biased allele expression - mainly cell maintenance, signal transduction, intracellular transport, and various metabolic processes (Figure 1. Panels B, for mentioned species). While ST had some overrepresented pathways involved in cell cycle and metabolism, the main affected pathways appeared to be related to embryonic morphogenesis and embryo development. Such a reproduction-related pathway overrepresentation among DEAs might be particularly relevant to understanding sexual conflict in ST, considering that ST exhibits the weakest differences between sexes in both DEA and SSA numbers.

The NOD-like receptor signaling pathway is enriched in genes with DEAs in NO and ST. Among them were NLR family member X1 (*nlrx1*), nucleotide-binding oligomerization domain-containing protein 2 (*nod2*) and TNF receptor associated factor 2b (*traf2b*) in NO, TNF receptor-associated factor 5 (*traf5*) and interleukin-1 beta (*il1b*) in ST. The nucleotide-binding oligomerization domain-like (NOD-like) receptors are a type of pattern recognition receptors that play a key role in innate immune system responses (Jones et al. 2016; Chou et al. 2023).

Sex-Specific Alleles (SSAs).

The only species that had enough genes with SSAs to conduct a meaningful enrichment analysis was HE. From a list of 353 genes with male SSAs, 5 pathways appeared to be overrepresented. From GO:BP - glycerol-3-phosphate metabolic process, from KEGG - VEGF signaling pathway, C-type lectin receptor signaling pathway, Herpes simplex virus 1 infection, and from Reactome - Adaptive Immune System.

Genes from “glycerol-3-phosphate metabolic process' ' were two glycerol-3-phosphate dehydrogenases and glycerol kinase 3. Most of the genes in overrepresented adaptive and innate immune pathways were related to signal transduction, metabolism and cell cycle, and only indirectly to the immune system. However, some genes with SSAs could support the suggested sexual immune dimorphism (Roth et al. 2011; Pappert et al. 2023), namely: integrin alpha 5 (*itga5*), caspase-9 (*casp9*), mechanistic target of rapamycin kinase (*mtor*), nuclear factor of activated T cells 2a (*nfatc2a*), and cluster of differentiation 40 (*cd40*).
Both ST and NO had genes with male and female SSAs. In NO - 24M/24F genes, and in ST 6F/4M. Since a meaningful enrichment analysis is impossible with the low number of genes involved, we scanned the lists manually and looked for genes that might be involved in SA. Considering sampled tissues functions, we expected to see signatures of SA in genes related to the innate and adaptive immune systems in both tissues, and cellular transport in gill, since fathers might need more oxygen when they are pregnant (Goncalves et al. 2015). Notable hits included: huntingtin (*htt*, NO-F), cryptochrome circadian regulator 2 (*cry2*, NO-F), macroH2A.1 histone (macroh2a1, NO-F), T cell immune regulator 1, ATPase H+ ransporting V0 subunit a3b (*tcirg1b*, NO-M), transporter associated with antigen processing 1 (*tap1*, NO-M), cathepsin K (*ctsk*, NO-M), ATP binding cassette subfamily B member 9 (*abcb9*, NO-M), and integrin beta 4 (*itgb4*, ST-F).

The search for sex-linked genome regions.

We next integrated results from our sex-biased SNPs analysis above with whole-genome intersex Fst scans. Genes were sorted according to their Fst and SNP density across the genome. We further compared our results to the study by Long et al. that focused on sex chromosomes in two seahorse species, *Hippocampus erectus* and *Hippocampus abdominalis*, and the pipefish species *Syngnathus scovelli* (Long et al. 2023). This required additional kmerGWAS analysis.

Genome-wide picture.

We calculated intersex Fst values for all three genomes in windows of 1kb, 10kb and 100kb. The 10kb and 100kb scans provided an overall look at the genome and facilitated the identification of regions or chromosomes with elevated Fst values that might be indicative of sex-linkage. If a sex chromosome or sex-linked region within a genome has accumulated sufficient fixed differences between the sexes, we would expect to see a distinct Fst peak at that location. For Weir Fst estimation we expect 0.5 Fst values for a fixed difference between the homozygous and the heterozygous sex (Gammerdinger et al. 2020). For 10kb and 100kb windows we used the threshold of 0.4. We considered a lower 0.4 threshold since values will be calculated over large genomic regions that might include sites with low Fst, hence lowering the overall Fst in a given window. All seemed to have a distinct pattern in a way Fst values are distributed across the genome (Figure 2A). ST had the lowest overall Fst values with regions of elevated Fst sparsely distributed along the genome (33 10Kb >=0.4 windows or ~0.08 per Mb), which is consistent with our previous observations on sex-biased allele expression and recently published data by Long et al. Both NO and HE had similar overall Fst levels, with NO having 1222 10Kb >=0.4 windows (~0.7 per Mb) and HE 721 10Kb >=0.4 windows (~1.7 per Mb).

However, in NO the elevated regions were scattered throughout the genome, while HE had three chromosome-size scaffolds differentiating from the rest of the genome with *chr14* having the highest number of windows with elevated Fst (183 windows or ~13.9 per Mb) and the next two *chr1* and *chr18* with fewer hits (31 and 21 windows or ~1.2 and ~2.1 per Mb respectively). The fragmented nature of the genome assembly and the genome size could be responsible for sparse distribution of elevated Fst windows in NO, i.e. most of the contig hits can potentially come from a single chromosome. But considering that NO possesses both male and female SSAs (like ST), we think it is unlikely that there is a single sex-linked region or chromosome.

Elevated Fst gene scan.

The results from a 1kb-window scans were used to identify individual genes that overlap with elevated Fst windows. Using this more sensitive approach putative sex-linked genes can be identified. We used the same approach to determine the Fst threshold as with previous scans. For HE and NO we used 0.5 threshold for the reasons described above and for ST we decreased it to 0.33 to see slight bias in allele frequencies. Whenever an annotation of a particular gene overlapped with an elevated Fst window, this gene was marked and the window was annotated as belonging to this gene.
Matching the above-described scans, there was no distinct region of elevated Fst. Both NO and HE had genes with intersex Fst >0.9, which implies near-complete fixation of alleles in each sex, i.e. both sexes are homozygous at this locus and possess different alleles. Three genes in NO - proteolipid protein 1b (*plp1b* - 1.0 Fst), tetratricopeptide repeat protein 3 (*ttc3* - 0.92 Fst) and transcobalamin beta b (*tcnbb* - 0.91 Fst). Two genes in HE - interferon regulatory factor 2 (*irf2* - 0.92Fst), and most notably anti-Müllerian hormone (*amh* - 0.91 Fst). Neither of those genes belong to the same contigs, or form a distinct cluster of genes with elevated Fst. In ST, the gene with the highest value was trafficking kinesin protein 2 (*trak2* - 0.8 Fst).

Fst-SBGs overlap.

Genes that show intersex bias in expression were also found to have elevated intersex Fst. In humans and flies the Fst scores of SBGs follow a “twin peaks” pattern, i.e. genes with an intermediate bias in expression show the highest intersex Fst , interpreted as targets of ongoing sex-specific selection on viability (Cheng and Kirkpatrick 2016).

Resolution of sexual conflict may be insufficient via bias in expression in some genes, causing alleles to diverge between the sexes. SBGs with elevated intersex Fst in guppies, however, were suggested to relate to the formation of sex-specific genetic architecture rather than intralocus sexual conflict (Wright et al. 2018).

To identify genomic signatures of SAS, we investigated an overlap between SBGs and the windows with elevated Fst. Further, to assess sex-linked regions, we investigated clustering of genomic locations of SBGs with elevated intersex Fst. Because differential gene expression analysis was done using transcriptome assemblies and Fst scans using genome annotations, the results had to be compared using orthology assignment. The search was first executed per organ and then for both organs combined.

The results support the previous Fst window-scan approach. As such, in HE, the same three chromosomal scaffolds appeared to contain the most SBGs with elevated Fst: *chr14* (196 hits or ~14.9 per Mb), *chr18* (46 hits or ~4.6 per Mb), and *chr1* (28 hits or ~1.15 per Mb). Among scaffolds with high number of SBGs and elevated Fst, we identified many unplaced scaffolds e.g. *ctg37* (128 hits), *ctg26* (92 hits), *ctg27* (67 hits), *ctg33* (54 hits), *ctg39* (19 hits). However, based on our synteny analysis, these scaffolds most likely belong to the same chromosome - *chr14*. The number of overlapping genes declines from here onwards. Notably, one of the well-known sex-determination genes, anti-Mullerian hormone (*amh*), appears among the most male-biased genes in both gill and head kidney of HE, and is also among the genes with the highest Fst in the genome - 0.91 (Figure 2B). The *amh* gene region, however, does not belong to either of the two candidate sex-linked scaffolds. As such, the amh-containing scaffold appeared during Fst scans of the genome, but was among candidates with low number of elevated Fst-windows (>300 in the top scaffolds vs 5 in the amh-containing scaffold).
In NO the number of SBGs with elevated Fst was not as significant as in HE, and there were more contigs that contained similar numbers of overlapping genes, which could be a consequence of a fragmented genome. There were no known sex-determining genes in the overlap.
ST showed the smallest number of SBGs with elevated Fst, and they could only be identified in the differential gene expression analysis from the head kidney. As in other species, we could not narrow it down to a single distinct region or contig enriched in overlapping genes. Most of the overlapping genes were involved in various metabolic processes, cellular transport, and aging. However, one of the most male-biased genes with elevated Fst (0.41) was doublesex and mab-3 related transcription factor 1 (*dmrt1*) - another well-known sex-determination gene in teleosts.

As we sampled blood-rich organs, the expression profile should not solely cover the specialized processes in the organs themselves but also show some of the blood expression patterns.
We examined the lists of SBGs with elevated Fst manually for all species and organs. Their putative functions was deduced based on the information on the Zfin database (Bradford et al. 2022) followed by a literature review.
Important to note is that it is difficult to determine their role in SA relationship due to interdependence of cellular processes and pleiotropic nature of genes (Stearns 2010). Furthermore, the roles of SBGs and the reasons for differential expression in a particular organ might not be directly related to the reasons for elevated Fst values at these loci. When discussing the functions of the genes, we payed more attention to their potential roles in SA.

In HE, among the SBGs with the highest log fold change and Fst in both organs was amh - one of the key genes involved in vertebrate gonadal differentiation pathway (Morrish and Sinclair 2002; Yatsu et al. 2016; Nagahama et al. 2021; Wagner et al. 2023). Besides that, amh also appears to have functions in the brain (Pfennig et al. 2015; Silva and Giacobini 2021).
Two genes are involved in spermatogenesis. The DEAD (Asp-Glu-Ala-Asp) box polypeptide 4 (*ddx4*) appeared within SBG/Fst lists in both organs. It is an RNA helicase, mostly known as a germline factor (Raz 2000; Hickford et al. 2011; Hansen and Pelegri 2021), has functions outside of germ cell-related processes (Yajima and Wessel 2011; Lasko 2013; Schwager et al.

2015; Yajima and Wessel 2015). Interestingly, *ddx4* was found to be highly expressed in certain lung cancer cells (Noyes et al. 2023). The exosome component 10 (exosc10) was shown to be important for spermatogenesis. Mutant mice with inactivated exosc10 in male germ cells had small testes, showed impaired germ cell differentiation and were subfertile (Jamin et al. 2017). The exosc10 is also involved in inactivation of X chromosome in somatic cells via regulation of Xist expression (Ciaudo et al. 2006).
Several genes under SAS were involved in synaptic and neuronal processes.
The low density lipoprotein receptor-related protein 4 (lrp4) is one of the regulators of the Wnt/β-catenin signaling pathway and is important in the vertebrate neuromuscular junction (Shen et al. 2015; Mosca et al. 2017; Walker et al. 2021). The gene has other roles, namely in bone homeostasis as shown in mice (Chang et al. 2014; Shen et al. 2015), and roles in tooth development as shown in zebrafish (Ahn et al. 2017). The interleukin 1 receptor accessory protein-like 1b (il1rapl1b) belongs to a family of IL1/Toll receptors, regulates dendrite and synapse formation and is responsible for the X-linked intellectual disabilities in humans (Yamagata et al. 2015; Montani et al. 2018). Il1rapl1-knockout mice showed mildly impaired working memory and remote fear memory, high locomotor activity, and decreased open-space and height anxiety (Yasumura et al. 2014). In zebrafish il1rapl1b is involved in synapse formation in olfactory sensory neurons (Yoshida and Mishina 2008). Another gene in the list involved in regulation of the Wnt/β-catenin signaling - Van Gogh-like (Vangl) planar cell polarity protein 1 (vangl1) (Mentink et al. 2018). Certain variants of vangl1 are associated with neural tube defects in zebrafish, human and mice (Torban et al. 2008; Kibar et al. 2009; Reynolds et al. 2010; Yang et al. 2017). The Triple-A syndrome (aaas) gene was among the genes under SAS in HE. Mutations in aaas are responsible for the Allgrove syndrome in humans which is associated with multiple neurological abnormalities (Handschug 2001; Brooks et al. 2004; Bizzarri et al. 2013).

Two immune-related genes also appeared among genes under SAS, namely protein tyrosine phosphatase non-receptor type 22 (ptpn22) - negative regulator of T cell receptor signaling variants of which are associated with severe autoimmune diseases (Fousteri et al. 2013; Armitage et al. 2021; Tizaoui et al. 2021), and limb development membrane protein 1-like (lmbr1l) - a gene that was shown to be essential for lymphopoiesis in mice, and mutations in which causing severe immunodeficiency (Choi et al. 2019). Interestingly, lmbr1l modulates lymphopoiesis via negative regulation of the Wnt/β-catenin signalling pathway.

In NO, there were several genes potentially under SAS involved in DNA break repair, and chromosome maintenance, inactivation, and DNA replication. The tonsoku-like (tonsl) gene is a DNA repair protein certain variants of which causing abnormalities in a skeletal development (Burrage et al. 2019). Moreover, tonsl was found to be regulating epithelial cell immortalisation in breast cancer (Khatpe et al. 2023). The minichromosome maintenance complex component 4 (mcm4) - part of the minichromosome maintenance (MCM) complex responsible for DNA replication (Sheu et al. 2014). In humans, mutation in mcm4 causes genomic instability and natural killer cell deficiency (Gineau et al. 2012). The partner and localizer of BRCA2 (palb2) - important DNA double-strand break gene. In humans mutations in palb2 are associated with breast cancers (Wu et al. 2020), in mice palb2 male mutants showed reduced fertility (Simhadri

et al. 2014), and in zebrafish palb2 (fancn in the publication) knockout caused female-to-male sex reversal phenotype (Ramanagoudr-Bhojappa et al. 2018).
The H2A.X variant histone (h2ax) - a histone variant that is one of the key actors in DNA break repair and chromatin remodeling (Pinto and Flaus 2010). The h2ax is involved in sex chromosome inactivation during meiosis in male mice (Fernandez-Capetillo et al. 2003; Turner et al. 2004).

The leucine-rich repeats and WD repeat domain containing 1 (lrwd1) is presumed to be the part of origin recognition complex (ORC). ORC is responsible for initiation of DNA replication and DNA repair (Shen et al. 2010; Vermeulen et al. 2010; Liu et al. 2023).
The final gene in this group, RAD51 paralog D (rad51d), is member of the RAD51 recombinases that are highly conserved across species (Lin et al. 2006; Greenhough et al. 2023). For successful recombination RAD51 is required to interact with palb2, that assists localisation of it to DNA damage sites (Zhang et al. 2009; Greenhough et al. 2023). Mutations in rad51d are associated with high risk of ovarian cancer in humans (Suszynska et al. 2020; Greenhough et al. 2023).

Besides DNA damage repair and replication, one of the genes - ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 3 (atp2a3) was found to be among genes under selection related to hypoosmotic regulation in Takifugu species (Zhang et al. 2020).

Itersex SNP-densities and kmerGWAS results in Hippocampus erectus.

Our analyses identified HE as the most prominent candidate for containing a sex chromosome among the three investigated species. However, we identified three potential candidates for a sex chromosome: *chr14*, *chr18*, and *chr1*. The *chr14* has the highest numbers of SSAs and DEAs and has the highest number of elevated Fst windows, which makes it the most likely sex-linked chromosome candidate. At the same time, the classical sex-determination gene - *amh*, was among genes with the highest Fst values and showed highest expression bias in males. The *amh* is located on scaffold *chr13* with an overall low Fst.
In order to give additional support to *chr14* as HE sex chromosome, we compared male/female SNP densities across the genome and ran a kmerGWAS analysis to exclude that scans on unphased genomes might have driven the outcome of this study, we run kmerGWAS, since the latter does not rely on genome information. To do so, we followed the approach suggested by the SexFindR workflow. Both analyses provided similar results, highlighting *chr14* as a definite sex-chromosome candidate. The results from male/female SNP density scans were much noisier. The KmerGWAS analysis resulted in a clear peak for *chr14* (Figure 3B).

Long et al. data. Comparison.

Long and colleagues had recently published a study investigating sex chromosomes in three syngnathids, including HE. In that study, the HE genome was reconstructed using the genome of *Hippocampus abdominalis* (HA), a species phylogenetically more derived species compared

to HE (Betancur-R et al. 2017) with a larger genome that underwent an additional chromosome fusion (He et al. 2022; Long et al. 2023).
To compare our results to Long et al. 2022, we performed a synteny analysis between the HA genome used by Long and the HE genome at our disposal (PRJNA613176) (Li et al. 2021). For this purpose, we used annotation data from these genomes, identified 1-1 orthologs with Orthofinder, and used locations of these orthologues in each genome to identify homologous chromosomes (Figure 3C). The sex chromosome that we identified (*chr14*) and *H. abdominalis chr4* identified by Long et al. were not homologous. For additional tests, the data published by Long et al. 2022 were used. We chose kmerGWAS because we obtained the cleanest signal using this method, and also because it was not based on information from the genome. The resulting sex-associated kmers were aligned to HE genome. The results of kmerGWAS confirmed Long's finding, although the signal was not as significant as in HE using our data (Figure 3A). This suggests that in HE, just like in some other teleost species, sex chromosomes are non stable and could be population/strain dependent.

DISCUSSION

Sex-linked regions, sex chromosomes, sex determination and sexual antagonism.

A large sex-linked region or sex-linked chromosome is normally assumed to contain a master SD gene or multiple SD genes, even when the evidence only supports sex-linkage and no actual SD locus was identified. In vertebrates, the “choice” of an SD gene is confined within the canonical gonadal differentiation pathway or genes that directly interact with it (Bertho et al. 2018). Even when a turnover event occurs, a young sex chromosome is expected to harbor an SD locus.
Latest study by Long and colleagues identified sex-linked chromosomes in *Hippocampus erectus (chr11 or* homologue of *abdominalis chr4)* and *Hippocampus abdominalis (chr6)* (HA). From the data, it was concluded that the sex-linked chromosomes in these two species are homomorphic, show low sequence divergence between the sexes and that they are probably young.
In our study we also identified a sex-linked chromosome in HE, however it was not homologous to the chromosome identified by Long et al. There are two possible explanations for such a result: i) methodological differences and errors, ii) both results represent real biological differences between the two breeds of HE.
Methodologically the two studies are different:

i Our study used RNA data to identify the sex-linked chromosome - they used DNA.
ii We used the chromosome-level HE genome - Long and colleagues reconstructed the HE genome based on a genome of a more distant species - HA.
iii Furthermore, we used different tools and metrics to identify sex-linkage.

It is unlikely that differences in the sources of sequencing data, RNA or DNA, would have caused such diverging results, especially considering that both studies used multiple

approaches to identify sex-linked regions. We decided to confirm the results described in Long et al. using their re-sequencing data from HE males and females, together with the chromosome-level assembly of HE from (Li et al. 2021) using kmerGWAS. This approach is alternative to the one used in Long et al., but the one that gave us the most conclusive results. This way we were able to compare our results.

The comparison of the previously published data with our results suggested that breeds of HE have independently evolved distinct solutions of sexual antagonistic conflict resulting in two different sex-linked chromosome systems. Differences in sex-liked regions between captive breeds and wild-types have been observed before in zebrafish and Siamese fighting fish breeds (Wilson et al. 2014; Panthum et al. 2022; Wang et al. 2022; Schartl et al. 2023) and cichlids (Lichilín et al. 2023). Zebrafish wild-types have a ZW SD system, and it has been hypothesized that domestication led to the loss of the Z chromosome. The WW individuals become either females or neomales, but the SD mechanism in such a breed is unknown. In Siamese fighting fish breeds some have a monogenic SD system with dmrt1 as a master SD gene, while some have polygenic SD.

These findings suggest that the identified sex-linked chromosomes in HE are younger than originally believed by Long. Both of our fish originated from cultured sources. Long's from a cultured population in Fujian (China), seed population of which consisted between 2,000 -3,000 wild-caught individuals. Our fish was obtained through a German aquarium breeder with unknown seed population size. It is possible that a "zebrafish scenario" repeated itself here, with two captive-bred populations having distinct sex-linked chromosomes due to domestication. It is another question whether these sex-linked chromosomes actually contain a master SD locus. As both in our case, and in Long's data, the Fst values across these chromosomes suggest incomplete linkage.

As sharply noted by Long et al., the recombination landscape could pre-date a turnover event. It has been shown before in many organisms that genes with sex-biased expression patterns also show elevated intersex Fst. In humans and flies this relationship followed a "Twin Peaks" pattern with genes exhibiting intermediate bias in expression showing the highest Fst values (Cheng and Kirkpatrick 2016). This, as suggested by authors, was a sign of ongoing sexual selection on those genes, whereas for others the sexual conflict was resolved via significant bias in expression.

In cichlids, the intersex Fst increased with bias in expression, i.e. the most extreme SBGs also had the highest intersex Fst (Lichilín et al. 2021). Similar results have been obtained in guppies (Wright et al. 2018). No matter what the underlying mechanisms might be, it is clear that bias in expression often is not enough to fully resolve sexual conflict.

Even though the identified sex-linked chromosome is young, we observed an enrichment of SBGs on it compared to the rest of the genome. Together with elevated intersex Fst pattern, kmerGWAS results or GWAS as in Long's study, one might assume that what we see here is indeed a true sex chromosome, i.e. the one that also harbours an SD gene, and elevated intersex Fst values are the result of ancestral low recombination rates. Long proposed several promising SD candidate genes, though refrains from giving a definitive answer.

On our chromosome, however, we were unable to identify potential candidates related to the canonical vertebrate gonadal differentiation pathway. At the same time we observed both

extreme intersex Fst (0.9) and extreme bias in expression for a well-known sex determination gene - the *amh*.
It has been recently suggested that many multi-locus SD systems actually represent transient states during turnover events (Schartl et al. 2023). We suggest that the observed sex-linked chromosome and the *amh* locus with near-fixation Fst represent such a transient state. We propose that initially, the chromosome that contained the *amh* gene as an SD locus was the only sex-linked chromosome in HE. The *amh* falls within the canonical vertebrate gonadal differentiation pathway and genes related to *amh,* like anti-Müllerian hormone type II receptor (*amhr2),* were found to be SD loci in other syngnathid species (10.1126/sciadv.abg5196). Due to the closed nature of the aquaria system and absence of natural predators, the effects of SS and SAS have amplified due to the absence of other selective forces that a population would have experienced in nature. The identified sex-linked chromosome was enriched for SBGs with elevated Fst values, which is unusual for such a young sex chromosome. Ongoing intralocus sexual conflict caused bias in expression and promoted genetic linkage. The captive breeding with a limited number of mates must have accelerated the fixation process. This could have created an overall reduced recombination landscape within a genome and facilitated linkage of *chr14* (or *chr11* in the other HE breed). While an entire chromosome became sex-linked, the master SD locus (*amh*) remained the same, thus creating a multi-locus SD system.
There are still questions regarding the speed of such transformation, the exact mechanisms for recombination slowdown, relationships between bias in expression and elevated Fst, and the exact role of SA in this process. However, if confirmed, this offers a lucrative system to study the effects of SS and SAS on the genome.

Ethics statement

Work was carried out in accordance with German animal welfare law and with the ethical approval given by the Ministerium für Energiewende, Landwirtschaft, Umwelt, Natur und Ditgitalisierung (MELUND) Schleswig-Holstein (permit no. V242-57983/2018). No wild endangered species were used in this investigation.

Authors’ contributions

A.D., O.R. and A.B. planned the project. J.P. collected and processed samples. A.D. analysed the data and interpreted the results with input from O.R. and A.B.. A.D. wrote the manuscript with input from all coauthors.

Funding

This project was supported by funding from the German Research Foundation (RO-4628/4-2) and from the European Research Council (ERC) under the European Union´s Horizon research and innovation program (MALEPREG: eu-repo/grantAgreement/EC/H2020/755659) to OR.

Data Access

All sequencing data generated in this project can be found under BioProject ID PRJNA947442.

References

Ahn Y, Sims C, Murray MJ, Kuhlmann PK, Fuentes-Antrás J, Weatherbee SD, Krumlauf R. 2017. Multiple modes of Lrp4 function in modulation of Wnt/β-catenin signaling during tooth development. *Development* [Internet]:dev.150680. Available from: https://journals.biologists.com/dev/article/doi/10.1242/dev.150680/264498/Multiple-modes-of-Lrp4-function-in-modulation-of

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* [Internet] 215:403–410. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0022283605803602

Armitage LH, Wallet MA, Mathews CE. 2021. Influence of PTPN22 Allotypes on Innate and Adaptive Immune Function in Health and Disease. *Front. Immunol.* [Internet] 12:636618. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2021.636618/full

Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman T-L, Hahn MW, Kitano J, Mayrose I, Ming R, et al. 2014. Sex Determination: Why So Many Ways of Doing It? *PLoS Biol.* [Internet] 12:e1001899. Available from: https://dx.plos.org/10.1371/journal.pbio.1001899

Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT. 2011. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* [Internet] 27:1691–1692. Available from: https://academic.oup.com/bioinformatics/article/27/12/1691/255399

Bar-Sagi D, Hall A. 2000. Ras and Rho GTPases. *Cell* [Internet] 103:227–238. Available from: https://linkinghub.elsevier.com/retrieve/pii/S009286740000115X

Benoit-Pilven C, Marchet C, Chautard E, Lima L, Lambert M-P, Sacomoto G, Rey A, Cologne A, Terrone S, Dulaurier L, et al. 2018. Complementarity of assembly-first and mapping-first approaches for alternative splicing annotation and differential analysis from RNAseq data. *Sci. Rep.* [Internet] 8:4307. Available from: https://www.nature.com/articles/s41598-018-21770-7

Bertho S, Herpin A, Branthonne A, Jouanno E, Yano A, Nicol B, Muller T, Pannetier M, Pailhoux E, Miwa M, et al. 2018. The unusual rainbow trout sex determination gene hijacked the canonical vertebrate gonadal differentiation pathway. *Proc. Natl. Acad. Sci.* [Internet] 115:12781–12786. Available from: https://pnas.org/doi/full/10.1073/pnas.1803826115

Betancur-R R, Wiley EO, Arratia G, Acero A, Bailly N, Miya M, Lecointre G, Ortí G. 2017. Phylogenetic classification of bony fishes. *BMC Evol. Biol.* [Internet] 17:162. Available from: https://bmcevolbiol.biomedcentral.com/articles/10.1186/s12862-017-0958-3

Bizuayehu TT, Lanes CF, Furmanek T, Karlsen BO, Fernandes JM, Johansen SD, Babiak I. 2012. Differential expression patterns of conserved miRNAs and isomiRs during Atlantic halibut development. *BMC Genomics* [Internet] 13:11. Available from: https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-13-11

Bizzarri C, Benevento D, Terzi C, Huebner A, Cappa M. 2013. Triple A (Allgrove) syndrome: an unusual association with syringomyelia. *Ital. J. Pediatr.* [Internet] 39:39. Available from: http://ijponline.biomedcentral.com/articles/10.1186/1824-7288-39-39

Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL. 2012. Domain enhanced lookup time accelerated BLAST. *Biol. Direct* [Internet] 7:12. Available from: http://biologydirect.biomedcentral.com/articles/10.1186/1745-6150-7-12

Boueid M-J, Mikdache A, Lesport E, Degerny C, Tawk M. 2020. Rho GTPases Signaling in Zebrafish Development and Disease. *Cells* [Internet] 9:2634. Available from: https://www.mdpi.com/2073-4409/9/12/2634

Bradford YM, Van Slyke CE, Ruzicka L, Singer A, Eagle A, Fashena D, Howe DG, Frazer K, Martin R, Paddock H, et al. 2022. Zebrafish information network, the knowledgebase for *Danio rerio* research.Wood V, editor. *Genetics* [Internet] 220:iyac016. Available from: https://academic.oup.com/genetics/article/doi/10.1093/genetics/iyac016/6528852

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* [Internet] 34:525–527. Available from: https://www.nature.com/articles/nbt.3519

Broad Institute. 2019. Picard toolkit. *Broad Inst. GitHub Repos.* [Internet]. Available from: https://broadinstitute.github.io/picard/

Brooks BP, Kleta R, Caruso RC, Stuart C, Ludlow J, Stratakis CA. 2004. Triple-A syndrome with prominent ophthalmic features and a novel mutation in the AAASgene: a case report. *BMC Ophthalmol.* [Internet] 4:7. Available from: http://bmcophthalmol.biomedcentral.com/articles/10.1186/1471-2415-4-7

Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF. 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* [Internet] 349:38–44. Available from: https://www.nature.com/articles/349038a0

Brůna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M. 2021. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics Bioinforma.* [Internet] 3:lqaa108. Available from: https://academic.oup.com/nargab/article/doi/10.1093/nargab/lqaa108/6066535

Brůna T, Lomsadze A, Borodovsky M. 2020. GeneMark-EP+: eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR Genomics Bioinforma.* [Internet] 2:lqaa026. Available from: https://academic.oup.com/nargab/article/doi/10.1093/nargab/lqaa026/5836691

Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* [Internet] 12:59–60. Available from: https://www.nature.com/articles/nmeth.3176

Burrage LC, Reynolds JJ, Baratang NV, Phillips JB, Wegner J, McFarquhar A, Higgs MR, Christiansen AE, Lanza DG, Seavitt JR, et al. 2019. Bi-allelic Variants in TONSL Cause SPONASTRIME Dysplasia and a Spectrum of Skeletal Dysplasia Phenotypes. *Am. J. Hum. Genet.* [Internet] 104:422–438. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002929719300072

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* [Internet] 10:421. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-10-421

Chang M-K, Kramer I, Huber T, Kinzel B, Guth-Gundel S, Leupin O, Kneissel M. 2014. Disruption of Lrp4 function by genetic deletion or pharmacological blockade increases bone mass and serum sclerostin levels. *Proc. Natl. Acad. Sci.* [Internet] 111. Available from: https://pnas.org/doi/full/10.1073/pnas.1413828111

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* [Internet] 34:i884–i890. Available from: https://academic.oup.com/bioinformatics/article/34/17/i884/5093234

Cheng C, Kirkpatrick M. 2016. Sex-Specific Selection and Sex-Biased Gene Expression in Humans and Flies.Nachman MW, editor. *PLOS Genet.* [Internet] 12:e1006170. Available from: https://dx.plos.org/10.1371/journal.pgen.1006170

Choi JH, Zhong X, McAlpine W, Liao T-C, Zhang D, Fang B, Russell J, Ludwig S, Nair-Gill E, Zhang Z, et al. 2019. LMBR1L regulates lymphopoiesis through Wnt/β-catenin signaling. *Science* [Internet] 364:eaau0812. Available from: https://www.science.org/doi/10.1126/science.aau0812

Chou W-C, Jha S, Linhoff MW, Ting JP-Y. 2023. The NLR gene family: from discovery to present day. *Nat. Rev. Immunol.* [Internet]. Available from: https://www.nature.com/articles/s41577-023-00849-x

Ciaudo C, Bourdet A, Cohen-Tannoudji M, Dietz HC, Rougeulle C, Avner P. 2006. Nuclear mRNA Degradation Pathway(s) Are Implicated in Xist Regulation and X Chromosome Inactivation.Reik W, editor. *PLoS Genet.* [Internet] 2:e94. Available from: https://dx.plos.org/10.1371/journal.pgen.0020094

Dainat J, Hereñú D, Dr. K. D. Murray, Davis E, Crouch K, LucileSol, Agostinho N, Pascal-Git, Zollman Z, Tayyrov. 2023. NBISweden/AGAT: AGAT-v1.2.0. Available from: https://zenodo.org/record/3552717

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. *Bioinformatics* [Internet] 27:2156–2158. Available from: https://academic.oup.com/bioinformatics/article/27/15/2156/402296

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* [Internet] 29:15–21. Available from: https://academic.oup.com/bioinformatics/article/29/1/15/272537

Dudaniec RY, Carey AR, Svensson EI, Hansson B, Yong CJ, Lancaster LT. 2022. Latitudinal clines in sexual selection, sexual size dimorphism and sex□specific genetic dispersal during a poleward range expansion. *J. Anim. Ecol.* [Internet] 91:1104–1118. Available from: https://onlinelibrary.wiley.com/doi/10.1111/1365-2656.13488

El Taher A, Ronco F, Matschiner M, Salzburger W, Böhne A. 2021. Dynamics of sex chromosome evolution in a rapid radiation of cichlid fishes. *Sci. Adv.* [Internet] 7:eabe8215. Available from: https://www.science.org/doi/10.1126/sciadv.abe8215

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* [Internet] 20:238. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1832-y

Fernandez-Capetillo O, Mahadevaiah SK, Celeste A, Romanienko PJ, Camerini-Otero RD, Bonner WM, Manova K, Burgoyne P, Nussenzweig A. 2003. H2AX Is Required for Chromatin Remodeling and Inactivation of Sex Chromosomes in Male Mouse Meiosis. *Dev. Cell* [Internet] 4:497–508. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1534580703000935

Fousteri G, Liossis S-NC, Battaglia M. 2013. Roles of the protein tyrosine phosphatase PTPN22 in immunity and autoimmunity. *Clin. Immunol.* [Internet] 149:556–565. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1521661613002647

Furman BLS, Metzger DCH, Darolti I, Wright AE, Sandkam BA, Almeida P, Shu JJ, Mank JE. 2020. Sex Chromosome Evolution: So Many Exceptions to the Rules.Fraser B, editor. *Genome Biol. Evol.* [Internet] 12:750–763. Available from: https://academic.oup.com/gbe/article/12/6/750/5823304

Gabriel L, Hoff KJ, Brůna T, Borodovsky M, Stanke M. 2021. TSEBRA: transcript selector for BRAKER. *BMC Bioinformatics* [Internet] 22:566. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-021-04482-0

Gammerdinger WJ, Toups MA, Vicoso B. 2020. Disagreement in F_{ST} estimators: A case study from sex chromosomes. *Mol. Ecol. Resour.* [Internet] 20:1517–1525. Available from: https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13210

Gineau L, Cognet C, Kara N, Lach FP, Dunne J, Veturi U, Picard C, Trouillet C, Eidenschenk C, Aoufouchi S, et al. 2012. Partial MCM4 deficiency in patients with growth retardation, adrenal insufficiency, and natural killer cell deficiency. *J. Clin. Invest.* [Internet] 122:821–832. Available from: http://www.jci.org/articles/view/61014

Goncalves IB, Ahnesjö I, Kvarnemo C. 2015. Embryo oxygenation in pipefish brood pouches: novel insights. *J. Exp. Biol.* [Internet] 218:1639–1646. Available from: https://journals.biologists.com/jeb/article/218/11/1639/13760/Embryo-oxygenation-in-pipefish-brood-pouches-novel

Gotoh O. 2008. A space-efficient and accurate method for mapping and aligning cDNA sequences onto genomic sequence. *Nucleic Acids Res.* [Internet] 36:2630–2638. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkn105

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* [Internet] 29:644–652. Available from: https://www.nature.com/articles/nbt.1883

Grath S, Parsch J. 2016. Sex-Biased Gene Expression. *Annu. Rev. Genet.* [Internet] 50:29–44. Available from: https://www.annualreviews.org/doi/10.1146/annurev-genet-120215-035429

Grayson P, Wright A, Garroway CJ, Docker MF. 2022. SexFindR: A computational workflow to identify young and old sex chromosomes. Evolutionary Biology Available from: http://biorxiv.org/lookup/doi/10.1101/2022.02.21.481346

Greenhough LA, Liang C-C, Belan O, Kunzelmann S, Maslen S, Rodrigo-Brenni MC, Anand R, Skehel M, Boulton SJ, West SC. 2023. Structure and function of the RAD51B–RAD51C–RAD51D–XRCC2 tumour suppressor. *Nature* [Internet] 619:650–657. Available from: https://www.nature.com/articles/s41586-023-06179-1

Gremme G, Steinbiss S, Kurtz S. 2013. GenomeTools: A Comprehensive Software Library for Efficient Processing of Structured Genome Annotations. *IEEE/ACM Trans. Comput. Biol. Bioinform.* [Internet] 10:645–656. Available from: http://ieeexplore.ieee.org/document/6529082/

Handschug K. 2001. Triple A syndrome is caused by mutations in AAAS, a new WD-repeat protein gene. *Hum. Mol. Genet.* [Internet] 10:283–290. Available from: https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/10.3.283

Hansen CL, Pelegri F. 2021. Primordial Germ Cell Specification in Vertebrate Embryos: Phylogenetic Distribution and Conserved Molecular Features of Preformation and Induction. *Front. Cell Dev. Biol.* [Internet] 9:730332. Available from: https://www.frontiersin.org/articles/10.3389/fcell.2021.730332/full

He L, Long X, Qi J, Wang Z, Huang Z, Wu S, Zhang X, Luo H, Chen X, Lin J, et al. 2022. Genome and gene evolution of seahorse species revealed by the chromosome□level genome of *Hippocampus abdominalis*. *Mol. Ecol. Resour.* [Internet] 22:1465–1477. Available from: https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.13541

Hickford DE, Frankenberg S, Pask AJ, Shaw G, Renfree MB. 2011. DDX4 (VASA) Is Conserved in Germ Cell Development in Marsupials and Monotremes1. *Biol. Reprod.* [Internet] 85:733–743. Available from: https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.111.091629

Hodge RG, Ridley AJ. 2016. Regulating Rho GTPases and their regulators. *Nat. Rev. Mol. Cell Biol.* [Internet] 17:496–510. Available from: https://www.nature.com/articles/nrm.2016.67

Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. 2016. BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics* [Internet] 32:767–769. Available from: https://academic.oup.com/bioinformatics/article/32/5/767/1744611

Hoff KJ, Lomsadze A, Borodovsky M, Stanke M. 2019. Whole-Genome Annotation with BRAKER. In: Kollmar M, editor. Gene Prediction. Vol. 1962. Methods in Molecular Biology. New York, NY: Springer New York. p. 65–95. Available from: http://link.springer.com/10.1007/978-1-4939-9173-0_5

Iversen LL, Svensson EI, Christensen ST, Bergsten J, Sand-Jensen K. 2019. Sexual conflict and intrasexual polymorphism promote assortative mating and halt population differentiation. *Proc. R. Soc. B Biol. Sci.* [Internet] 286:20190251. Available from: https://royalsocietypublishing.org/doi/10.1098/rspb.2019.0251

Iwata H, Gotoh O. 2012. Benchmarking spliced alignment programs including Spaln2, an extended version of Spaln that incorporates additional species-specific features. *Nucleic Acids Res.* [Internet] 40:e161–e161. Available from: https://academic.oup.com/nar/article/40/20/e161/2414522

Jackman SD, Vandervalk BP, Mohamadi H, Chu J, Yeo S, Hammond SA, Jahesh G, Khan H, Coombe L, Warren RL, et al. 2017. ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. *Genome Res.* [Internet] 27:768–777. Available from: http://genome.cshlp.org/lookup/doi/10.1101/gr.214346.116

Jaffe AB, Hall A. 2005. RHO GTPASES: Biochemistry and Biology. *Annu. Rev. Cell Dev. Biol.* [Internet] 21:247–269. Available from: https://www.annualreviews.org/doi/10.1146/annurev.cellbio.21.020604.150721

Jamin SP, Petit FG, Kervarrec C, Smagulova F, Illner D, Scherthan H, Primig M. 2017. EXOSC10/Rrp6 is post-translationally regulated in male germ cells and controls the onset of spermatogenesis. *Sci. Rep.* [Internet] 7:15065. Available from: https://www.nature.com/articles/s41598-017-14643-y

Jing J, Wu J, Liu W, Xiong S, Ma W, Zhang J, Wang W, Gui J-F, Mei J. 2014. Sex-Biased miRNAs in Gonad and Their Potential Roles for Testis Development in Yellow Catfish.Qu L-H, editor. *PLoS ONE* [Internet] 9:e107946. Available from: https://dx.plos.org/10.1371/journal.pone.0107946

Jones JDG, Vance RE, Dangl JL. 2016. Intracellular innate immune surveillance devices in plants and animals. *Science* [Internet] 354:aaf6395. Available from: https://www.science.org/doi/10.1126/science.aaf6395

Juanchich A, Bardou P, Rué O, Gabillard J-C, Gaspin C, Bobe J, Guiguen Y. 2016. Characterization of an extensive rainbow trout miRNA transcriptome by next generation

sequencing. *BMC Genomics* [Internet] 17:164. Available from: http://www.biomedcentral.com/1471-2164/17/164

Juanchich A, Le Cam A, Montfort J, Guiguen Y, Bobe J. 2013. Identification of Differentially Expressed miRNAs and Their Potential Targets During Fish Ovarian Development1. *Biol. Reprod.* [Internet] 88. Available from: https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.112.105361

Kent WJ. 2002. BLAT —The BLAST -Like Alignment Tool. *Genome Res.* [Internet] 12:656–664. Available from: http://genome.cshlp.org/lookup/doi/10.1101/gr.229202

Khatpe AS, Dirks R, Bhat-Nakshatri P, Mang H, Batic K, Swiezy S, Olson J, Rao X, Wang Y, Tanaka H, et al. 2023. TONSL Is an Immortalizing Oncogene and a Therapeutic Target in Breast Cancer. *Cancer Res.* [Internet] 83:1345–1360. Available from: https://aacrjournals.org/cancerres/article/83/8/1345/725108/TONSL-Is-an-Immortalizing-Oncogene-and-a

Kibar Z, Bosoi CM, Kooistra M, Salem S, Finnell RH, De Marco P, Merello E, Bassuk AG, Capra V, Gros P. 2009. Novel mutations in *VANGL1* in neural tube defects. *Hum. Mutat.* [Internet] 30:E706–E715. Available from: https://onlinelibrary.wiley.com/doi/10.1002/humu.21026

Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM. 2019. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res.* [Internet] 47:D807–D811. Available from: https://academic.oup.com/nar/article/47/D1/D807/5160989

Lasko P. 2013. The DEAD-box helicase Vasa: Evidence for a multiplicity of functions in RNA processes and developmental biology. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* [Internet] 1829:810–816. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1874939913000680

Lau K, Lai KP, Bao JYJ, Zhang N, Tse A, Tong A, Li JW, Lok S, Kong RYC, Lui WY, et al. 2014. Identification and Expression Profiling of MicroRNAs in the Brain, Liver and Gonads of Marine Medaka (Oryzias melastigma) and in Response to Hypoxia.Antoniewski C, editor. *PLoS ONE* [Internet] 9:e110698. Available from: https://dx.plos.org/10.1371/journal.pone.0110698

Levins R, Lewontin RC. 1985. The dialectical biologist. Cambridge (Mass.) London: Harvard university press

Li C, Olave M, Hou Y, Qin G, Schneider RF, Gao Z, Tu X, Wang X, Qi F, Nater A, et al. 2021. Genome sequences reveal global dispersal routes and suggest convergent genetic adaptations in seahorse evolution. *Nat. Commun.* [Internet] 12:1094. Available from: https://www.nature.com/articles/s41467-021-21379-x

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* [Internet] 25:2078–2079. Available from: https://academic.oup.com/bioinformatics/article/25/16/2078/204688

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* [Internet] 22:1658–1659. Available from: https://academic.oup.com/bioinformatics/article/22/13/1658/194225

Lichilín N, El Taher A, Böhne A. 2021. Sex-biased gene expression and recent sex chromosome turnover. *Philos. Trans. R. Soc. B Biol. Sci.* [Internet] 376:20200107. Available from: https://royalsocietypublishing.org/doi/10.1098/rstb.2020.0107

Lichilín N, Salzburger W, Böhne A. 2023. No evidence for sex chromosomes in natural populations of the cichlid fish *Astatotilapia burtoni*.Betran E, editor. *G3 GenesGenomesGenetics* [Internet] 13:jkad011. Available from: https://academic.oup.com/g3journal/article/doi/10.1093/g3journal/jkad011/6989787

Lin Z, Kong H, Nei M, Ma H. 2006. Origins and evolution of the *recA* / *RAD51* gene family: Evidence for ancient gene duplication and endosymbiotic gene transfer. *Proc. Natl. Acad. Sci.* [Internet] 103:10328–10333. Available from: https://pnas.org/doi/full/10.1073/pnas.0604232103

Liu D, Sonalkar J, Prasanth SG. 2023. ORChestra coordinates the replication and repair music. *BioEssays* [Internet] 45:2200229. Available from: https://onlinelibrary.wiley.com/doi/10.1002/bies.202200229

Lomsadze A. 2005. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* [Internet] 33:6494–6506. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gki937

Lomsadze A, Burns PD, Borodovsky M. 2014. Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Res.* [Internet] 42:e119–e119. Available from: https://academic.oup.com/nar/article/42/15/e119/2434516

Long X, Charlesworth D, Qi J, Wu R, Chen M, Wang Z, Xu L, Fu H, Zhang X, Chen X, et al. 2023. Independent Evolution of Sex Chromosomes and Male Pregnancy–Related Genes in Two Seahorse Species.Wilson M, editor. *Mol. Biol. Evol.* [Internet] 40:msac279. Available from: https://academic.oup.com/mbe/article/doi/10.1093/molbev/msac279/6964685

Lopez-Maestre H, Brinza L, Marchet C, Kielbassa J, Bastien S, Boutigny M, Monnin D, Filali AE, Carareto CM, Vieira C, et al. 2016. SNP calling from RNA-seq data without a reference genome: identification, quantification, differential analysis and impact on the protein sequence. *Nucleic Acids Res.* [Internet]:gkw655. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw655

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* [Internet] 15:550. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8

Lui W-Y, Lee WM, Cheng CY. 2003. Rho GTPases and spermatogenesis. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* [Internet] 1593:121–129. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0167488902003488

Lumley AJ, Michalczyk Ł, Kitson JJN, Spurgin LG, Morrison CA, Godwin JL, Dickinson ME, Martin OY, Emerson BC, Chapman T, et al. 2015. Sexual selection protects against extinction. *Nature* [Internet] 522:470–473. Available from: https://www.nature.com/articles/nature14419

Mank JE. 2017. Population genetics of sexual conflict in the genomic era. *Nat. Rev. Genet.* [Internet] 18:721–730. Available from: https://www.nature.com/articles/nrg.2017.83

Mank Judith E, Hall DW, Kirkpatrick M, Avise JC. 2006. Sex chromosomes and male ornaments: a comparative evaluation in ray-finned fishes. *Proc. R. Soc. B Biol. Sci.* [Internet] 273:233–236. Available from: https://royalsocietypublishing.org/doi/10.1098/rspb.2005.3334

Mank Judith E., Promislow DEL, Avise JC. 2006. Evolution of alternative sex-determining mechanisms in teleost fishes: FISH SEX DETERMINATION MECHANISMS. *Biol. J. Linn. Soc.* [Internet] 87:83–93. Available from: https://academic.oup.com/biolinnean/article-lookup/doi/10.1111/j.1095-8312.2006.00558.x

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes.Kelley J, editor. *Mol. Biol. Evol.* [Internet] 38:4647–4654. Available from: https://academic.oup.com/mbe/article/38/10/4647/6329644

Martínez-Ruiz C, Knell RJ. 2017. Sexual selection can both increase and decrease extinction probability: reconciling demographic and evolutionary factors.Bassar R, editor. *J. Anim. Ecol.* [Internet] 86:117–127. Available from: https://onlinelibrary.wiley.com/doi/10.1111/1365-2656.12601

Mentink RA, Rella L, Radaszkiewicz TW, Gybel T, Betist MC, Bryja V, Korswagen HC. 2018. The planar cell polarity protein VANG-1/Vangl negatively regulates Wnt/β-catenin signaling through a Dvl dependent mechanism.Chisholm AD, editor. *PLOS Genet.* [Internet] 14:e1007840. Available from: https://dx.plos.org/10.1371/journal.pgen.1007840

Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* [Internet] 34:i142–i150. Available from: https://academic.oup.com/bioinformatics/article/34/13/i142/5045727

Montani C, Gritti L, Beretta S, Verpelli C, Sala C. 2018. The synaptic and neuronal functions of the X□linked intellectual disability protein Interleukin□1 receptor accessory protein like 1 (IL 1 RAPL 1). *Dev. Neurobiol.* [Internet]:dneu.22657. Available from: https://onlinelibrary.wiley.com/doi/10.1002/dneu.22657

Morrish B, Sinclair A. 2002. Vertebrate sex determination: many means to an end. *Reproduction* [Internet]:447–457. Available from: https://rep.bioscientifica.com/view/journals/rep/124/4/447.xml

Mosca TJ, Luginbuhl DJ, Wang IE, Luo L. 2017. Presynaptic LRP4 promotes synapse number and function of excitatory CNS neurons. *eLife* [Internet] 6:e27347. Available from: https://elifesciences.org/articles/27347

Nagahama Y, Chakraborty T, Paul-Prasanth B, Ohta K, Nakamura M. 2021. Sex determination, gonadal sex differentiation, and plasticity in vertebrate species. *Physiol. Rev.* [Internet] 101:1237–1308. Available from: https://journals.physiology.org/doi/10.1152/physrev.00044.2019

Noyes C, Kitajima S, Li F, Suita Y, Miriyala S, Isaac S, Ahsan N, Knelson E, Vajdi A, Tani T, et al. 2023. The germline factor DDX4 contributes to the chemoresistance of small cell lung cancer cells. *Commun. Biol.* [Internet] 6:65. Available from: https://www.nature.com/articles/s42003-023-04444-7

Nygård M, Kvarnemo C, Ahnesjö I, Braga Goncalves I. 2019. Pipefish embryo oxygenation, survival, and development: egg size, male size, and temperature effects.St Mary C, editor. *Behav. Ecol.* [Internet] 30:1451–1460. Available from: https://academic.oup.com/beheco/article/30/5/1451/5525172

Ou S, Jiang N. 2018. LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons. *Plant Physiol.* [Internet] 176:1410–1422. Available from: https://academic.oup.com/plphys/article/176/2/1410-1422/6117145

Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, Lugo CSB, Elliott TA, Ware D, Peterson T, et al. 2019. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* [Internet] 20:275. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1905-y

Panthum T, Jaisamut K, Singchat W, Ahmad SF, Kongkaew L, Wongloet W, Dokkaew S, Kraichak E, Muangmai N, Duengkae P, et al. 2022. Something Fishy about Siamese Fighting Fish (Betta splendens) Sex: Polygenic Sex Determination or a Newly Emerged Sex-Determining Region? *Cells* [Internet] 11:1764. Available from: https://www.mdpi.com/2073-4409/11/11/1764

Pappert FA, Dubin A, Torres G, Roth O. 2023. Sex or sex role – unravelling the evolutionary route of sex-specific senescence. Evolutionary Biology Available from: http://biorxiv.org/lookup/doi/10.1101/2023.05.02.539036

Parker J, Dubin A, Roth O. 2023. Genome rearrangements, male pregnancy and immunological tolerance – the curious case of the syngnathid immune system. *Front. Mar. Sci.* [Internet] 10:1099231. Available from: https://www.frontiersin.org/articles/10.3389/fmars.2023.1099231/full

Parsch J, Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nat. Rev. Genet.* [Internet] 14:83–87. Available from: https://www.nature.com/articles/nrg3376

Perry JA, Maddox AS. 2019. Uncovering the secret life of Rho GTPases. *eLife* [Internet] 8:e53276. Available from: https://elifesciences.org/articles/53276

Pfennig F, Standke A, Gutzeit HO. 2015. The role of Amh signaling in teleost fish – Multiple functions not restricted to the gonads. *Gen. Comp. Endocrinol.* [Internet] 223:87–107. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0016648015002695

Pinhal D, Bovolenta LA, Moxon S, Oliveira AC, Nachtigall PG, Acencio ML, Patton JG, Hilsdorf AWS, Lemke N, Martins C. 2018. Genome-wide microRNA screening in Nile tilapia reveals pervasive isomiRs' transcription, sex-biased arm switching and increasing complexity of expression throughout development. *Sci. Rep.* [Internet] 8:8248. Available from: https://www.nature.com/articles/s41598-018-26607-x

Pinto DMS, Flaus A. 2010. Structure and Function of Histone H2AX. In: Nasheuer H-P, editor. Genome Stability and Human Diseases. Vol. 50. Subcellular Biochemistry. Dordrecht: Springer Netherlands. p. 55–78. Available from: http://link.springer.com/10.1007/978-90-481-3471-7_4

Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van Der Auwera GA, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, et al. 2017. Scaling accurate genetic variant discovery to tens of thousands of samples. Genomics Available from: http://biorxiv.org/lookup/doi/10.1101/201178

Press CMcL, Evensen Ø. 1999. The morphology of the immune system in teleost fishes. *Fish Shellfish Immunol.* [Internet] 9:309–318. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1050464898901814

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, et al. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* [Internet] 81:559–575. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002929707613524

Qiu W, Zhu Y, Wu Y, Yuan C, Chen K, Li M. 2018. Identification and expression analysis of microRNAs in medaka gonads. *Gene* [Internet] 646:210–216. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0378111917311319

Qu M, Liu Y, Zhang Y, Wan S, Ravi V, Qin G, Jiang H, Wang X, Zhang Huixian, Zhang B, et al. 2021. Seadragon genome analysis provides insights into its phenotype and sex determination locus. *Sci. Adv.* [Internet] 7:eabg5196. Available from: https://www.science.org/doi/10.1126/sciadv.abg5196

Ramanagoudr-Bhojappa R, Carrington B, Ramaswami M, Bishop K, Robbins GM, Jones M, Harper U, Frederickson SC, Kimble DC, Sood R, et al. 2018. Multiplexed CRISPR/Cas9-mediated knockout of 19 Fanconi anemia pathway genes in zebrafish revealed their roles in growth, sexual development and fertility.Barsh GS, editor. *PLOS Genet.* [Internet] 14:e1007821. Available from: https://dx.plos.org/10.1371/journal.pgen.1007821

Rankin DJ, Dieckmann U, Kokko H. 2011. Sexual Conflict and the Tragedy of the Commons. *Am. Nat.* [Internet] 177:780–791. Available from: https://www.journals.uchicago.edu/doi/10.1086/659947

Raz E. 2000. The function and regulation of vasa-like genes in germ-cell development. *Genome Biol.* [Internet] 1:reviews1017.1. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/gb-2000-1-3-reviews1017

Reynolds A, McDearmid JR, Lachance S, Marco PD, Merello E, Capra V, Gros P, Drapeau P, Kibar Z. 2010. VANGL1 rare variants associated with neural tube defects affect convergent extension in zebrafish. *Mech. Dev.* [Internet] 127:385–392. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0925477309015007

Rice WR, Holland B. 1997. The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. *Behav. Ecol. Sociobiol.* [Internet] 41:1–10. Available from: http://link.springer.com/10.1007/s002650050357

Roth O, Scharsack JP, Keller I, Reusch TBH. 2011. Bateman's principle and immunity in a sex-role reversed pipefish: Sexual immune dimorphism in pipefish. *J. Evol. Biol.* [Internet] 24:1410–1420. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1420-9101.2011.02273.x

Roth O, Solbakken MH, Tørresen OK, Bayer T, Matschiner M, Baalsrud HT, Hoff SNK, Brieuc MSO, Haase D, Hanel R, et al. 2020. Evolution of male pregnancy associated with remodeling of canonical vertebrate immunity in seahorses and pipefishes. *Proc. Natl. Acad. Sci.* [Internet] 117:9431–9439. Available from: https://pnas.org/doi/full/10.1073/pnas.1916251117

Rowe L, Chenoweth SF, Agrawal AF. 2018. The Genomics of Sexual Conflict. *Am. Nat.* [Internet] 192:274–286. Available from: https://www.journals.uchicago.edu/doi/10.1086/698198

Sacomoto GA, Kielbassa J, Chikhi R, Uricaru R, Antoniou P, Sagot M-F, Peterlongo P, Lacroix V. 2012. KIS SPLICE: de-novo calling alternative splicing events from RNA-seq data. *BMC Bioinformatics* [Internet] 13:S5. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-13-S6-S5

Schartl M, Georges A, Marshall Graves JA. 2023. Polygenic sex determination in vertebrates – is there any such thing? *Trends Genet.* [Internet] 39:242–250. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168952522003110

Schwager EE, Meng Y, Extavour CG. 2015. vasa and piwi are required for mitotic integrity in early embryogenesis in the spider Parasteatoda tepidariorum. *Dev. Biol.* [Internet] 402:276–290. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0012160614004382

Shen C, Xiong W-C, Mei L. 2015. LRP4 in neuromuscular junction and bone development and diseases. *Bone* [Internet] 80:101–108. Available from: https://linkinghub.elsevier.com/retrieve/pii/S8756328215001830

Shen Z, Sathyan KM, Geng Y, Zheng R, Chakraborty A, Freeman B, Wang F, Prasanth KV, Prasanth SG. 2010. A WD-Repeat Protein Stabilizes ORC Binding to Chromatin. *Mol. Cell*

[Internet] 40:99–111. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1097276510007495

Sheu Y-J, Kinney JB, Lengronne A, Pasero P, Stillman B. 2014. Domain within the helicase subunit Mcm4 integrates multiple kinase signals to control DNA replication initiation and fork progression. *Proc. Natl. Acad. Sci.* [Internet] 111. Available from: https://pnas.org/doi/full/10.1073/pnas.1404063111

Shi J, Liang C. 2019. Generic Repeat Finder: A High-Sensitivity Tool for Genome-Wide De Novo Repeat Detection. *Plant Physiol.* [Internet] 180:1803–1815. Available from: https://academic.oup.com/plphys/article/180/4/1803-1815/6117697

Silva MSB, Giacobini P. 2021. New insights into anti-Müllerian hormone role in the hypothalamic–pituitary–gonadal axis and neuroendocrine development. *Cell. Mol. Life Sci.* [Internet] 78:1–16. Available from: https://link.springer.com/10.1007/s00018-020-03576-x

Simhadri S, Peterson S, Patel DS, Huo Y, Cai H, Bowman-Colin C, Miller S, Ludwig T, Ganesan S, Bhaumik M, et al. 2014. Male Fertility Defect Associated with Disrupted BRCA1-PALB2 Interaction in Mice. *J. Biol. Chem.* [Internet] 289:24617–24629. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0021925820319931

Smith SH, Hsiung K, Böhne A. 2023. Evaluating the role of sexual antagonism in the evolution of sex chromosomes: new data from fish. *Curr. Opin. Genet. Dev.* [Internet] 81:102078. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0959437X23000588

Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S. 2016. TransRate: reference-free quality assessment of de novo transcriptome assemblies. *Genome Res.* [Internet] 26:1134–1144. Available from: http://genome.cshlp.org/lookup/doi/10.1101/gr.196469.115

Soneson C, Love MI, Robinson MD. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* [Internet] 4:1521. Available from: https://f1000research.com/articles/4-1521/v1

Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics* [Internet] 24:637–644. Available from: https://academic.oup.com/bioinformatics/article/24/5/637/202844

Stanke M, Schöffmann O, Morgenstern B, Waack S. 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* [Internet] 7:62. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-7-62

Stearns FW. 2010. One Hundred Years of Pleiotropy: A Retrospective. *Genetics* [Internet] 186:767–773. Available from: https://academic.oup.com/genetics/article/186/3/767/6063631

Su W, Gu X, Peterson T. 2019. TIR-Learner, a New Ensemble Method for TIR Transposable Element Annotation, Provides Evidence for Abundant New Transposable Elements in the Maize Genome. *Mol. Plant* [Internet] 12:447–460. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1674205219300620

Suszynska M, Ratajska M, Kozlowski P. 2020. BRIP1, RAD51C, and RAD51D mutations are associated with high susceptibility to ovarian cancer: mutation prevalence and precise risk estimates based on a pooled analysis of ~30,000 cases. *J. Ovarian Res.* [Internet] 13:50. Available from: https://ovarianresearch.biomedcentral.com/articles/10.1186/s13048-020-00654-3

Svensson EI. 2018. On Reciprocal Causation in the Evolutionary Process. *Evol. Biol.* [Internet] 45:1–14. Available from: http://link.springer.com/10.1007/s11692-017-9431-x

Svensson EI. 2019. Eco□evolutionary dynamics of sexual selection and sexual conflict.Deere J, editor. *Funct. Ecol.* [Internet] 33:60–72. Available from: https://onlinelibrary.wiley.com/doi/10.1111/1365-2435.13245

Svensson Erik I. 2023. Phenotypic selection in natural populations: what have we learned in 40 years? *Evolution* [Internet] 77:1493–1504. Available from: https://academic.oup.com/evolut/article/77/7/1493/7146186

Svensson Erik I. 2023. Why high-condition males might be bad news for females and populations. *Proc. Natl. Acad. Sci.* [Internet] 120:e2301989120. Available from: https://pnas.org/doi/10.1073/pnas.2301989120

Takahashi Y, Kagawa K, Svensson EI, Kawata M. 2014. Evolution of increased phenotypic diversity enhances population performance by reducing sexual harassment in damselflies. *Nat. Commun.* [Internet] 5:4468. Available from: https://www.nature.com/articles/ncomms5468

Tao W, Sun L, Shi H, Cheng Y, Jiang D, Fu B, Conte MA, Gammerdinger WJ, Kocher TD, Wang D. 2016. Integrated analysis of miRNA and mRNA expression profiles in tilapia gonads at an early stage of sex differentiation. *BMC Genomics* [Internet] 17:328. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-016-2636-z

Tizaoui K, Terrazzino S, Cargnin S, Lee KH, Gauckler P, Li H, Shin JI, Kronbichler A. 2021. The role of PTPN22 in the pathogenesis of autoimmune diseases: A comprehensive review. *Semin. Arthritis Rheum.* [Internet] 51:513–522. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0049017221000342

Torban E, Patenaude A-M, Leclerc S, Rakowiecki S, Gauthier S, Andelfinger G, Epstein DJ, Gros P. 2008. Genetic interaction between members of the Vangl family causes neural tube defects in mice. *Proc. Natl. Acad. Sci.* [Internet] 105:3449–3454. Available from: https://pnas.org/doi/full/10.1073/pnas.0712126105

Turner JMA, Aprelikova O, Xu X, Wang R, Kim S, Chandramouli GVR, Barrett JC, Burgoyne PS, Deng C-X. 2004. BRCA1, Histone H2AX Phosphorylation, and Male Meiotic Sex Chromosome Inactivation. *Curr. Biol.* [Internet] 14:2135–2142. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0960982204009194

Varì R, Scazzocchio B, Filardi T, Citarella A, Bellenghi M, Masella R, Santangelo C. 2021. Significance of Sex Differences in ncRNAs Expression and Function in Pregnancy and Related Complications. *Biomedicines* [Internet] 9:1509. Available from: https://www.mdpi.com/2227-9059/9/11/1509

Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, et al. 2010. Quantitative Interaction Proteomics and Genome-wide

Profiling of Epigenetic Histone Marks and Their Readers. *Cell* [Internet] 142:967–980. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0092867410009517

Voichek Y, Weigel D. 2020. Identifying genetic variants underlying phenotypic variation in plants without complete genomes. *Nat. Genet.* [Internet] 52:534–540. Available from: https://www.nature.com/articles/s41588-020-0612-7

Wagner S, Whiteley SL, Castelli M, Patel HR, Deveson IW, Blackburn J, Holleley CE, Marshall Graves JA, Georges A. 2023. Gene expression of male pathway genes sox9 and amh during early sex differentiation in a reptile departs from the classical amniote model. *BMC Genomics* [Internet] 24:243. Available from: https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-023-09334-0

Walker LJ, Roque RA, Navarro MF, Granato M. 2021. Agrin/Lrp4 signal constrains MuSK-dependent neuromuscular synapse development in appendicular muscle. *Development* [Internet] 148:dev199790. Available from: https://journals.biologists.com/dev/article/148/21/dev199790/272655/Agrin-Lrp4-signal-constrains-MuSK-dependent

Walton OC, Stevens M. 2018. Avian vision models and field experiments determine the survival value of peppered moth camouflage. *Commun. Biol.* [Internet] 1:118. Available from: https://www.nature.com/articles/s42003-018-0126-3

Wang F, Jia Y, Wang P, Yang Q, Du Q, Chang Z. 2017. Identification and profiling of Cyprinus carpio microRNAs during ovary differentiation by deep sequencing. *BMC Genomics* [Internet] 18:333. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-017-3701-y

Wang L, Sun F, Wan ZY, Yang Z, Tay YX, Lee M, Ye B, Wen Y, Meng Z, Fan B, et al. 2022. Transposon-induced epigenetic silencing in the X chromosome as a novel form of dmrt1 expression regulation during sex determination in the fighting fish. *BMC Biol.* [Internet] 20:5. Available from: https://bmcbiol.biomedcentral.com/articles/10.1186/s12915-021-01205-y

Whittington CM, Griffith OW, Qi W, Thompson MB, Wilson AB. 2015. Seahorse Brood Pouch Transcriptome Reveals Common Genes Associated with Vertebrate Pregnancy. *Mol. Biol. Evol.*

[Internet]:msv177. Available from: https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msv177

Wilson CA, High SK, McCluskey BM, Amores A, Yan Y, Titus TA, Anderson JL, Batzel P, Carvan MJ, Schartl M, et al. 2014. Wild Sex in Zebrafish: Loss of the Natural Sex Determinant in Domesticated Strains. *Genetics* [Internet] 198:1291–1308. Available from: https://academic.oup.com/genetics/article/198/3/1291/6065698

Wright AE, Fumagalli M, Cooney CR, Bloch NI, Vieira FG, Buechel SD, Kolm N, Mank JE. 2018. Male-biased gene expression resolves sexual conflict through the evolution of sex-specific genetic architecture. *Evol. Lett.* [Internet] 2:52–61. Available from: https://academic.oup.com/evlett/article/2/2/52/6697426

Wu S, Zhou J, Zhang K, Chen H, Luo M, Lu Y, Sun Y, Chen Y. 2020. Molecular Mechanisms of PALB2 Function and Its Role in Breast Cancer Management. *Front. Oncol.* [Internet] 10:301. Available from: https://www.frontiersin.org/article/10.3389/fonc.2020.00301/full

Wu TD, Watanabe CK. 2005. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* [Internet] 21:1859–1875. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bti310

Xiao J, Zhong H, Zhou Y, Yu F, Gao Y, Luo Y, Tang Z, Guo Z, Guo E, Gan X, et al. 2014. Identification and Characterization of MicroRNAs in Ovary and Testis of Nile Tilapia (Oreochromis niloticus) by Using Solexa Sequencing Technology.Guan X-Y, editor. *PLoS ONE* [Internet] 9:e86821. Available from: https://dx.plos.org/10.1371/journal.pone.0086821

Xiong W, He L, Lai J, Dooner HK, Du C. 2014. HelitronScanner uncovers a large overlooked cache of *Helitron* transposons in many plant genomes. *Proc. Natl. Acad. Sci.* [Internet] 111:10263–10268. Available from: https://pnas.org/doi/full/10.1073/pnas.1410068111

Xu Z, Wang H. 2007. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* [Internet] 35:W265–W268. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkm286

Yajima M, Wessel GM. 2011. The DEAD-box RNA helicase Vasa functions in embryonic mitotic progression in the sea urchin. *Development* [Internet] 138:2217–2222. Available from: https://journals.biologists.com/dev/article/138/11/2217/44377/The-DEAD-box-RNA-helicase-Vasa-functions-in

Yajima M, Wessel GM. 2015. Essential elements for translation: the germline factor Vasa functions broadly in somatic cells. *Development* [Internet] 142:1960–1970. Available from: https://journals.biologists.com/dev/article/142/11/1960/46767/Essential-elements-for-translation-the-germline

Yamagata A, Yoshida T, Sato Y, Goto-Ito S, Uemura T, Maeda A, Shiroshima T, Iwasawa-Okamoto S, Mori H, Mishina M, et al. 2015. Mechanisms of splicing-dependent trans-synaptic adhesion by PTPδ–IL1RAPL1/IL-1RAcP for synaptic differentiation. *Nat. Commun.* [Internet] 6:6926. Available from: https://www.nature.com/articles/ncomms7926

Yan H, Liu Q, Jiang J, Shen X, Zhang L, Yuan Z, Wu Y, Liu Y. 2021. Identification of sex differentiation-related microRNA and long non-coding RNA in Takifugu rubripes gonads. *Sci. Rep.* [Internet] 11:7459. Available from: https://www.nature.com/articles/s41598-021-83891-w

Yang W, Garrett L, Feng D, Elliott G, Liu X, Wang N, Wong YM, Choi NT, Yang Y, Gao B. 2017. Wnt-induced Vangl2 phosphorylation is dose-dependently required for planar cell polarity in mammalian development. *Cell Res.* [Internet] 27:1466–1484. Available from: https://www.nature.com/articles/cr2017127

Yasumura M, Yoshida T, Yamazaki M, Abe M, Natsume R, Kanno K, Uemura T, Takao K, Sakimura K, Kikusui T, et al. 2014. IL1RAPL1 knockout mice show spine density decrease, learning deficiency, hyperactivity and reduced anxiety-like behaviours. *Sci. Rep.* [Internet] 4:6613. Available from: https://www.nature.com/articles/srep06613

Yatsu R, Miyagawa S, Kohno S, Parrott BB, Yamaguchi K, Ogino Y, Miyakawa H, Lowers RH, Shigenobu S, Guillette LJ, et al. 2016. RNA-seq analysis of the gonadal transcriptome during Alligator mississippiensis temperature-dependent sex determination and differentiation. *BMC Genomics* [Internet] 17:77. Available from: https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-016-2396-9

Yoshida T, Mishina M. 2008. Zebrafish orthologue of mental retardation protein IL1RAPL1 regulates presynaptic differentiation. *Mol. Cell. Neurosci.* [Internet] 39:218–228. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1044743108001656

Zhang F, Ma J, Wu J, Ye L, Cai H, Xia B, Yu X. 2009. PALB2 Links BRCA1 and BRCA2 in the DNA-Damage Response. *Curr. Biol.* [Internet] 19:524–529. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0960982209007234

Zhang H, Hou J, Liu H, Zhu H, Xu G, Xu J. 2020. Adaptive evolution of low-salinity tolerance and hypoosmotic regulation in a euryhaline teleost, Takifugu obscurus. *Mar. Biol.* [Internet] 167:90. Available from: https://link.springer.com/10.1007/s00227-020-03705-x

Zhang N, Liang J, Tian Y, Yuan L, Wu L, Miao S, Zong S, Wang L. 2010. A Novel Testis-specific GTPase Serves as a Link to Proteasome Biogenesis: Functional Characterization of RhoS/RSA-14-44 in Spermatogenesis.Kaibuchi K, editor. *Mol. Biol. Cell* [Internet] 21:4312–4324. Available from: https://www.molbiolcell.org/doi/10.1091/mbc.e10-04-0310

Zhang R-G, Li G-Y, Wang X-L, Dainat J, Wang Z-X, Ou S, Ma Y. 2022. TEsorter: An accurate and fast method to classify LTR-retrotransposons in plant genomes. *Hortic. Res.* [Internet] 9:uhac017. Available from: https://academic.oup.com/hr/article/doi/10.1093/hr/uhac017/6532211

Zhu H, Wen Z, Zhang A, Liu D, Wang H, Cheng Y, Yang X, Xiao Y, Li J, Sun D, et al. 2023. RhoGDIα regulates spermatogenesis through Rac1/cofilin/F-actin signaling. *Commun. Biol.* [Internet] 6:214. Available from: https://www.nature.com/articles/s42003-023-04579-7

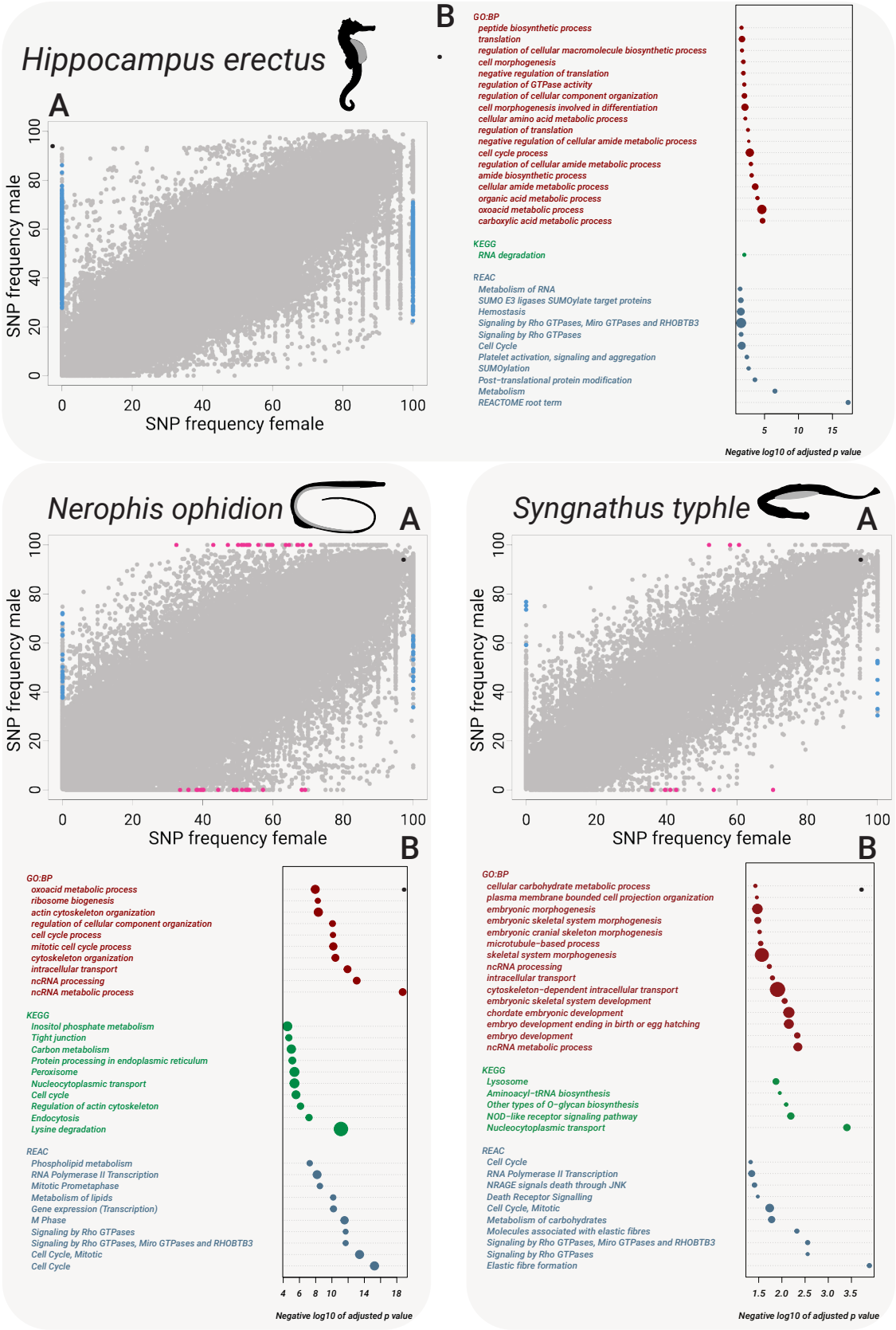

Figure 1. Panels - *Hippocampus erectus*, *Nerophis ophidion*, *Syngnathus typhle*. **Subpanel A** - male vs female allele frequencies of Kissplice-called SNPs. Coloured grey are the variants that show neither sex-specificity, nor differential expression between sexes. Marked with blue are male-specific SNPs, marked with pink - female specific SNPs. **Subpanel B** - results of an enrichment analysis for 3 data sources: GO BP (red), KEGG (green) and Reactome (blue). Each point represents an enriched term, with name of the term located on the Y-axis, and padj values on the X-axis. The size of the point represents a relationship between the number of genes from the list that belong to this term vs the term size. Meaning, point size represents a proportion of the term is that is represented but the overlap genes.

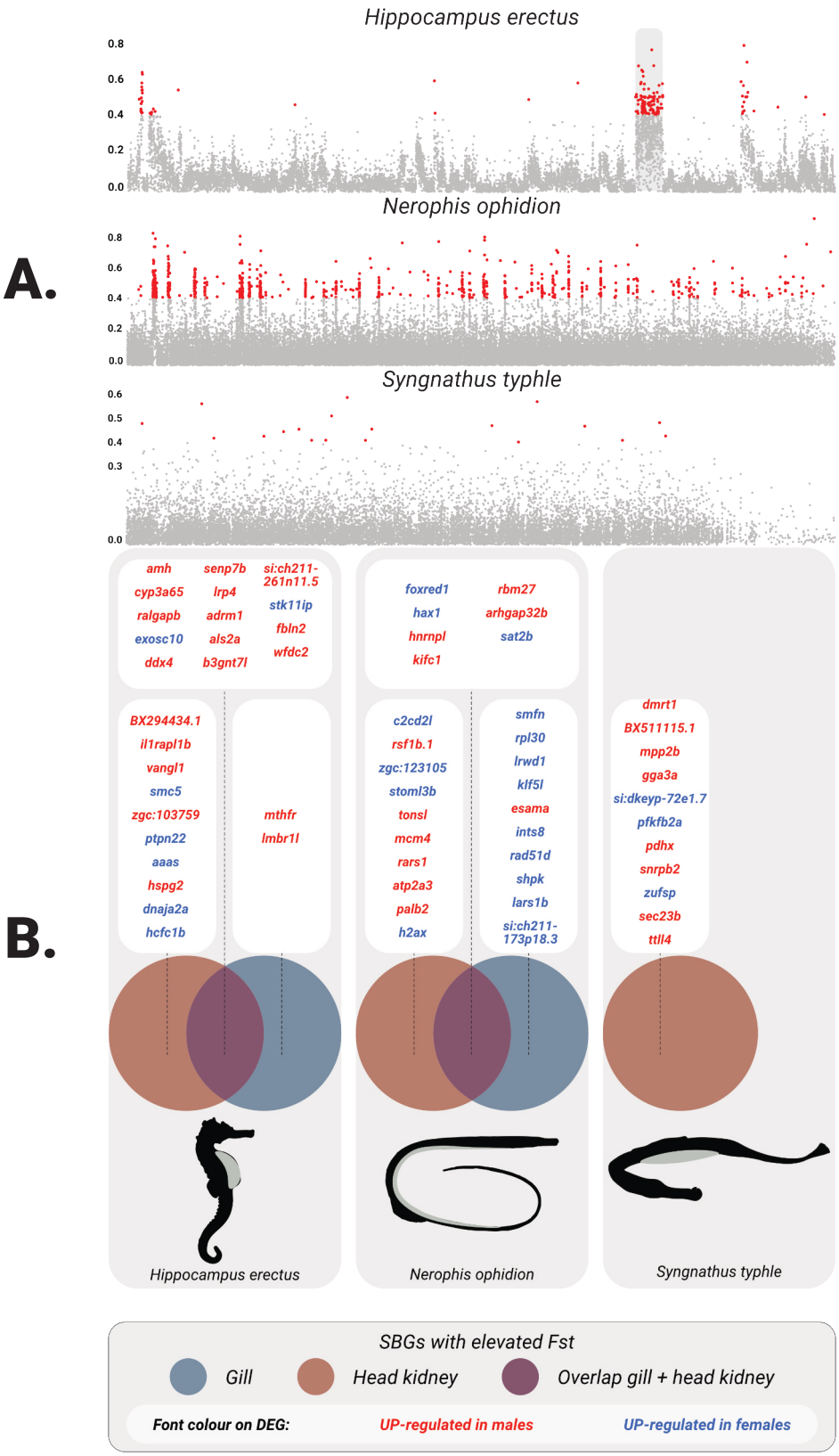

Figure 2. **Panel A** - Fst values per 10 kb windows across the genome in *H.erectus*, *N.ophidion*, *S.typhle*. Y-axis - Fst value, X-axis -position within a genome. Colored points - ones that pass the 0.4 threshold. On HE plot, grey rounded rectangle marks candidate sex chromosome. **Panel B** - the overlap between genes with elevated Fst and DEG in *H.erectus*, *N.ophidion*, *S.typhle*. Blue colored circles represent gill DEGs, red colored circles represent head kidney DEGs, purple - the overlap between lists of genes.Each panel with gene lists is connected to the appropriate overlap sector with a dotted line. The lists are ordered based on *padj* value and colored based on sex (red - up-regulation in males, blue - up-regulation in females).

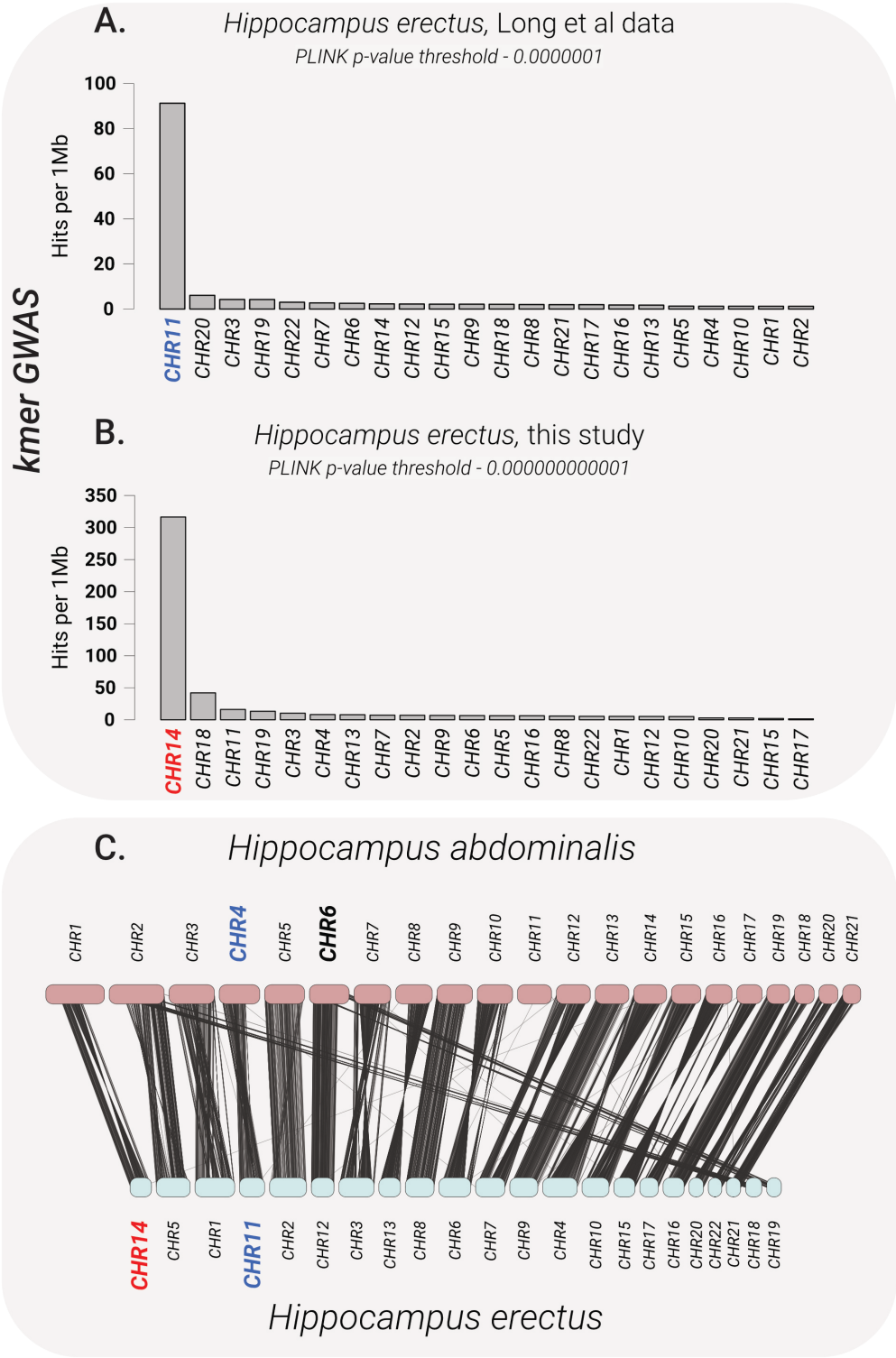

Figure 3. **Panels A and B** - results of kmerGWAS analysis. Y-axis - number of ABYSS contigs that aligned to the genome with BLASTn per 1Mb (i.e. adjusted for size), X-axis - the chromosome. **Panel C** - synteny analysis between *H.abdominalis* and *H.erectus* genome. Each line connects 1-1 orthologues in HA and HE. HA chromosomes were ordered based on their length, HE chromosomes were ordered by their relation to HA genome. Red marks sex-linked chr identified in our study. Blue - HE sex-linke chr identified by Long, Bold Black - *H.abdominalis* sex-linked chr.