Independent evolution of sex chromosomes and male pregnancy-related genes in two seahorse species

Xin Long[1,9], Deborah Charlesworth[2], Jianfei Qi[3], Ruiqiong Wu[4], Meiling Chen[4], Zongji Wang[1], Luohao Xu[5], Honggao Fu[4], Xueping Zhang[4], Xinxin Chen[3], Libin He[3], Leyun Zheng[3,*], Zhen Huang[4,6,*], Qi Zhou[1,5,7,8,*]

1. MOE Laboratory of Biosystems Homeostasis and Protection and Zhejiang Provincial Key Laboratory for Cancer Molecular Cell Biology, Life Sciences Institute, Zhejiang University, Hangzhou, 310058, China.
2. Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, West Mains Road, EH9 3LF, UK
3. Fisheries Research Institute of Fujian, Xiamen, 361013, China.
4. Fujian Key Laboratory of Developmental and Neural Biology & Southern Center for Biomedical Research, College of Life Sciences, Fujian Normal University, Fuzhou, Fujian, China.
5. Department of Neuroscience and Developmental Biology, University of Vienna, Vienna, 1090, Austria.
6. Fujian-Macao Science and Technology Cooperation Base of Traditional Chinese Medicine-Oriented Chronic Disease Prevention and Treatment, Innovation and Transformation Center, Fujian University of Traditional Chinese Medicine, Fuzhou, 350108,China.
7. Center for Reproductive Medicine, The 2nd Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310052, China
8. Evolutionary & Organismal Biology Research Center, School of Medicine, Zhejiang University, Hangzhou, 310058, China
9. Research Center for Intelligent Computing Platforms, Zhejiang Lab, Hangzhou, 311100, China

* Correspondence author: Leyun Zheng: 981668817@qq.com or Zhen Huang: zhuang@fjnu.edu.cn or Qi Zhou: zhouqi1982@zju.edu.cn

Abstract

Unlike birds and mammals, many teleosts have homomorphic sex chromosomes and changes in the chromosome carrying the sex-determining locus, termed “turnovers”, are common. Recent turnovers allow studies of several interesting questions. One question is whether the new sex-determining regions evolve to become completely non-recombining, and if so, how and why. Another is to test the prediction that evolutionary changes that benefit one sex will accumulate in the newly sex-linked region. To study these questions, we analyzed the genome sequences of two seahorse species of the Syngnathidae, a fish group in which many species evolved a unique structure, the male brood pouch. We find that both seahorse species have XY sex chromosome systems, and their sex chromosome pairs are not homologs, implying that at least one turnover event has occurred. The Y-linked regions respectively occupy 63.9% and 95.1% of the entire chromosome of the two species, and do not exhibit extensive sequence divergence with their X-linked homologs. We find evidence for occasional recombination between the extant sex chromosomes that may account for their homomorphic pattern. We argue that these Y-linked regions did not evolve by recombination suppression after the turnover. Instead, it can be explained by the ancestral nature of low crossover rate at the corresponding chromosome location. With such an ancestral crossover landscape, a turnover can instantly create an extensive Y-linked region. Finally, we investigate the adaptive evolution of male pouch related genes after they become Y-linked in the seahorse.

Introduction

Studies of young sex-linked regions are important for understanding the processes involved in the initial evolution of sex chromosomes, including how their lack of recombination arose, and the time-course of subsequent adaptive and degenerative changes (reviewed by Bachtrog, et al. 2014). The discovery that sex-determining regions (SDR) have evolved independently in different fish species (reviewed by (Pan, et al. 2016)) offers opportunities for such studies. A major hypothesis to explain the lack of recombination between sex chromosomes is the sexually antagonistic (SA) polymorphism hypothesis, involving selection generated by conflicts or fitness trade-offs between alleles' effects in the two sexes (Rice 1987). Such selection can establish a two-locus polymorphism with alleles at the locus under SA selection associated with alleles at the sex-determining locus, which generates selection for closer linkage between the two genes (Bull 1983). This process of expanding the non-recombining region can be repeated, causing successive recombination suppression events, at different evolutionary times, each with a distinctive level of sequence divergence between Y-X or W-Z sex-linked gene pairs. Such so-called "evolutionary strata" have been detected in mammalian XY chromosome pairs (Lahn and Page 1999; Bellott, et al. 2014; Cortez, et al. 2014). It was proposed that these recombination suppression events are caused by inversions (Lahn and Page 1999), and and there is evidence supporting this for at least one human stratum (Lemaitre, et al. 2009), and a stratum in payaya (Wang, et al. 2012), though not in the threespine stickleback (Peichel, et al. 2020). Recombination suppression does not require an inversion, and can also evolve via recombination modifiers that control crossovers in specific genome locations. Other species with female heterogamety, including birds, also appear to have evolutionary strata (Zhou, et al. 2014).

In contrast to the ancient mammalian and bird sex chromosomes, non-homology of the sex chromosomes of closely related species has been documented in several teleosts, amphibians and plants, indicating that changes, or "turnover events" have occurred (reviewed by (Vicoso 2019), sometimes even between different populations of the same species (e.g., frog (Miura, et al. 2012)). Turnovers might also involve SA polymorphism, as such a polymorphism can favor the appearance of a new sex-determining gene in its closely linked genomic region, or can become established near a sex-determining gene that has appeared in a new genome region for some other reason (van Doorn and Kirkpatrick 2007).

Turnover events can offer opportunities to test the SA polymorphism hypothesis. This hypothesis is extremely difficult to test (see, for example(Olito, et al. 2022)), because documenting that a balanced polymorphism that is being maintained is challenging; and the additional requirement for evidence that SA selection is involved makes the task even harder, though population genomic approaches are starting to be developed (Qiu, et al. 2013; Dagilis, et al. 2022).

While old-established turnovers can be investigated to test whether new completely sex-linked strata have subsequently evolved, recent turnover events can reveal alternative origins of large sex-linked genome regions. Because SA polymorphisms are expected to be maintained mainly at loci very closely linked to the sex-determining locus (Jordan and Charlesworth 2012), they are unlikely to lead to evolution of large sex-linked genome regions unless a large inversion happens to suppress recombination across the region that includes the SA and sex-determining loci, and becomes fixed in the population; invoking recurrent such events in multiple organisms seems implausible (Olito, et al. 2022). Therefore, if an extensive sex-linked region is found after a recent turnover, alternative possibilities should be considered.

A complementary approach is to test whether the genome region linked to a new male-determining locus starts accumulating SA mutations. Under male heterogamety, for example, mutations can spread if they benefit aspects of male fitness enough to outweigh any deleterious effects in females. Even if the restrictive conditions are not satisfied for such mutations to establish polymorphisms (Jordan and Charlesworth 2012) and create selection for linkage with the male-determining gene, their accumulation or fixation in the population (replacing both the ancestral Y- and X-linked alleles) can potentially be detected without needing to detect balancing selection. The Syngnathidae are suitable for such studies because, uniquely among animals, they have evolved the male brood pouch, in which males fertilize the eggs and nurture the offspring until hatching (Whittington, et al. 2015).

We here study several species of Syngnathidae, a teleost group that includes seahorses, seadragons and pipefish (Stiller, et al. 2022). The syngnathids' male pregnancy reproductive behavior is unique among animals, and its evolution appears to have involved modifications of the immune gene repertoire, including loss or rapid sequence evolution of several major histocompatibility complex (MHC) genes, which may reflect trade-offs between immunological tolerance and embryo rejection (Roth, et al. 2020). There is evidence that male Gulf pipefish actively abort embryos of less attractive females (Paczolt and Jones 2010). Therefore this group of teleosts has been frequently used to study the evolution of novel morphological traits, parent-offspring conflicts, sexual conflicts and sexual selection (Vincent, et al. 1992; Lin, et al. 2016; Li, et al. 2021).

Our study focuses on the genomes of two seahorse species, the big-belly seahorse, *Hippocampus abdominalis* and the lined seahorse, *H. erectus*, and two outgroup species within the Syngnathidae, the Gulf pipefish, *Syngnathus scovelli*, and common seadragon, *Phyllopteryx taeniolatus* (**Figure 1A**). We recently generated a high-quality chromosome level genome assembly of a male big-belly seahorse (He, et al. 2021). Using this as a reference, we newly produced and analysed genomic and transcriptomic data from both sexes of this species, and those from its related species, the lined seahorse. Here we now (i) identify the sex-determining regions of the two seahorse

species, and show that at least one of them specifically evolved within the seahorse lineage, (ii) describe evidence that both seahorse species have large Y-linked regions, and investigate how these regions' reduced recombination evolved, and (iii) test whether these Y-linked regions have undergone genetic degeneration and/or adaptive changes, including whether SA selection is suggested. This test is based on the prediction that seahorse Y-linked regions should become enriched with male-specific pouch-expressed genes. Acquisition of such genes may involve movements from other chromosomes, probably requiring a long evolutionary time. If, however, a Y-linked region carrying many genes evolved recently, as the case of the sex chromosome of big-belly seahorse shown here, we might expect to detect an earlier stage of adaptation, in which pouch-expressed genes that are located in the region start evolving higher expression.

Results

Chromosome evolution of seahorse and Gulf pipefish species

We first describe our new results establishing that both seahorse species studied have extensive sex-linked regions that include many genes. We also evaluate the extent of sex-linkage and the evidence that it evolved recently in at least one of the seahorse lineages represented by our study species. Little was previously known about seahorse sex chromosomes, though a cytogenetic study suggested that the lined seahorse (*Hippocampus erectus*) might have a ZW sex chromosome pair (Liu, et al. 2020), which our results below do not support.

Cytogenetic studies reported diploid chromosome numbers ranging from 2n=44 (in *H. erectus* and the Gulf pipefish, *S. scovelli*) to 2n=58 among syngnathids, with most seahorse chromosomes being acrocentric, while those of Gulf pipefish are (sub)metacentric (Vitturi, et al. 1998; Liu, et al. 2020). The teleost ancestor was inferred to have 24 pairs of acrocentric chromosomes (Ohno 1974; Vitturi, et al. 1995; Mank and Avise 2006; Yoshida and Kitano 2021) suggesting two chromosome fusion or translocation events in the Syngnathidae lineage, as already suggested for the Gulf pipefish (Small, et al. 2016). In *H. abdominalis*, over 99% of the haploid genome is assembled into 21 chromosomes (He, et al. 2021), indicating one additional fusion or translocation event specific to *H. abdominalis* (**Supplementary Fig. S1**). Most *H. abdominalis* chromosomes each show one-to-one sequence homology with a single *S. scovelli* or Nile tilapia chromosome, but distinct parts of Chr11 are homologous to separate *S. scovelli* chromosomes, LG15 and LG17 (**Figure 1B**) and Nile tilapia (**Figure 1C**). This suggests a chromosome fusion after the split from *H. erectus*, which is estimated to have occurred 11.4 MY ago (He, et al. 2021).

Sequences of the *H. abdominalis* Chr6 (which, as shown below, is this species' sex chromosome pair) can be aligned to two chromosomes of the distantly related non-teleost fish, *Lepisosteus oculatus* (the spotted gar, more details below in **Fig. 3D**). This

rearrangement is additional to the two chromosome fusion or translocation events within the Syngnathidae lineage (see above). It probably occurred before the teleost whole genome duplication and species radiation (Bian, et al. 2016), since homologs of these two gar chromosomes are also fused in many other teleosts, including zebrafish (Howe, et al. 2013), stickleback (Braasch, et al. 2016), arowana (Bian, et al. 2016), medaka and the guppy, in which the fused chromosome is the one named LG16 (**Supplementary Fig. S2**).

Compared with *S. scovelli*, the *H. abdominalis* genome is 44.8% larger, and all three seahorse species analysed have higher repeat contents (1.38-fold for *H. erectus,* 2.07-fold for *H. abdominalis*, and 1.40-fold for *H. comes*, the tiger-tail seahorse, see **Supplementary Fig. S3**). This indicates *Hippocampus*-specific expansions of several transposable element (TE) families (**Supplementary Table S2**), especially the DNA transposon TcMar-Tigger, Long Terminal Repeat (LTR) Ngaro subfamilies (**Figure 1A**), as well as two subfamilies of short interspersed nuclear element (SINE) tRNA-Core-L2 and DNA transposon Zator, which are both concentrated at the chromosome fusion junction in the middle of Chr11 of *H. abdominalis* (**Figure 1D**). About 55.2% and 15.9% of the entire genome's tRNA-Core-L2 and Zator sequences, respectively, are within this junction region (**Supplementary Table S3**). These two repeats are more abundant in the *H. abdominals* genome than in *H. erectus* (1.48 and 2.12 times higher, respectively; **Supplementary Table S4**), suggesting that they have recently expanded in *H. abdominals*. As we did not detect any telomeric repeats at the junction site, the tRNA-Core-L2 and Zator repeats might mark *H. abdominalis* centromeric regions.

The same two transposon sequences were indeed enriched at one end of 18 of the other 20 *H. abdominalis* chromosomes (**Supplementary Table S5**). Centromeric repeats commonly differ between species (Melters, et al. 2013), and the sequences identified here did not show homology with centromeric repeats reported in other teleosts. By examining regions homologous to the fusion junction site of *H. abdominalis* Chr11, we also identified putative centromeric repeats in *H. erectus* and *S. scovelli*. They are also enriched at the ends of the two chromosomes corresponding to linkage groups LG15 and LG17 of *S. scovelli* (**Supplementary Table S6**, **Figure 1D**), supporting the centric fusion hypothesis outlined above for *H. abdominals* Chr11.

Chromosome-wide recombination patterns in seahorses

To initially assess recombination landscapes before searching for sex-linked regions in our seahorse genomes, we describe features of the chromosome sequences of *H. abdominalis*, *H. erectus* (Lin, et al. 2017) and *S. scovelli* (Small, et al. 2016) that relate to recombination rates. First, the chromosomes exhibit spikes of repeat content at the inferred centromeric ends in all three species (Vitturi, et al. 1998; Liu, et al. 2020). The chromosomes also display pronounced patterns in GC content: based on the

centromeric positions inferred above, the other ends of the chromosomes (which we term “tip regions”) show spikes of GC content in both seahorse genome assemblies (**Figure 1E**, $P < 0.001$, by Mann–Whitney U tests), resembling findings in many other teleosts (Costantini, et al. 2007; Matoulek, et al. 2020; Cheng, et al. 2021). The tip regions with GC spikes identified by change-point analyses of GC content within each chromosome (see **Materials and Methods** and **Supplementary Fig. S4**) occupy 4.93% of the entire genome. This probably reflects GC-biased gene conversion (gGBC) associated with highly localized crossovers (Arbeithuber, et al. 2015) at one chromosome end in at least one sex, as suggested for the guppy, in which crossover rates in males correlate with intronic GC values (Charlesworth, et al. 2020). Correspondingly, GC contents of pericentromeric regions of each seahorse chromosome (defined as 3Mb regions flanking the putative centromeric marks in **Figure 1D**) are significantly lower than in the rest of the respective chromosomes ($P < 2.2e-16$, Wilcoxon test, **Supplementary Fig. S5**). These regions account for at least 18% of the *H. abdominalis* genome. Chromosome-wide GC patterns are much less clear in *S. scovelli*. Based on the centromeric positions inferred above in *S. scovelli*, (**Supplementary Fig. S6, Supplementary Table S5**) 8 of its 22 chromosomes are inferred to be metacentric, submetacentric or subtelocentric, and intrachromosomal rearrangements may have changed the overall GC patterns.

No dense genetic maps are yet available for seahorses. However, sex-averaged recombination rates (rho values), estimated by analyzing linkage disequilibrium across *H. abdominalis* chromosomes (using 30 males and 29 females from a cultured population, see **Materials and Methods**), show the expected tendency to be highest in regions with high GC content (Pearson correlation $r = 0.35$, $P < 2.2e-16$; **Supplementary Fig. S7**), consistent with high meiotic crossover rates in those regions, and low rates in other regions in one or both sexes.

Although direct evidence from sex-specific genetic mapping data is not yet available in the seahorse species studied here, independent evidence supports the conclusion that strong crossover localization to small chromosome end regions occurs specifically in males. Genetic mapping of microsatellite markers estimated a 10-fold higher recombination rate in females than in males in the Western Australian seahorse, *H. angustus* (Jones, et al. 1998). This suggests that, in males, recombining regions are physically small, so that most markers will be in large regions that rarely cross over.

Sex-linked regions of three Syngnathidae species

To identify the regions that include the sex-determining genes, we sequenced the genomes of 30 and 10 individuals, respectively, of each sex of captive populations of *H. abdominalis* and *H. erectus* at around 15x coverage per individual. Since no chromosomal assembly of *H. erectus* was available, we assembled the *H. erectus* sequences using

RaGOO (Alonge, et al. 2019) to order and orient scaffolds into chromosomes using the *H. abdominalis* genome assembly as a reference.

Given the known great diversity of teleost sex-linked regions (Kottler and Schartl 2018; El Taher, et al. 2021; Pan, Feron, et al. 2021), and the absence of prior information from seahorses, other than that sex chromosome hetermorphism has not been detected (Liu, et al. 2020; Feron, et al. 2021), we combined multiple approaches to search for completely sex-linked regions. These include comparing males and females for genomic read coverage and sequence diversity (or SNP density) and estimating F_{ST} between the sexes. If a Y- or W-linked region has become highly degenerated, the genomic coverage of the X- or Z-linked region should be twice as high in the homogametic as the heterogametic sex, and Y- or W-specific sequences should have coverage close to zero in the homogametic sex. SNP density and F_{ST} between the sexes can potentially detect less degenerated fully sex-linked regions, which may nevertheless include many sex-specific variants, especially if the regions are extensive, while genome-wide association studies (GWAS) treating sex as a binary trait can detect the sex-linked regions whose sequences are even less differentiated, or are physically small, given a sufficiently high density of sequence variants (Palmer, et al. 2019; Vicoso 2019).

Neither of the two seahorse species, nor the Gulf pipefish, exhibits a pronounced sex difference in coverage or SNP density in any chromosomal region, except for the highly repetitive fusion site in Chr11 in *H. abdominalis* (**Supplementary Fig. S8-14**). Their sex-linked regions may have originated too recently for genetic degeneration to have led to major gene loss, or for extensive sequence divergence to have evolved. Although accurate read mapping can be hindered by polymorphic repetitive elements within Y-linked regions, this should lead to low coverage in males, and is unlikely to lead to falsely suggest a recent origin. Also, neither seahorse species has any genomic region whose read coverage suggests a sex-specific duplication. The *AMHR2* gene has evolved by such a duplication to be a male-determiner (termed *AMHR2Y*) in several other teleosts (Herpin and Schartl 2015; Pan, et al. 2016), and a duplicate putatively sex-determining *AMHR2Y* gene is found in two other syngnathids (Qu, et al. 2021). In both seahorse species, however, this gene has an equal coverage in both sexes (**Supplementary Fig. S15**), consistent with the previous conclusion that it is absent from the seahorse lineage (**Figure 1A**) (Qu, et al. 2021). A previous GWAS study based on RAD-seq data also failed to identify any sex-linked region in *H. abdominalis* (Feron, et al. 2021).

In contrast to the approaches just outlines, GWAS results using whole-genome resequencing data of individuals from captive populations (**Materials and Methods**) did detect sex-linked regions in both seahorse species studied here, *H. abdominalis* and *H. erectus*. Indeed, significantly sex-associated SNPs are found across about two thirds of the *H. abdominalis* Chr6, based on 30 individuals of each sex from the same family

(**Figure 2C, Supplementary Fig. S16**), and almost the entire *H. erectus* Chr4, based on 10 individuals of each sex (**Figure 2D**). We denote these chromosomes by HaChr6 and HeChr4, respectively. In the regions just mentioned, 82% of *H. abdominalis* sex-associated SNPs are heterozygous in most, but not all, males examined (see below), and homozygous in most females (**Figure 2E-F**). In *H. erectus*, 89.3% of the sex-associated SNPs are heterozygous in most males, and homozygous in most females. None of the sex-associated SNPs in *H. abdominalis*, and only 0.4% in *H. erectus* shows higher frequencies of heterozygotes in females. The concentration within large contiguous chromosome regions of sites that are heterozygous much more often in males than females indicates that these extensive regions are either fully or partially, but closely, linked to the sex-determining loci in these seahorses.

In the same captive samples of both species, the candidate sex-linked regions exhibits higher differentiation between the sexes (measured by F_{ST}) than the rest of the chromosome, or other chromosomes (P<0.001, by Mann–Whitney U tests) (**Figure 2C**), and, importantly, heterozygote frequencies are higher in males than females, as detected by negative F_{IS} values across the sex-differentiated region (P<0.001 by Mann–Whitney U tests, see **Supplementary Figs. S17-18**). These results all support male heterogamety with extensive Y-linked regions in both species. Although some other genome regions also show high F_{ST} values between the sexes (including part of *H. erectus* Chr13, **Supplementary Fig. S18, 20**), only the candidate sex-linked regions on HeChr4 and HaChr6 are supported by GWAS analyses. The elevated F_{ST} between the sexes, and strongly negative F_{IS} specific to the differentiated region, are especially clear in HeChr4. We conclude that the previously suggested ZW system in *H. erectus* (Liu, et al. 2020) is not correct, at least for our material (it is possible that heterogamety might vary, as in the platyfish or *Gambusia* species (Black and Howell 1979; Volff and Schartl 2001). Overall, these results, together with those in the previous section, suggest that both the seahorse species studied have highly similar, homomorphic sex chromosomes, or physically small completely sex-linked regions. Therefore, either complete sex-linkage evolved recently or these regions recombine often enough to prevent Y chromosome divergence (see below).

We also analysed the published restriction site-associated DNA sequencing (RAD-Seq) data with on average 9.5x sequencing coverage per individual from 167 male and 57 female *S. scovelli* individuals from a wild population (Flanagan and Jones 2017). Neither GWAS nor F_{ST} analyses, nor analysis of coverage, identified any contiguous sex-linked region (**Figure 2A, Supplementary Fig. S11**), and GWAS analysis of the RAD-seq data did not identify the heterogametic sex (**Figure 2A, Supplementary Fig. S19**): among the 419 sites that were significantly associated with sex in the natural population studied, with $P < 0.05$ by Chi-squared tests, 248 sites were heterozygous only in males, suggesting male heterogamety, but 171 suggested female heterogamety

(**Figure 2B**). These variants are distributed across many different chromosomes. RAD-seq variants are probably too sparse for inferring the actual sex-determining locus. These data, together with the previously reported female-biased sex-ratios (Bolland and Boettcher 2005) in a wild *S. scovelli* population suggest that *S. scovelli* might have environmental sex determination (ESD), though this remains uncertain.

Sex chromosome turnover(s) in seahorses

An autosome must have evolved into a sex chromosome in one or both of the two seahorse species studied. Two outgroup species, *P. taeniolatus* and *S. biaculeatus*, also have XY systems (Qu, et al. 2021). The Syngnathidae ancestor could therefore either have had an XY system or environmental sex-determination. The homolog of Chr4, with the inferred Y-linked region in *H. erectus*, is also a sex chromosome in several cichlid species (El Taher, et al. 2021), and in *P. taeniolatus*. However, it is unlikely that *H. erectus* inherited it as a sex chromosome from a syngnathid ancestor, because as mentioned above, both Syngnathidae species so far studied carry the *AMHR2Y* duplication, but the HeChr4 chromosome does not, and must carry a different male-determining gene. A turnover in *P. taeniolatus* and *S. biaculeatus* lineage that includes the two species with the *AMHR2Y* duplication cannot currently be excluded. Among the cichlid species that were shown to have the homologous sex chromosome (numbered LG07 in cichlids, **Figure 1C**), some species have ZW sex systems, and Chr4 is likely to have become a sex chromosomes convergently in different tribes (El Taher, et al. 2021). Taken together, the data show that the Y-linked region on chr6 of *H. abdominalis*, and probably also that on chr4 of *H. erectus* must have evolved recently, after their divergence from the common ancestor of seahorses.

Defining the boundaries between sex-linked regions and the adjacent pseudoautosomal regions (PARs)

We next describe further genomic evidence in support of the idea above that the seahorse sex chromosomes have large regions that recombine rarely, and small pseudo-autosomal regions. Based on the locations enriched with the centromeric repeat marks noted above (the DNA transposons Zator in *H. abdominalis*, and hAT and LTR Pao in *H. erectus*, **Figure 1D**, **Supplementary Fig. S21**), we inferred that the HaChr6 and HeChr4 sex chromosomes are metacentric or submetacentric, like their autosomal homologs in the other species studied (**Figure 1D**, **Supplementary Fig. S21**). Within HaChr6 and HeChr4, change-point analyses using 50kb sliding windows identified significant changes between adjacent regions in the between-sex F_{ST} values, or using the *P*-values of our GWAS analyses, or the densities of heterozygous sites in males (**Figure 3** and **Supplementary Fig. S17**). These significant changepoints allow us to define the boundaries between regions with different strengths of sex-linkage. In *H. erectus,* a

single clearly defined PAR was inferred, occupying only 1.28 Mb (4.9% of the HeChr4 length); a small and very fragmented region at the other chromosome end is homologous to the high-GC region of the homologous *H. abdominalis* chromosome, and could be a second PAR (**Supplementary Fig. S22**). As HaChr6 is derived via an ancient centric fusion between two chromosomes, two PARs might be expected, because the evolution of sex-linkage on one arm might not be expected to lead to suppressed recombination on the other arm. We indeed found a PAR at each end of HaChr6 (**Figure 3A-B**). GWAS analysis suggests two possible PARs (7.04Mb and 1.79Mb long, respectively) together occupying 8.83 Mb in total (36.1% of HaChr6), but the transitions in F_{ST} and density of male heterozygotes defining them (and especially the right-hand one, on the arm derived from LG11 of *L. oculatus*, see **Figure 3D**), are not as sharp as the one defining the *H. erectus* PAR boundary.

The regions inferred as "non-PAR" probably still recombine, at least occasionally, as associations between genotypes of the sex-linked variable sites and phenotypic sex (based on brood pouch development in mature fish) were incomplete, in both species (some apparently male homozygotes or female heterozygotes were found within the putatively sex-linked regions; **Figure 2E-F**). This might reflect sex-reversals, as reported in European tree frogs (Stöck, et al. 2011), and in the wild in a seahorse species *H. reidi* (Freret-Meurer, et al. 2021). We genotyped 10 individuals of each phenotypic sex for 10 randomly picked *H. abdominalis* sex-associated SNP sites scattered within the candidate completely sex-linked region detected by our initial GWAS analyses. Seven tested sites were heterozygous more often in males than in females (**Supplementary Fig. S23, Supplementary Table S7-8**), but three were homozygous in all 20 *H. abdominalis* individuals tested, of both sexes. Among 59 *H. abdominalis* individuals and 20 *H. erectus* individuals from captive populations whose whole genomes were sequenced in this study (**Materials and Methods**), 15% and 10% of the individuals, respectively, had genotypes expected for the sex other than their phenotypic sex at almost all polymorphic sites ascertained as showing sex-associations (**Figure 2E-F**). Sex linkage is therefore not complete in either sex.

Despite the physically large sizes of the apparently sex-linked regions of HeChr4 and especially HaChr6, their sex-determining systems (especially in *H. erectus*) could have evolved recently. This can be further tested in the future by studying other more closely related seahorse species, such as the sister species of *H. erectus*, *H. hippocampus* (Li, et al. 2021). The sex-linked regions of either or both species studied here could either have evolved as a direct consequence of ancestral crossover localisation (as proposed in the guppy (Bergero, et al. 2019)), or it could have evolved subsequently. In either case, given enough time, this would allow evolution of sequence variants associated with the male-determining allele (linkage disequilibrium, or LD). If recombination was not completely suppressed, Y-X divergence would remain low, and

genetic degeneration would not occur, but F_{ST} between the sexes could be higher than elsewhere in the genomes, as it reflects LD (Charlesworth, et al. 1997).

Ancestral low recombination rates and seahorse sex chromosome evolution

In both *H. abdominalis* HaChr6 and *H. erectus* HeChr4, the centromere positions inferred above are consistent with (peri)centromeric regions being within the sex-linked regions (**Figure 3A-B**). This is also supported by the fact that, in chromatin interaction data from *H. abdominalis* (He, et al. 2021), 76.9% of the HaCh6 sex-linked region is classified as repressive chromatin domains (the B compartment in **Figure 3A**); as expected, A compartment genes have significantly higher expression (for example in brain and testis) than B compartment ones ($P < 2.2e-16$, Mann–Whitney U tests). The sex-determining genes of both seahorse species may therefore have evolved within pericentromeric regions with ancestrally low recombination rates. The autosomal regions homologous to each of these regions in the other seahorse species and the Gulf pipefish also have significantly lower GC content than other genome regions, consistent with repressive chromatin states being established before these regions became sex-linked ($P <0.001$ by Mann–Whitney *U* tests, **Figure 3C**).

We next asked whether the PAR of either species was formerly larger, and has become smaller by chromosome rearrangements, enlarging the ancestrally sex-linked regions. We searched the PAR boundaries and their flanking regions for evidence of inversions that might have suppressed recombination, by examining the paired-end relationships of Illumina reads and male PacBio reads of both species spanning the regions, but found no evidence for such rearrangements (**Materials and Methods, Supplementary Fig. S24-S25**). Alignment between HaCh6 and the homologous LG3 of Gulf pipefish also did not reveal inversions at the PAR boundary (**Supplementary Fig. S26**). Within these species' sex-linked regions, we also did not detect any signatures of new evolutionary strata; there are no consistent regional differences in male-female F_{ST} values, nor in the numbers of sites that are predominantly heterozygous in males but homozygous in females (**Figure 3A-B, Supplementary Fig. S17**). Based on sex-linked genes in which all variable sites are heterozygous in all *H. abdominalis* males but homozygous in females (1.12% of the sex-linked genes), and after excluding individuals that likely have undergone sex-reversals (**Figure 2E-F**), we estimated the average synonymous site divergence between the X and Y (K_S) to be 0.002; in *H. erectus* the corresponding K_S estimate was 0.006 (in which 423, or 41.6%, of the sex-linked genes satisfied this criterion). As this must over-estimate the divergence, the very low values indicate that both species' sex-linked regions either originated very recently, or, more likely, based on the evidence above, still recombine, albeit at a very low rate. Even in HaChr6, in which F_{ST} between the sexes is quite low (**Figure 3A**), significant changes in F_{ST} and GC content coincide, supporting the conclusion that their recombination rates

have been different over considerable evolutionary times, and thus our classification into PAR versus MSY regions.

One of the *H. abdominalis* PAR boundaries identified above coincides with an ancient chromosome fusion site (**Figure 3D**), suggesting that the sex-linked region evolved near an ancestral centromere created in a centric fusion event. The HaChr6 sex-linked region represents the larger of two ancient fish microchromosomes (the homolog of the spotted gar LG11), and the PAR represents the smaller one (homologous with the spotted gar LG24) involved in the fusion, with the centromeric region separating them (**Figure 3D**). The chromosome arm that evolved a male-determining factor might already have been largely non-recombining in males, or may have stopped recombining subsequently, but the other arm could have remained unaffected, and would be unlikely to stop recombining by a pericentric inversion, as these are rare events (although, as mentioned in the Introduction, recombination suppression can evolve without an inversion). Selection may not have favored suppressed recombination on this arm, perhaps because no SA polymorphisms became established in this small chromosome arm in the short evolutionary time since HaChr6 became a sex chromosome.

Different TE subfamilies are specifically enriched at the PAR boundaries of both seahorse species studied, compared with their genome averages. In *H. abdominalis*, long interspersed nuclear element (LINE) L1 elements showed a 6.82-fold enrichment, and the DNA transposon Mule-MuDRs showed a 6.03-fold enrichment. In *H. erectus*, LINE/L1-Tx1 sequences were 6.67-fold enriched at the PAR boundary (**Figure 3A-B**). However, analyses of Illumina reads spanning the TE insertions yielded no evidence that they are male-specific. In the other seahorse and Gulf pipefish species, neither LINE/L1 nor LINE/L1-Tx1 elements were enriched in the homologous regions (**Supplementary Fig. S27**), suggesting that these are recent species-specific insertions of these TEs.

Candidate sex differentiation genes in the two seahorse species

The different locations of the sex-linked regions in the two seahorse species (**Figure 2D**) implies independent acquisition of a male-determining gene in at least one species, and parts of HeChr4 and HaChr6 must have been evolving under at least partial sex-linkage for different evolutionary times. These regions may have become enriched for genes participating in gonadogenesis and gonad differentiation, and, as mentioned in the introduction, changes in expression of genes in the region with sex-related functions, including in the male pouch, might also evolve after their male-determining factors evolved (Ellegren and Parsch 2007). To identify genes of these kinds, we searched gonad transcripts from *H. abdominalis* and *H. erectus* (**Supplementary Fig. S28, Materials and Methods)** for genes transcribed at different levels in gonads of either sex compared with other tissues, or whose transcription patterns have changed compared with their

autosomal orthologs in the other seahorse or Gulf pipefish species. As it is unclear exactly when in development sex determination is initiated in seahorse species, we sampled tissues spanning different developmental stages, from the time when gonads first become morphologically visible until they become distinctly different in the two sexes. The genes described below are therefore more likely to function in sex differentiation processes than in sex-determination, especially if they are only transiently expressed in one sex. Nevertheless, sex differences in expression of the sex-determining genes sometimes extend to late developmental stages. For example, in amphibians *AMH* is predominantly expressed in differentiating testes, after the sex-determining stage (Pan, Kay, et al. 2021).

The HeChr4 sex-linked region includes 1,018 identified genes, one of which, *CYP11A1*, may act during male gonad differentiation. In both seahorse species, *CYP11A1* showed testis-biased transcription (relative to other male somatic tissues, including muscle, brain, kidney, and brood pouch). Its transcript levels increased during testis development and male gestation in *H. erectus*, consistent with an important male function having evolved in this species (**Supplementary Fig. S29).** In *H. abdominalis*, this gene is not sex-linked, and the average expression level across different testis development stages was 5.38 times lower than in *H. erectus*, and transcript levels decreased during testis development (**Figure 4A, Supplementary Fig. S29**). Its mammalian ortholog encodes a key steroidogenic enzyme (Morohashi, et al. 1993), and it was reported to have sex-biased transcription which changes during gonad development in other teleosts, including Nile tilapia and olive flounder (Ijiri, et al. 2008; Liang, et al. 2018). Another candidate male-determining gene is *WT1*, which was reported to be essential for mammalian male determination (Hossain and Saunders 2001), and showed higher transcript levels in *H. erectus* testis than somatic tissues. Its transcript level also increased during testis development and male gestation (**Supplementary Fig. S29**), but was 2.5 times lower on average than in *H. abdominalis*.

The HaCh6 sex-linked region includes 624 identified genes, none of which is orthologous to known vertebrate sex-determining or sex-differentiation genes. This species' sex determination may have arisen via a turnover event, as outlined above, and may involve a gene with no previously known function in gonad development, as was found for the rainbow trout male-determining gene *sdY* (Bertho, et al. 2018). One interesting sex-differentiation candidate gene is present, however; this is *MSH5*, whose mouse ortholog is required for proper meiosis (de Vries, et al. 1999) (mouse mutants show reduced gonad size and infertility in both sexes (Edelmann, et al. 1999). This gene is transcribed at a 47-fold higher level in the immature *H. abdominalis* testis and 24-fold higher in the mature testis, relative to ovary or somatic tissues at the same stage; the *H. erectus* ortholog shows no such bias, and had an average transcript level in testis 6.51 times lower than in *H. abdominalis* (**Figure 4A**, **Supplementary Fig. S29**). *MSH5* is located close to the putative HaCh6 centromere.

Sex chromosome turnovers may also be followed by changes in transcription of genes involved in the sex differentiation pathway located elsewhere in the genome. To test this, we compiled a list of 72 genes previously reported to be involved in gonad development, sex determination or gonad differentiation of other teleost species (Chen, et al. 2018; Guiguen, et al. 2018; Shen and Wang 2018; Taboada, et al. 2018) (**Figure 4A, Supplementary Fig. S30**). Most of these genes are autosomal in both seahorse species studied here, and most showed gonad transcription patterns similar to those of their orthologs in the other seahorse species or other teleosts. For example, orthologs of the generally male-determining genes *DMRT1* and *AMH* showed testis-specific transcription in both seahorse species. The ortholog of *CYP19A*, whose protein product converts androgens to estrogens and plays an important role in female-determination in many other teleost species (Shen and Wang 2014), is specifically transcribed in the ovary of *H. abdominalis*. The orthologs of two germline genes, *VASA* and *PIWI*, are transcribed specifically in gonads of both sexes in both seahorse species. Finally, a few genes had the highest transcript levels in the gonads of the opposite sexes in different species (**Figure 4A**).

However, we detected intriguing changes in some sex-differentiation pathway genes in *H. abdominalis* (**Figure 4A**), as might be expected given that a turnover must have occurred in this species (although the event is recent enough that changes might not have occurred). None of these genes is in the newly evolved sex-linked region of HaChr6. One example is *FOXL2*, a highly conserved gene that is predominantly transcribed in ovary granulosa cells of both mammals and some other teleosts (Bertho, et al. 2016), but shows testis-specific transcription in *H. abdominalis* but not *H. erectus*. In contrast, the *GSDF* gene, which is the male-determining gene in several other teleosts (Pan, et al. 2016) has become specifically transcribed in the ovary of *H. abdominalis*.

Evolution of gene transcription in the newly evolved sex-linked region of *H. abdominalis*

Finally, to investigate possible changes in gene expression after genome regions newly became sex linked, we compared transcription patterns, particularly in the sex-specific organs (gonads and male brood pouch), in the two seahorse species. During its development, the pouch undergoes drastic tissue remodeling (Whittington, et al. 2015), with 4,505 genes (19.7% of all *H. abdominalis* annotated genes) changing their transcription between immature and mature pouch stages. The set of genes that were differentially expressed between these pouch stages was also enriched in genes with expression differences between male pouch and female epidermal tissue (representing the likely ancestral transcription state) in reproductively mature animals, using a threshold of at least twice the mean TPM of the female tissue from specimens ($P < 2.2e-16$ by Fisher's exact test; **Supplementary Fig. S31**). We also defined 17 genes as pouch-biased, because they had higher expression than in other male tissues at the same developmental stage (muscle, brain, kidney). 8 of these showed higher transcript levels

in the male pouch than in female abdominal epidermal tissues of *H. abdominalis*. We compared the representation of such genes, and genes transcribed at higher levels in mature male brood pouch tissue, relative to other tissues at the same developmental stage (muscle, brain, kidney), in different *H. abdominalis* genome regions against the genome-wide background by Fisher's exact tests. This analysis suggested over-representation (P <0.05, Fisher's test) of such genes in the region of HaCh6 that recently became sex-linked (which includes 17 such genes), but not among HaCh6 PAR genes (though chromosome 7 also showed an enrichment in such genes, see **Figure 4B** and **Supplementary Fig. S32**). Of the 13 autosomal orthologs of the 17 pouch-biased genes in *H. erectus*, 2 are pouch-biased, versus 11 that are not, again consistent with changes having evolved in the *H. abdominalis* genome (no orthologs were found on the homologous *H. erectus* chromosome for 4 of these genes). Thus some HaChr6 genes probably evolved increased expression in the male brood pouch after this *H. abdominalis* chromosome arm acquired the present male-determining gene, as expected for male-beneficial mutations.

Supporting the hypothesis that their male-biased transcription relates to pouch development, the *H. abdominalis* genes with pouch-biased expression are enriched for GO terms including immune response (e.g., 'T cell receptor signaling pathway', 'innate immune response', 'cellular homeostasis' etc.) and hemopoiesis ('leukocyte activation and differentiation', 'hematopoietic organ development' etc.; **Supplementary Table S9**), consistent with previous results showing that evolution of male pregnancy in seahorses is associated with turnovers of immune-related genes (Roth, et al. 2020). Genes with higher transcription levels in female epidermal tissue than male pouch tissue showed no such enriched GO terms (**Figure 4C**). Mutants of the mouse orthologs of these seahorse genes are also enriched in phenotypic defects in immune or hematopoietic systems (P < 0.01, Fisher's exact test), while the female epidermal tissue biased genes are not (**Figure 4D, Supplementary Table S10**).

The *H. erectus* HeChr4 sex-linked region is enriched for testis-biased genes. However, the enrichment probably pre-dates the divergence of these two species, as its autosomal homolog in *H. abdominalis* shows the same enrichment (P <0.01 for both, regardless of different definitions of tissue-biased genes (see **Materials and Methods and Supplementary Fig. S33**).

Discussion

Sex chromosome turnovers in seahorses and the origins of their extensive non-recombining regions

Sex chromosome changes in the Syngnathidae species appear to have preserved male heterogamety, including in the two seahorse species studied here, which diverged about 11.4 MYA. This is consistent with the prediction that turnover events preserving

the same heterogametic sex should be more frequent than those changing the heterogamety (Bull 1983). This has also been found in large comparative studies of true frogs (Jeffries, et al. 2018) and cichlids (Gammerdinger and Kocher 2018; El Taher, et al. 2021). It has been suggested that sex chromosome turnovers are responses to genetic degeneration and accumulated deleterious mutational load (Blaser, et al. 2013), to sexual conflict(van Doorn and Kirkpatrick 2007), or both together in the 'hot potato' model of (Blaser, et al. 2014)(which predicts conservation of patterns of heterogamety and non-random recruitment of certain autosomes as sex chromosomes during their transition), or to meiotic drive (Úbeda, et al. 2015), or that they are selectively neutral changes under genetic drift (Veller, et al. 2017). The reasons for individual turnover events are unknown, but many species, including the two seahorse species studied here, do not have strongly degenerated sex-linked regions chromosomes as required by the hot potato model, which suggests that transitions can occur without extensive accumulated deleterious mutations.

A factor that is rarely mentioned is the possibility that the recombination landscape might pre-date a turnover. Ancestrally low recombination across large genomic regions, either in males or in both sexes, favours sex chromosome transitions. This is because, if a new male-determining factor arises within such a region, this immediately creates a region of complete (or almost complete) sex linkage that is transmitted like a sex chromosome (Bergero, et al. 2019). The same may have occurred in turnover events in tilapias, which are members of the cichlid group of fish (Tao, et al. 2021).

Localization of crossovers to the ends of chromosomes in meiosis of males (or both sexes) may be widespread among other teleosts (Sardell and Kirkpatrick 2020), as well as in amphibians (Dufresnes, et al. 2021) and plants (Charlesworth 2019; Rifkin, et al. 2021). Accurate fine-scale sex-specific genetic maps are scarce for teleosts, apart from the guppy (Bergero, et al. 2019) , some sticklebacks (Sardell, et al. 2018; Sardell, et al. 2021), and the Atlantic halibut (Edvardsen, et al. 2022) and Atlantic herring (Pettersson, et al. 2019), so only crossover localization in both sexes will often be detectable. Among these species, strong male-specific crossover localization was found in all except for the Atlantic herring. Our results in seahorse species suggest that recombination was indeed ancestrally low in the current Y-linked regions, though no sex-specific recombination map is available in *H. abdominalis* or *H. erectus*, and our genome-wide estimates provide only sex-averaged recombination rates. However, as mentioned above, microsatellite analyses in one seahorse species *H. angustus* (Jones, et al. 1998) found a much higher recombination rate in females than in males, and the GC spikes at the tip regions of chromosomes in the two seahorses species studied here (**Figure 1**) suggest a chromosomal recombination landscape similar to that in the other fish species just mentioned. This should be further tested by genetic mapping in these species. Newly developed methods using linked-read sequencing of sperm DNA

(Dréau, et al. 2019; Yoshitake, et al. 2022) hold promise for future examination of crossover patterns in seahorse species.

We also found that the present boundary between the PAR and putatively sex-linked region of HaCh6 coincides with the location of an ancient centric fusion between two chromosomes, which occurred long before this chromosome became a sex chromosome. After the fusion, this genome region must have been a centromeric or pericentromeric region, probably with a lower recombination rate than other parts of the new (fused) Chr6, even in females. If crossovers in this ancestral species were localized to the chromosome ends in males, the fusion could have created larger regions of low recombination on one or both arms. If the chromosome gained a male-determining factor in a turnover event, this would create a new Y-linked region.

Non-randomness of chromosomes involved in sex chromosome turnovers, and subsequent evolutionary changes in new sex-linked regions

Sex chromosome turnovers can occur in two ways (Pan, et al. 2016; Vicoso 2019), either through translocation of a gene or gene duplication with a sex-determining function to a new genomic location (as in the house fly, see (Sharma, et al. 2017), or common seadragon and alligator pipefish (Qu, et al. 2021)) , or by acquisition of sex-determining function by a mutation in a pre-existing gene (termed “diversification”). In the former case, any chromosome might gain a new sex-determining factor, though this is most likely when a sexually antagonistic polymorphism is maintained in the new location (van Doorn and Kirkpatrick 2010). If so, new sex-determiners are expected to arise in genome regions already carrying genes involved in gonad functions, or other functions where sexually antagonistic mutations may arise. In the diversification process, chromosomes whose ancestral gene content is enriched for genes involved in the sex-determining cascade are more likely than others to become sex chromosomes. In either case, some chromosomes will be expected to gain sex-determining factors more often than others, as is indeed observed in amniotes (O'Meally, et al. 2012; Kratochvíl, et al. 2021) and even within the cichlids (Gammerdinger and Kocher 2018; El Taher, et al. 2021). The ancestral gene content could thus have facilitated the gain of a new male-determining gene.

Chr4 is enriched for testis-biased genes in both studied seahorse species (**Figure 4B**). It is also the XY sex chromosome in another Syngnathidae species, the common sea dragon (Qu, et al. 2021) (**Supplementary Fig. S34**) due to a recent insertion creating the *AMHR2Y*; it is homologous to a part (21Mb or 32.7%) of a Nile tilapia chromosome LG7 (**Supplementary Fig. S35**) that has convergently evolved to become a sex chromosomes in multiple other cichlid species, in Lake Tanganyika (including in the families of Eretmodini, Bathybatini, Perissodini, and in the species *Hemibates stenosoma* (El Taher, et al. 2021); **Figure 2D**). These results are consistent with the hypothesis that the ancestral chromosome carried genes capable of acquiring mutations conferring an

upstream male-determining function, making it especially likely to become a sex chromosome (rather than that recent evolution of Y-linkage triggered changes towards higher expression in testis).

In contrast, after part of HaChr6 became sex-linked, we detect increased pouch-specific transcription that was not seen in the autosomal orthologs in *H. erectus*. This may reflect the predicted 'masculinization', in which newly Y-linked alleles with male-related functions undergo adaptive evolution. In a seahorse species, any pouch genes in the new Y-linked region may be the first to adapt to such a genomic change. Pronounced genetic degeneration is probably expected to evolve later than such adaptive changes, as it depends on slower processes involving genetic drift of deleterious mutations, or their hitch-hiking effects along with advantageous mutations, under complete sex-linkage (Bachtrog 2008). Consistent with this, the evidence presented here shows that the seahorse Y-linked regions have not lost high proportions of genes, despite being large genome regions. They may have evolved recently, or degeneration of these regions may have been slowed by occasional recombination with the X, as is suggested by the small Y-X divergence estimates for both HaChr6 and HeChr4.

Materials and Methods

DNA sampling and sequencing

The big-belly seahorse and lined seahorse samples used in this project were obtained from cultured populations in Fujian (China), with approval by the Animal Ethical and Welfare Committee of the Fisheries Research Institute of Fujian (Approval No. IACUC-2019-031). The original population of the big-belly seahorse was established from about 10,000 individuals caught from the wild, and that of lined seahorse from about 2,000 to 3,000 individuals. Before tissue collection, all the seahorses were cultured in indoor breeding tanks (150L), with seawater from the sea and adjusted to a salinity of 31 ‰ and temperature 16 -18 ℃. The seahorses were fed three times per day (08:00, 12:00 and 16:00) with mysid. All the seahorses were sacrificed after being anaesthetized for 30 min by MS-222 solution at a concentration of 50 mg/L, which is approved by FDA (Food and Drug Administration, USA). To infer the sex-linked regions, we collected 30 male and 29 female individuals of *H. abdominalis*, and 10 male and 10 female *H. erectus*. Muscle tissue from each individual was dissected under a stereomicroscope and stored at -80°C until processing. DNA extraction was carried out using EasyPure® Genomic DNA Kit (TransGen Biotech, China). A paired-end library was constructed with an insert size of 250 base pairs (bp) according to the protocol provided by the manufacturer, and was sequenced on an Illumina X ten platform by Annoroad Gene Technology Co. Ltd. (http://www.annoroad.com).

mRNA sequencing and gene expression analysis

We collected tissues from brain, kidney, testis, ovary, male brood pouch and female abdominal skin of *H. abdominalis* individuals, at both immature (3 months old) and mature (5 months) stages. For each tissue sample, we prepared two biological replicates for total RNA extractions using the EasyPure RNA Kit (TransGen Biotech, China) following the manufacturer's protocol. Unstranded total RNA library was constructed and sequenced with a read length of 150 bp using the Illumina HiSeq 2000 sequencing platform. The recently assembled *H. abdominalis* genome (He, et al. 2021), as well as the published genomes of *H. erectus, S. scovelli* and *O. niloticus* (Small, et al. 2016; Lin, et al. 2017; Tao, et al. 2021) were used as references for gene expression analyses. A chromosome level genome assembly of *H. erectus* was based on the *H. abdominalis* reference genome (Alonge, et al. 2019) using RaGOO. To investigate biased gene expression in different tissues in *H. erectus*, we also collected published transcriptome data from *H. erectus* testis, ovary, brood pouch, muscle and kidney (Lin, et al. 2016). RNA-seq reads were mapped to the respective genomes using Hisat2 (Kim, et al. 2015)(2.1.0) and the reads mapped to each gene were counted uing featureCounts version 1.6.2 (Liao, et al. 2014). Read counts were normalized using the TPM (Transcripts Per Kilobase per Million reads) method. Differentially expressed gene analyses to compare tissue types, developmental stages and sexes were performed with the edgeR R packge (Robinson, et al. 2010). Genes with low expression were filtered with the filterByExpr function in the edgeR package, using the default parameters (min.count = 10, min.total.count = 15, min.prop = 0.7). We used the Benjamini and Hochberg's algorithm (BH) to control the false discovery rate (FDR). We classified only genes with BH-adjusted *P*-value <0.05 (exact negative binomial test) as significantly differentially expressed. After excluding genes not expressed (TPM<1) in any tissue, we estimated each gene's tissue-specificity according to the tau (τ) algorithm(Yanai, et al. 2005), which performed best in recognizing tissue-specific genes in a benchmark study(Kryuchkova-Mostacci and Robinson-Rechavi 2017). To enable comparisons across different tissues, we performed a quantile normalization on the entire dataset before calculating tau values using the tispec R-package (https://rdrr.io/github/roonysgalbi/tispec). The threshold for a gene to be classified as tissue-specific was set to 0.8, following a previous study (Kryuchkova-Mostacci and Robinson-Rechavi 2017). If a tissue-specific gene defined by tau value also had a top-three normalized expression level in a tissue, we classified it as expressed specifically in that tissue. We also did analyses after defining tissue-biased genes as genes with at least two-fold greater TPM than the mean TPM in other tissues of the same developmental stage, to test whether the conclusions our robust to different definitions.

Annotation of repetitive sequences

We used RepeatModeler version 2.0 (Flynn, et al. 2020) for *de novo* identification of repetitive elements in the genome sequences of the species studied here, and to

classify them based on their sequence similarity to known repeat families from other organisms. We then combined the all repeat families identified with annotated repeats in the Dfam_3.1(Hubley, et al. 2016) and RepBase-20170127 databases and used RepeatMasker version 4.0.9 (Tarailo-Graovac and Chen 2009) to annotate the repeat sequences in our genome sequences.

Comparative genomics analysis

Chromosome-level genome assemblies of teleost species, the Gulf pipefish, *S. scovelli* (Small, et al. 2016), Nile tilapia, *Oreochromis niloticus* (Tao, et al. 2021) and Guppy, *Poecilia reticulata* (Kunstner, et al. 2016), as well as the genome of spotted gar, *L. oculatus* (Braasch, et al. 2016) were aligned to the chromosome-level assembly of the big-belly seahorse, *Hippocampus abdominalis* (He, et al. 2021), using a large-scale alignment tool, LAST (https://gitlab.com/mcfrith/last), using the commands: lastdb -uNEAR -R01 ./out.ref ref.genome.fa; lastal -P32 ./out.ref query.genome.fa | last-split > ./out.maf; maf-swap ./out.maf | last-split > ./out.1to1.reverse.maf; maf-swap ./out.1to1.reverse.maf > ./out.1to1.maf. Syntenic relationships were visualized with the Circos software, version 0.69.9 (Krzywinski, et al. 2009).

Identification of the candidate sex determining regions

To attempt to identify the sex-linked region of *Syngnathus scovelli* (the Gulf pipefish), demultiplexed ddRAD-seq data (Flanagan and Jones 2017) from males and females (accession #: PRJNA358088) were downloaded from NCBI Sequence Read Archive (SRA) . Demultiplexed reads were analyzed with the RADSex computational workflow (Feron, et al. 2021) using the *radsex* software version 1.1.0 (https://github.com/SexGenomicsToolkit/radsex). Tables of sex-associated markers and their depths of coverage were created with *process* command. The distribution of markers in the two sexes was computed with *distrib,* and markers significantly associated with phenotypic sex were extracted with *signif* using variants with a minimum depth of 5 (-d 5) for both commands, and the default for all other settings. 611,484 markers with minimum depth = 5 were found in at least one individual, among which 419 were significantly associated with sex (*radsex signif*, $P < 0.05$, Chi-squared test with Bonferroni correction, highlighted tiles in **Supplementary Fig. S15**). Among these 419 markers, 53 (12.6%) had reads aligned to unanchored scaffolds, and 263 (62.8%) had reads not uniquely aligned with mapping quality higher than 15. The remaining 103 markers were distributed sparsely among the 22 LGs, with between a and 15 markers on each LG. F_{ST} between the Gulf pipefish male and female populations was computed with the --fstat option in the Stacks software, version 2.5.3 (Rochette, et al. 2019).

We performed whole genome resequencing of 30 male and 29 female individuals to infer the sex-linked region in *H. abdominalis*, plus 10 male and 10 female individuals of *H. erectus*. For each individual, sequencing reads were mapped to the reference genome with BWA-mem (Li and Durbin 2009) (0.7.17-r1188). We called the SNPs using HaplotypeCaller in GATK (DePristo, et al. 2011) (3.8.1.0) and then called the genotypes

with GenotypeGVCFs. The SNP results were filtered with following filtering expression: "QD < 2.0 || FS > 60.0 || MQRankSum < -12.5 || RedPosRankSum < -8.0 || SOR > 3.0 || MQ < 40.0". Only biallelic sites with minor allele frequency greater than 0.05 were kept. A total of 6,115,914 and 5,603,399 SNPs in *H. abdominalis* and *H. erectus*, respectively, passed the quality control. We used Beagle version 5.0 (Browning, et al. 2018) to impute missing genotypes, and SHAPIT v2.r904 (Delaneau, et al. 2013) for haplotype estimation (phasing). A genome wide association study (GWAS) was performed by mixed-model association using eXpedited (beta-07Mar2010 version of EMMAX); (Kang, et al. 2010), using sex as a phenotype. Phased genotypes were processed with PLINK v1.90b6.10 (Purcell, et al. 2007); http://pngu.mgh.harvard.edu/purcell/plink/) to generate the input for EMMAX. The threshold for genome-wide significance in the GWAS was set at *P*-value = 8.18e-9 for *H. abdominalis*, and 8.92e-9 for *H. erectus* by dividing 0.05 with the number of total SNPs in each species. For *H. abdominalis*, the sex-linked region was inferred based on presence of SNPs significantly associated with sex on HeChr6 by the GWAS analysis. Phasing either with or without imputation did not influence the enrichment of sex-associated SNPs in the sex-linked region of *H. abdominlais* Chr6 or *H. erectus* Chr4. The boundaries of the HeChr4 region were inferred on the same basis and using the density of sites at which males are heterozygous and females are homozygous. We used nucmer (Marçais, et al. 2018) to infer the homologous regions of identified seahorse sex-linked regions in other species. Only one-to-one homologous alignments were kept. We validated the genotypes of 10 randomly picked *H. abdominalis* sex-associated SNP sites by performing polymerase chain reaction (PCR) experiments using FastLong PCR MasterMix (PC8001, Aidlab) according to the manufacturer's protocol, followed by sequencing to identify the SNP genotypes. The primers used are listed in **Supplementary Table S8**.

We used 50kb overlapping sliding windows with step size 25kb to calculate the F_{ST} values between male and female populations, using vcftools version 0.1.13 (Danecek, et al. 2011), and coverage depths of merged reads from males or females, using sambamba version 0.7.0 (Tarasov, et al. 2015). Heterozygous sites found only in males were defined after excluding the possibly sex-reversed individuals (see the Results section). Synonymous site divergence between X and Y was calculated based on such male-specific SNPs using the KaKs_Calculator software (Zhang, et al. 2006) and Nei and Gojobori's (Nei and Gojobori 1986) correction for multiple changes. F_{IS} was calcaulated according to Nei (Nei 1973), $F_{IS} = 1–H_o/H_e$, where H_o and H_e represent observed and expected heterozygosities of each SNP.

Recombination rates (rho) were estimated after excluding SNPs with >10% missing individuals in our sample, or located within repetitive regions, using FastEPRR2.0 (Gao, et al. 2016) with parameters winLength=50000, stepLength=25000, winDXThreshold=10. Changepoint analyses of GC%, F_{ST}, the density of male-specific heterozygous sites, and log-transformed P values in our GWAS, were conducted with

the R package "changepoint" (Killick and Eckley 2014) to implement the binary segmentation method (Bai 1997), with a maximum of 3 changepoints; change-points were detected based on both mean value and variance.

Searches for inversions in the *H. abdominalis* and *H. erectus* genomes

We searched for read pairs supporting inversions in our candidate sex-linked region based on their orientation and mapping coordinates, using TIDDIT (Eisfeldt, et al. 2017) (https://github.com/SciLifeLab/TIDDIT). Read pairs aligning on the same chromosome at a distance higher than the insert size, and having the same orientation, were considered as candidates supporting inversions. The fractions of read pairs supporting inversion events were used to define the genotypes in these regions, defining fractions less than 100% and greater than 25% as candidate heterozygotes, and less than 25% as homozygotes for the reference arrangement. Inversion events were identified in PacBio sequences using npINV (Shao, et al. 2018). PacBio reads containing a pair of alignments mapping to the same chromosome, but with a different orientation, were considered to support inversions.

A/B compartments

We processed the Hi-C data (He, et al. 2021) with HiCPro (Servant, et al. 2015). Mapped read-pairs were used to generate raw Hi-C contact matrix at a 50kb resolution using hicBuildMatrix of the HiCExplorer (2.2.1) suite (Wolff, et al. 2018; Wolff, et al. 2020). We used the ICE method implemented in hicCorrectMatrix to remove the bins with zero or low number and extremely high number of reads because the former bins usually tend to indicate repetitive regions and the latter bins may represent copy number variations. After correcting the Hi-C matrices (hicCorrectMatrix), we used cooltools (0.3.2)(https://cooltools.readthedocs.io/) and its call-compartments function to obtain the first eigenvector values (PC1) of each chromosome of the Hi-C matrices with a 50kb resolution. Regions with positive PC1 values were assigned as A (active) compartments and those with negative PC1 values were assigned as B (inactive) compartments, adjusted by the GC percentage of the region.

Gene Ontology (GO) and phenotype enrichment analysis

Orthologs of *H. abdominalis* and *H. erectus* found in the *Danio rerio* (zebrafish) genome assembly GRCz11 (danRer11) and *Mus musculus* (house mouse) genome assembly GRCm39 (mm39) from NCBI (National Center for Biotechnology Information) were used as input in Gene Ontology and phenotype enrichment analysis. We performed the GO analyses by Metascape (Zhou, et al. 2019) with default parameters (P<0.01, min overlap=3 and min enrichment=1.5). We used phenome2 (Weng and Liao 2017)(http://evol.nhri.org.tw/phenome2/) to detect the enriched phenotypes in each gene set. Zebrafish and mouse phenotypes with Bonferroni corrected P-value < 0.01 (Fisher's exact test) were considered as enriched.

Data Accessibility

Whole genome sequencing reads from both sexes of *H. abdominalis* and *H. erectus* individuals, as well as RNA-seq reads of *H. abdominalis* tissues have been deposited on NCBI Short Reads Archive under the BioProject Accession Number PRJNA736169. Other published data used in this project were listed in **Supplementary Table S11**.

Acknowledgments

Animal icons are from https://thenounproject.com (including "Testis" icon by Becris, "Ovary" icon by Dairy Free Design, "Kidney" icon by Alexander Blagochevsky, "Muscle" icon by tezar tantular, "Brain" icon by Pixel Bazaar and "Garfish" icon by Phạm Thanh Lộc). This work is supported by the National Natural Science Foundation of China (32170415, 32061130208), Natural Science Foundation of Zhejiang Province (LD19C190001), European Research Council Starting Grant (grant agreement 677696) to Q.Z. and the 13th Five-Year Plan for the Marine Innovation and Economic Development Demonstration Projects (FZHJ11) to Z.H.

Reference

Alonge M, Soyk S, Ramakrishnan S, Wang X, Goodwin S, Sedlazeck FJ, Lippman ZB, Schatz MC. 2019. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. Genome Biol. 20:224.

Arbeithuber B, Betancourt AJ, Ebner T, Tiemann-Boege I. 2015. Crossovers are associated with mutation and biased gene conversion at recombination hotspots. Proc Natl Acad Sci U S A 112:2109-2114.

Bachtrog D. 2008. The temporal dynamics of processes underlying Y chromosome degeneration. Genetics 179:1513-1525.

Bai J. 1997. Estimating Multiple Breaks One at a Time. Econometric Theory 13:315-352.

Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho T-J, Koutseva N, Zaghlul S, Graves T, Rock S, et al. 2014. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. Nature 508:494-499.

Bergero R, Gardner J, Bader B, Yong L, Charlesworth D. 2019. Exaggerated heterochiasmy in a fish with sex-linked male coloration polymorphisms. Proc. Natl. Acad. Sci. U. S. A. 116:6924-6931.

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. U. S. A. 115:12781-12786.

Bertho S, Pasquier J, Pan Q, Le Trionnaire G, Bobe J, Postlethwait JH, Pailhoux E, Schartl M, Herpin A, Guiguen Y. 2016. Foxl2 and Its Relatives Are Evolutionary Conserved Players in Gonadal Sex Differentiation. Sex Dev. 10:111-129.

Bian C, Hu Y, Ravi V, Kuznetsova IS, Shen X, Mu X, Sun Y, You X, Li J, Li X, et al. 2016. The Asian arowana (Scleropages formosus) genome provides new insights into the evolution of an early lineage of teleosts. Sci. Rep. 6:24501.

Black DA, Howell WM. 1979. The North American mosquitofish, Gambusia affinis: a unique case in sex chromosome evolution. Copeia:509-513.

Blaser O, Grossen C, Neuenschwander S, Perrin N. 2013. Sex-chromosome turnovers induced by deleterious mutation load. Evolution 67:635-645.

Blaser O, Neuenschwander S, Perrin N. 2014. Sex-chromosome turnovers: the hot-potato model. Am. Nat. 183:140-146.

Bolland J, Boettcher A. 2005. Population structure and reproductive characteristics of the gulf pipefish,Syngnathus scovelli, in Mobile Bay, Alabama. Estuaries 28:957-965.

Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, Amores A, Desvignes T, Batzel P, Catchen J, et al. 2016. The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. Nat. Genet. 48:427-437.

Browning BL, Zhou Y, Browning SR. 2018. A One-Penny Imputed Genome from Next-Generation Reference Panels. Am. J. Hum. Genet. 103:338-348.

Bull JJ. 1983. Evolution of Sex Determining Mechanisms: Benjamin-Cummings Publishing Company.

Charlesworth B, Nordborg M, Charlesworth D. 1997. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet Res 70:155-174.
Charlesworth D. 2019. Young sex chromosomes in plants and animals. New Phytol 224:1095-1107.
Charlesworth D, Zhang Y, Bergero R, Graham C, Gardner J, Yong L. 2020. Using GC content to compare recombination patterns on the sex chromosomes and autosomes of the guppy, Poecilia reticulata, and its close outgroup species. Mol Biol Evol.
Chen S-L, Zhou Q, Shao C-W. 2018. Genomic and Epigenetic Aspects of Sex Determination in Half-Smooth Tongue Sole. In. Sex Control in Aquaculture. p. 525-545.
Cheng P, Huang Y, Lv Y, Du H, Ruan Z, Li C, Ye H, Zhang H, Wu J, Wang C, et al. 2021. The American paddlefish genome provides novel insights into chromosomal evolution and bone mineralization in early vertebrates. Mol Biol Evol 38:1595-1607.
Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A, Waters PD, Grützner F, Kaessmann H. 2014. Origins and functional evolution of Y chromosomes across mammals. Nature 508:488-493.
Costantini M, Auletta F, Bernardi G. 2007. Isochore patterns and gene distributions in fish genomes. Genomics 90:364-371.
Dagilis AJ, Sardell JM, Josephson MP, Su Y, Kirkpatrick M, Peichel CL. 2022. Searching for signatures of sexually antagonistic selection on stickleback sex chromosomes. Philos Trans R Soc Lond B Biol Sci 377:20210205.
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 27:2156-2158.
de Vries SS, Baart EB, Dekker M, Siezen A, de Rooij DG, de Boer P, te Riele H. 1999. Mouse MutS-like protein Msh5 is required for proper chromosome synapsis in male and female meiosis. Genes Dev. 13:523-531.
Delaneau O, Zagury J-F, Marchini J. 2013. Improved whole-chromosome phasing for disease and population genetic studies. Nat. Methods 10:5-6.
DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43:491-498.
Dréau A, Venu V, Avdievich E, Gaspar L, Jones FC. 2019. Genome-wide recombination map construction from single individuals using linked-read sequencing. Nat. Commun. 10:4309.
Dufresnes C, Brelsford A, Baier F, Perrin N. 2021. When sex chromosomes recombine only in the heterogametic sex: heterochiasmy and heterogamety in Hyla tree frogs. Mol Biol Evol 38:192-200.
Edelmann W, Cohen PE, Kneitz B, Winand N, Lia M, Heyer J, Kolodner R, Pollard JW, Kucherlapati R. 1999. Mammalian MutS homologue 5 is required for chromosome pairing in meiosis. Nat. Genet. 21:123-127.
Edvardsen RB, Wallerman O, Furmanek T, Kleppe L, Jern P, Wallberg A, Kjaerner-Semb E, Maehle S, Olausson SK, Sundstrom E, et al. 2022. Heterochiasmy and the establishment of gsdf as a novel sex determining gene in Atlantic halibut. PLoS Genet 18:e1010011.
Eisfeldt J, Vezzi F, Olason P, Nilsson D, Lindstrand A. 2017. TIDDIT, an efficient and comprehensive structural variant caller for massive parallel sequencing data. F1000Res. 6:664.
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 7:eabe8215.
Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. Nature Reviews Genetics 8:689-698.
Feron R, Pan Q, Wen M, Imarazene B, Jouanno E, Anderson J, Herpin A, Journot L, Parrinello H, Klopp C, et al. 2021. RADSex: A computational workflow to study sex determination using restriction site-associated DNA sequencing data. Mol Ecol Resour 21:1715-1731.
Flanagan SP, Jones AG. 2017. Genome-wide selection components analysis in a fish with male pregnancy. Evolution 71:1096-1105.
Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. U. S. A. 117:9451-9457.
Freret-Meurer NV, Vaccani ADC, Cabiro GDS. 2021. Evidence of feminization in seahorses from a tropical estuary. J Fish Biol 99:695-699.
Gammerdinger WJ, Kocher TD. 2018. Unusual Diversity of Sex Chromosomes in African Cichlid Fishes. Genes 9.

Gao F, Ming C, Hu W, Li H. 2016. New Software for the Fast Estimation of Population Recombination Rates (FastEPRR) in the Genomic Era. G3 6:1563-1571.
Guiguen Y, Fostier A, Herpin A. 2018. Sex Determination and Differentiation in Fish. In. Sex Control in Aquaculture. p. 35-63.
He L, Long X, Qi J, Wang Z, Huang Z, Wu S, Zhang X, Luo H, Chen X, Lin J, et al. 2021. Genome and gene evolution of seahorse species revealed by the chromosome-level genome of Hippocampus abdominalis. Mol. Ecol. Resour.
Herpin A, Schartl M. 2015. Plasticity of gene-regulatory networks controlling sex determination: of masters, slaves, usual suspects, newcomers, and usurpators. EMBO Rep. 16:1260-1274.
Hossain A, Saunders GF. 2001. The human sex-determining gene SRY is a direct target of WT1. J. Biol. Chem. 276:16817-16823.
Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, et al. 2013. The zebrafish reference genome sequence and its relationship to the human genome. Nature 496:498-503.
Hubley R, Finn RD, Clements J, Eddy SR, Jones TA, Bao W, Smit AFA, Wheeler TJ. 2016. The Dfam database of repetitive DNA families. Nucleic Acids Res. 44:D81-89.
Ijiri S, Kaneko H, Kobayashi T, Wang D-S, Sakai F, Paul-Prasanth B, Nakamura M, Nagahama Y. 2008. Sexual dimorphic expression of genes in gonads during early differentiation of a teleost fish, the Nile tilapia Oreochromis niloticus. Biol. Reprod. 78:333-341.
Jeffries DL, Lavanchy G, Sermier R, Sredl MJ, Miura I, Borzée A, Barrow LN, Canestrelli D, Crochet P-A, Dufresnes C, et al. 2018. A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. Nat. Commun. 9:4088.
Jones AG, Kvarnemo C, Moore GI, Simmons LW, Avise JC. 1998. Microsatellite evidence for monogamy and sex-biased recombination in the Western Australian seahorse Hippocampus angustus. Mol. Ecol. 7:1497-1505.
Jordan CY, Charlesworth D. 2012. The potential for sexually antagonistic polymorphism in different genome regions. Evolution 66:505-516.
Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S-Y, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. 42:348-354.
Killick R, Eckley I. 2014. changepoint:an R package for changepoint analysis. J. Stat. Softw. 58:19.
Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat. Methods 12:357-360.
Kottler VA, Schartl M. 2018. The Colorful Sex Chromosomes of Teleost Fish. Genes 9.
Kratochvíl L, Gamble T, Rovatsos M. 2021. Sex chromosome evolution among amniotes: is the origin of sex chromosomes non-random? Philos. Trans. R. Soc. Lond. B Biol. Sci. 376:20200108.
Kryuchkova-Mostacci N, Robinson-Rechavi M. 2017. A benchmark of gene expression tissue-specificity metrics. Briefings in bioinformatics 18:205-214.
Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639-1645.
Kumar S, Stecher G, Suleski M, Hedges SB. 2017. TimeTree: a resource for timelines, timetrees, and divergence times. Molecular biology and evolution 34:1812-1819.
Kunstner A, Hoffmann M, Fraser BA, Kottler VA, Sharma E, Weigel D, Dreyer C. 2016. The Genome of the Trinidadian Guppy, Poecilia reticulata, and Variation in the Guanapo Population. PLoS One 11:e0169087.
Lahn BT, Page DC. 1999. Four evolutionary strata on the human X chromosome. Science 286:964-967.
Lemaitre C, Braga MD, Gautier C, Sagot MF, Tannier E, Marais GA. 2009. Footprints of inversions at present and past pseudoautosomal boundaries in human sex chromosomes. Genome Biol Evol 1:56-66.
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. 12:1094.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-1760.
Liang D, Fan Z, Zou Y, Tan X, Wu Z, Jiao S, Li J, Zhang P, You F. 2018. Characteristics of Cyp11a during Gonad Differentiation of the Olive Flounder Paralichthys olivaceus. International Journal of Molecular Sciences 19:2641.
Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923-930.
Lin Q, Fan S, Zhang Y, Xu M, Zhang H, Yang Y, Lee AP, Woltering JM, Ravi V, Gunter HM, et al. 2016. The seahorse genome and the evolution of its specialized morphology. Nature 540:395-399.
Lin Q, Qiu Y, Gu R, Xu M, Li J, Bian C, Zhang H, Qin G, Zhang Y, Luo W, et al. 2017. Draft genome of the lined seahorse, Hippocampus erectus. Gigascience 6:1-6.

Liu X, Zhang D, Lin T, Zhou L. 2020. Chromosome preparation and karyotype of the lined seahorse (Hippocampus erectus). JOURNAL OF FISHERIES OF CHINA 44:907-914.
Mank JE, Avise JC. 2006. Phylogenetic conservation of chromosome numbers in Actinopterygiian fishes. Genetica 127:321-327.
Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: A fast and versatile genome alignment system. PLoS Comput. Biol. 14:e1005944.
Matoulek D, Borůvková V, Ocalewicz K, Symonová R. 2020. GC and Repeats Profiling along Chromosomes—The Future of Fish Compositional Cytogenomics. Genes 12:50.
Melters DP, Bradnam KR, Young HA, Telis N, May MR, Ruby JG, Sebra R, Peluso P, Eid J, Rank D. 2013. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. Genome biology 14:1-20.
Miura I, Ohtani H, Ogata M. 2012. Independent degeneration of W and Y sex chromosomes in frog Rana rugosa. Chromosome Res 20:47-55.
Morohashi K, Zanger UM, Honda S, Hara M, Waterman MR, Omura T. 1993. Activation of CYP11A and CYP11B gene promoters by the steroidogenic cell-specific transcription factor, Ad4BP. Mol. Endocrinol. 7:1196-1204.
Nei M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences 70:3321-3323.
Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418-426.
O'Meally D, Ezaz T, Georges A, Sarre SD, Graves JAM. 2012. Are some chromosomes particularly good at sex? Insights from amniotes. Chromosome Res. 20:7-19.
Ohno S. 1974. Animal cytogenetics, volume 4: Chordata 1. Protochordata, Cyclostomata and Pisces. Gebrüder-Borntraeger, Berlin.
Olito C, Ponnikas S, Hansson B, Abbott JK. 2022. Consequences of partially recessive deleterious genetic variation for the evolution of inversions suppressing recombination between sex chromosomes. Evolution 76:1320-1330.
Paczolt KA, Jones AG. 2010. Post-copulatory sexual selection and sexual conflict in the evolution of male pregnancy. Nature 464:401-404.
Palmer DH, Rogers TF, Dean R, Wright AE. 2019. How to identify sex chromosomes and their turnover. Mol. Ecol. 28:4709-4724.
Pan Q, Anderson J, Bertho S, Herpin A, Wilson C, Postlethwait JH, Schartl M, Guiguen Y. 2016. Vertebrate sex-determining genes play musical chairs. C R Biol 339:258-262.
Pan Q, Feron R, Jouanno E, Darras H, Herpin A, Koop B, Rondeau E, Goetz FW, Larson WA, Bernatchez L, et al. 2021. The rise and fall of the ancient northern pike master sex-determining gene. Elife 10.
Pan Q, Kay T, Depincé A, Adolfi M, Schartl M, Guiguen Y, Herpin A. 2021. Evolution of master sex determiners: TGF-β signalling pathways at regulatory crossroads. Philosophical Transactions of the Royal Society B 376:20200091.
Peichel CL, McCann SR, Ross JA, Naftaly AFS, Urton JR, Cech JN, Grimwood J, Schmutz J, Myers RM, Kingsley DM, et al. 2020. Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. Genome Biol 21:177.
Pettersson ME, Rochus CM, Han F, Chen J, Hill J, Wallerman O, Fan G, Hong X, Xu Q, Zhang H, et al. 2019. A chromosome-level assembly of the Atlantic herring genome-detection of a supergene and other signals of selection. Genome Res 29:1919-1928.
Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559-575.
Qiu S, Bergero R, Charlesworth D. 2013. Testing for the footprint of sexually antagonistic polymorphisms in the pseudoautosomal region of a plant sex chromosome pair. Genetics 194:663-672.
Qu M, Liu Y, Zhang Y, Wan S, Ravi V, Qin G, Jiang H, Wang X, Zhang H, Zhang B, et al. 2021. Seadragon genome analysis provides insights into its phenotype and sex determination locus. Sci Adv 7.
Rice WR. 1987. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. Evolution 41:911-914.
Rifkin JL, Beaudry FEG, Humphries Z, Choudhury BI, Barrett SCH, Wright SI. 2021. Widespread recombination suppression facilitates plant sex chromosome evolution. Mol Biol Evol 38:1018-1030.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.
Rochette NC, Rivera-Colón AG, others. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Mol. Ecol.
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. U. S. A. 117:9431-9439.
Sardell JM, Cheng C, Dagilis AJ, Ishikawa A, Kitano J, Peichel CL, Kirkpatrick M. 2018. Sex Differences in Recombination in Sticklebacks. G3 (Bethesda) 8:1971-1983.
Sardell JM, Josephson MP, Dalziel AC, Peichel CL, Kirkpatrick M. 2021. Heterogeneous Histories of Recombination Suppression on Stickleback Sex Chromosomes. Mol Biol Evol 38:4403-4418.
Sardell JM, Kirkpatrick M. 2020. Sex differences in the recombination landscape. Am Nat 195:361-379.
Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, Heard E, Dekker J, Barillot E. 2015. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. 16:259.
Shao H, Ganesamoorthy D, Duarte T, Cao MD, Hoggart CJ, Coin LJM. 2018. npInv: accurate detection and genotyping of inversions using long read sub-alignment. BMC Bioinformatics 19:261.
Sharma A, Heinze SD, Wu Y, Kohlbrenner T, Morilla I, Brunner C, Wimmer EA, van de Zande L, Robinson MD, Beukeboom LW, et al. 2017. Male sex in houseflies is determined by , a paralog of the generic splice factor gene. Science 356:642-645.
Shen Z-G, Wang H-P. 2018. Environmental Sex Determination and Sex Differentiation in Teleosts – How Sex Is Established. In. Sex Control in Aquaculture. p. 85-115.
Shen ZG, Wang HP. 2014. Molecular players involved in temperature-dependent sex determination and sex differentiation in Teleost fish. Genet Sel Evol 46:26.
Small CM, Bassham S, Catchen J, Amores A, Fuiten AM, Brown RS, Jones AG, Cresko WA. 2016. The genome of the Gulf pipefish enables understanding of evolutionary innovations. Genome Biol. 17:258.
Stiller J, Short G, Hamilton H, Saarman N, Longo S, Wainwright P, Rouse GW, Simison WB. 2022. Phylogenomic analysis of Syngnathidae reveals novel relationships, origins of endemic diversity and variable diversification rates. BMC Biol 20:75.
Stöck M, Horn A, Grossen C, Lindtke D, Sermier R, Betto-Colliard C, Dufresnes C, Bonjour E, Dumas Z, Luquet E, et al. 2011. Ever-young sex chromosomes in European tree frogs. PLoS Biol. 9:e1001062.
Taboada X, Robledo D, Bouza C, Piferrer F, Viñas AM, Martínez P. 2018. Reproduction and Sex Control in Turbot. In. Sex Control in Aquaculture. p. 565-582.
Tao W, Xu L, Zhao L, Zhu Z, Wu X, Min Q, Wang D, Zhou Q. 2021. High-quality chromosome-level genomes of two tilapia species reveal their evolution of repeat sequences and sex chromosomes. Mol. Ecol. Resour. 21:543-560.
Tarailo-Graovac M, Chen N. 2009. Using RepeatMasker to identify repetitive elements in genomic sequences. Curr. Protoc. Bioinformatics Chapter 4:Unit 4.10.
Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. Bioinformatics 31:2032-2034.
Úbeda F, Patten MM, Wild G. 2015. On the origin of sex chromosomes from meiotic drive. Proc. Biol. Sci. 282:20141932.
van Doorn GS, Kirkpatrick M. 2010. Transitions between male and female heterogamety caused by sex-antagonistic selection. Genetics 186:629-645.
van Doorn GS, Kirkpatrick M. 2007. Turnover of sex chromosomes induced by sexual conflict. Nature 449:909-912.
Veller C, Muralidhar P, Constable GWA, Nowak MA. 2017. Drift-Induced Selection Between Male and Female Heterogamety. Genetics 207:711-727.
Vicoso B. 2019. Molecular and evolutionary dynamics of animal sex-chromosome turnover. Nat Ecol Evol 3:1632-1641.
Vincent A, Ahnesjö I, Berglund A, Rosenqvist G. 1992. Pipefishes and seahorses: Are they all sex role reversed? Trends Ecol. Evol. 7:237-241.
Vitturi R, Catalano E, Colomba M, Montagnino L, Pellerito L. 1995. Karyotype analysis of Aphanius fasciatus (Pisces, Cyprinodontiformes): Ag-NORs and C-band polymorphisms in four populations from Sicily. Biol Zent Bl 114.
Vitturi R, Libertini A, Campolmi M, Calderazzo F, Mazzola A. 1998. Conventional karyotype, nucleolar organizer regions and genome size in five Mediterranean species of Syngnathidae (Pisces, Syngnathiformes). J. Fish Biol. 52:677-687.
Volff J-N, Schartl M. 2001. Variability of genetic sex determination in poeciliid fishes. Genetica 111:101-110.

Wang J, Na JK, Yu Q, Gschwend AR, Han J, Zeng F, Aryal R, VanBuren R, Murray JE, Zhang W, et al. 2012. Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. Proc Natl Acad Sci U S A 109:13710-13715.

Weng M-P, Liao B-Y. 2017. modPhEA: model organism Phenotype Enrichment Analysis of eukaryotic gene sets. Bioinformatics 33:3505-3507.

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 32:3114-3131.

Wolff J, Bhardwaj V, Nothjunge S, Richard G, Renschler G, Gilsbach R, Manke T, Backofen R, Ramírez F, Grüning BA. 2018. Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. Nucleic Acids Res. 46:W11-W16.

Wolff J, Rabbani L, Gilsbach R, Richard G, Manke T, Backofen R, Grüning BA. 2020. Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and single-cell Hi-C data analysis, quality control and visualization. Nucleic Acids Res. 48:W177-W184.

Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, Bar-Even A, Horn-Saban S, Safran M, Domany E. 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformatics 21:650-659.

Yoshida K, Kitano J. 2021. Tempo and mode in karyotype evolution revealed by a probabilistic model incorporating both chromosome number and morphology. PLoS genetics 17:e1009502.

Yoshitake K, Ishikawa A, Yonezawa R, Kinoshita S, Kitano J, Asakawa S. 2022. Construction of a chromosome-level Japanese stickleback species genome using ultra-dense linkage analysis with single-cell sperm sequencing. NAR Genom Bioinform 4:lqac026.

Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J. 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinformatics 4:259-263.

Zhou Q, Zhang J, Bachtrog D, An N, Huang Q, Jarvis ED, Gilbert MTP, Zhang G. 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. Science 346:1246338.

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. 10:1523.

Figure Legend

Figure 1. Chromosome evolution and repeat landscapes of seahorse and Gulf pipefish species. **(A)** Phylogenetic tree showing estimated times of appearance (Kumar, et al. 2017) of Syngnathidae species, with the Nile tilapia (*O. niloticus*) as an outgroup. The red dot labels the node corresponding to the time of the previously reported duplication creating the putatively sex determining gene *AMHR2Y* specific to the two outgroup species to seahorses, *P. taeniolatus* and *S. biaculeatus* (see main text). The distributions of sequence divergence from the consensus sequences of the repeat families that have expanded in the two studied seahorse species are also shown. The Y-axis represents the genome-wide percentages of repeat families with different divergence values. DNA transposons hAT-Charlie and TcMar-Tigger, as well as LTR Ngaro families have expanded in *H. abdominalis*, compared to *H. erectus*, *S. scovelli* and *O. niloticus*. The divergence times of these species are labelled on the cladogram at the left of the repeat landscapes. (**B-C)** Chromosome synteny between *H. abdominalis* and two other teleosts: Gulf pipefish (*S. scovelli*) and Nile tilapia (*O. niloticus*). Whole genome alignments suggest that most chromosomes of *S. scovelli* and *O. niloticus* have one-to-one homologous relationships with those of *H. abdominalis* except for Chr11. **(D)** Percentage of candidate centromeric repeats in *H. abdominalis, H. erectus* and *S. scovelli*, shown in the 5kb-sliding window. Candidate centromeric repeats in different species are shown with different colors. A red triangle indicates the putative fusion site of Chr11 in *H. abdominalis*, which corresponds to two homologous chromosomes in the other two teleosts. **(E)** GC (blue) and overall repeat content (yellow) *of H. abdominalis, H. erectus* and *S. scovelli* (calculated in each 50kb-sliding window).

Figure 2. Independent evolution of sex-linked regions of Syngnathid species.
(A) Sex-associated RAD markers in *S. scovelli*. Markers significantly associated with sex are labelled either in blue (when they suggest male heterogamety) or red (female heterogamety) in the outer track of the circos plot. F_{ST} values between males and females are shown in the inner track. Neither suggests a sex chromosome. **(B)** Heatmap showing individual sequencing depth for all markers present in at least 5 *S. scovelli* individuals. Each column corresponds to a marker, and each row corresponds to one male or female individual. Males (blue) and females (red) cluster in separate groups, except for one female and 19 male outliers marked by red arrows. **(C-D)** Sex-linked regions of *H. abdominalis* and *H. erectus.* The figure shows various metrics calculated in 50kb sliding windows across each chromosome, with blue tracks for the -log10 *P*-values of GWAS tests shown (for sex-associated SNPs with $P < 10^{-4}$ in *H. abdominalis* and *H. erectus*), and green tracks for the numbers of sites that are heterozygous in males and homozygous in females after excluding the possible sex-reversed individuals, while purple tracks show the sites that are heterozygous in females and homozygous in males, and pink tracks show the mean F_{ST} value between the sexes. Both the GWAS results and increased F_{ST} values suggest Y-linked regions occupying large parts of the *H. abdominalis* chromosome 6 and the *H. erectus* chromosome 4. **(E-F)** Genotypes of the 100 most strongly associated SNPs with sex in the GWAS analyses of *H. abdominalis* and *H. erectus* (ranked by their *P*-values). Each row corresponds to one male or female individual, and each column corresponds to a SNP within the sex-linked region. In both species, males were mostly heterozygous and females were mostly homozygous at these sites, indicating male heterogamety. Individuals with the inferred genotype of the opposite sex at all these sites are indicated by red arrows. (**G**) Phylogenetic tree and sex chromosome of Syngnathidae and the more distant outgroup Cichlidae species. Sex chromosomes of Syngnathidae and Cichlidae species are colored to indicate XY (blue) and ZW (red) sex-determining systems.

Figure 3 Repeat and putative recombination landscapes at the seahorse sex-linked regions. (**A**) Part of the sex-linked region inferred on the *H. abdominalis* Chr6 (7.03Mb - 22.68Mb). The dark blue track shows the sex-associated SNPs identified by our GWAS; the gray track shows the overall repeat content, with the inferred centromeric region (based on enrichment of the DNA element Zator, see **Figure 1**) labelled by a red triangle, and the candidate sex-determining gene *MSH5* indicated with a red dashed line; the green track shows the numbers of sites per 50kb that are heterozygous in males but homozygous in females, after excluding the possibly sexually reversed individuals; the pink track shows the mean F_{ST} values between males and females in 50kb windows, and the light blue track shows the GC content values; the last track shows the A/B or active/repressive compartment distribution inferred from Hi-C data. (**B**) Part of the sex-linked region of *H. erectus* Chr4 inferred based on a high density of sex-associated SNPs in our GWAS (1.28Mb - 26.1Mb). The centromeric region was inferred from enrichment of the DNA element hAT (red triangle), and the candidate sex-determining genes *WT1* and *CYP11A1* are indicated by red dashed lines. Red dashed vertical lines also label the boundary between the putatively sex-linked regions and the PAR. Horizontal black lines indicate the different mean values of each quantity inferred by change-point analysis. (**C**) GC content of the HeChr4 and HaChr6 sex-linked regions. (**D**) Syntenic regions in HaChr6 of *H. abdominalis* and LG24 and Chr11 of *L. oculatus*, the spotted gar. Homologous sequences located within the sex-linked region of *H. abdominalis* are shown with yellow lines, while gray lines show those in the PAR.

Figure 4 Evolution of sex-related genes associated with the independent sex chromosome evolution

(A) Heatmap showing the transcription patterns of putative sex-determining genes and gonad development genes in different tissues of *H. abdominalis* (blue) and *H. erectus* (yellow). Darker colors in the first row represent later developmental stages. Different intensities of blue indicate the number of days after fertilization (DAF), from 90 (light) to 150 (deep), while yellow intensities indicate later stages, 150 DAF (lightest), 210 DAF, 240 DAF (pregnant) and 240 DAF (post-pregnant, deepest). Gray boxes represent missing orthologs in *H. erectus*. Genes are labelled with different colors: blue for male determining or testis development genes, red for female determining or ovary development genes and yellow for genes with functions in both sexes. Blue triangles represent genes whose gonad transcription patterns are similar to those of their orthologs in the other seahorse or a teleost species, and red triangles represent genes whose gonad transcription patterns differ between the species. The suffix "-1" added to the gene name indicates that it has multiple copies in the genome. **(B)** Enrichment on different chromosomes of tissue-biased genes (relative to male muscle, brain and kidney tissues) and male pouch-biased genes (relative to female abdominal skin). Heatmap showing the scaled log10 P-values of Fisher's tests for chromosomal enrichment. Chromosomes significantly enriched for genes with higher levels in pouch relative to other male tissues are labelled with asterisks (Fisher's exact test, $P<0.05$(*), $P<0.01$(**), $P<0.001$(***)). **(C)** GO enrichment of different categories of genes. Immature male-based: genes that have higher transcription levels in the immature (3 months old) than the mature (5 months old) pouch, and also than in the female abdominal skin at the same stage. Immature female-biased: genes that have a higher transcription level in the immature (3 months old) than the mature pouch (5 months old), but a higher transcription level in the female abdominal skin than the pouch etc. Each node represents one enriched GO term. Nodes of related GO functional terms are shown in similar colors, and the widths of the lines connecting the nodes indicate the similarity between the terms, while the node sizes represent the percentage of input genes belonging to each term. **(D)** Enrichment analysis of homologous mouse genes' mutant phenotypes for different types of *H. abdominalis* genes, with different colors indicating the *P*-values of Bonferroni corrected Fisher's exact tests. IFB: immature female-biased, MFB: mature female-biased, IMB: immature male-biased, MMB: mature male-biased. Immune system related phenotypes were enriched in both immature and mature male-biased genes.

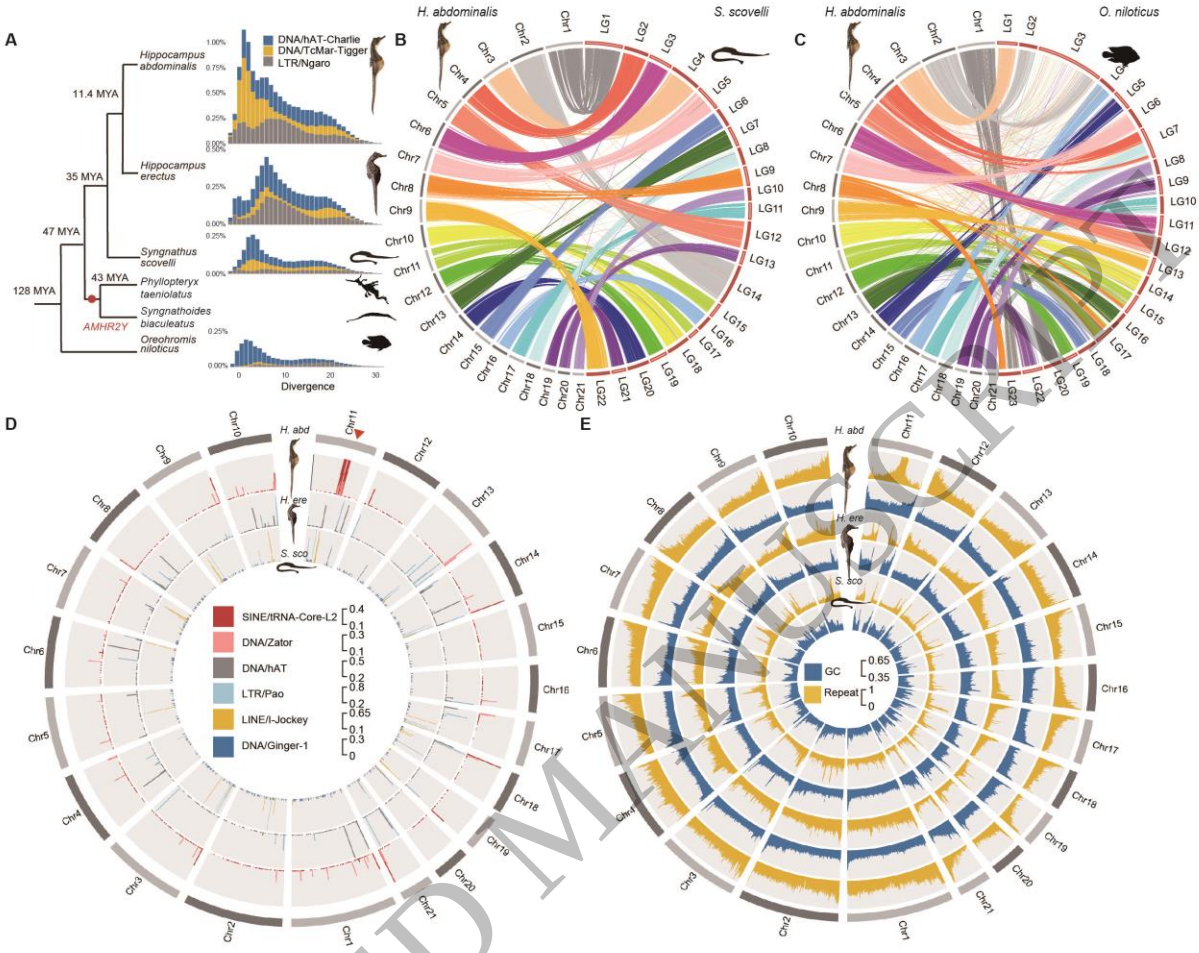

Figure 1
207x163 mm (x DPI)

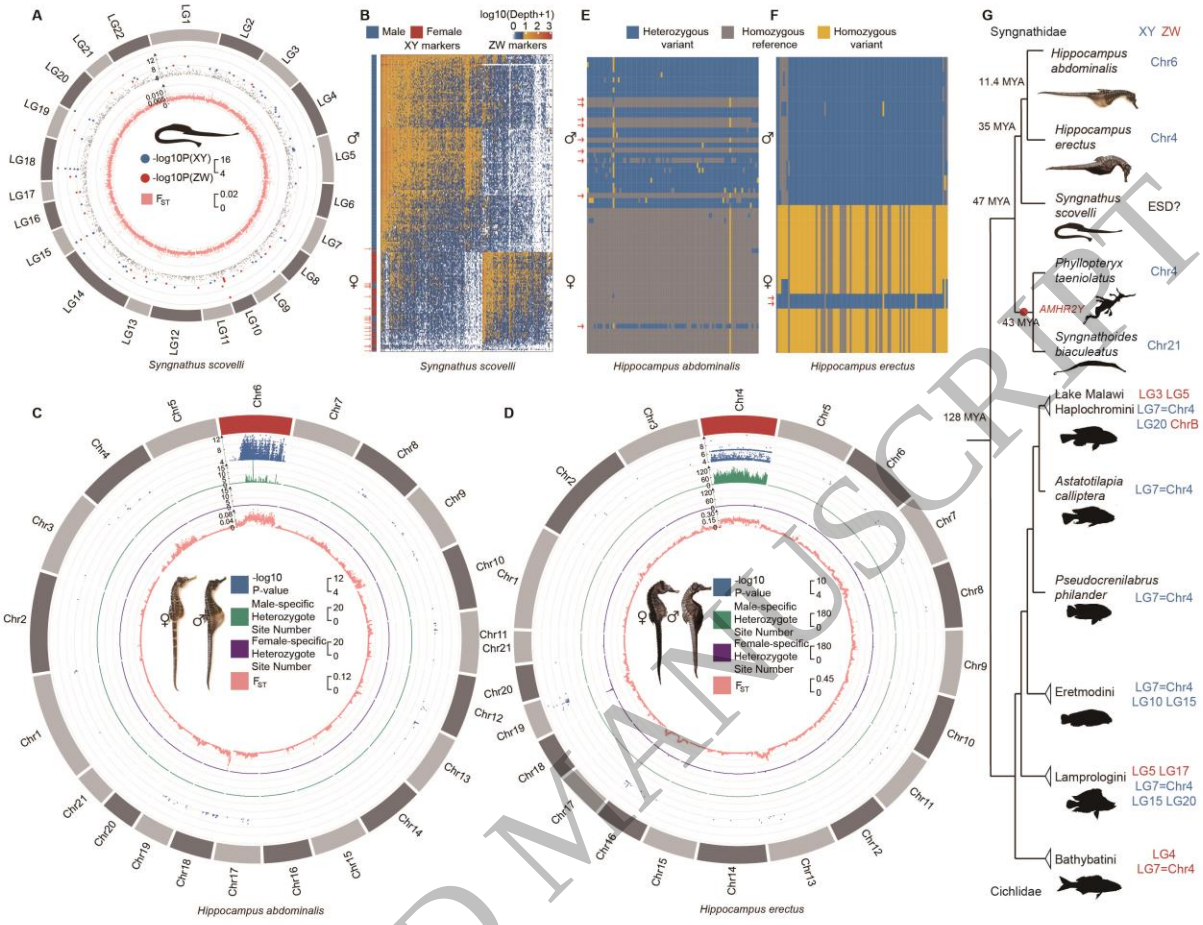

Figure 2
247x188 mm (x DPI)

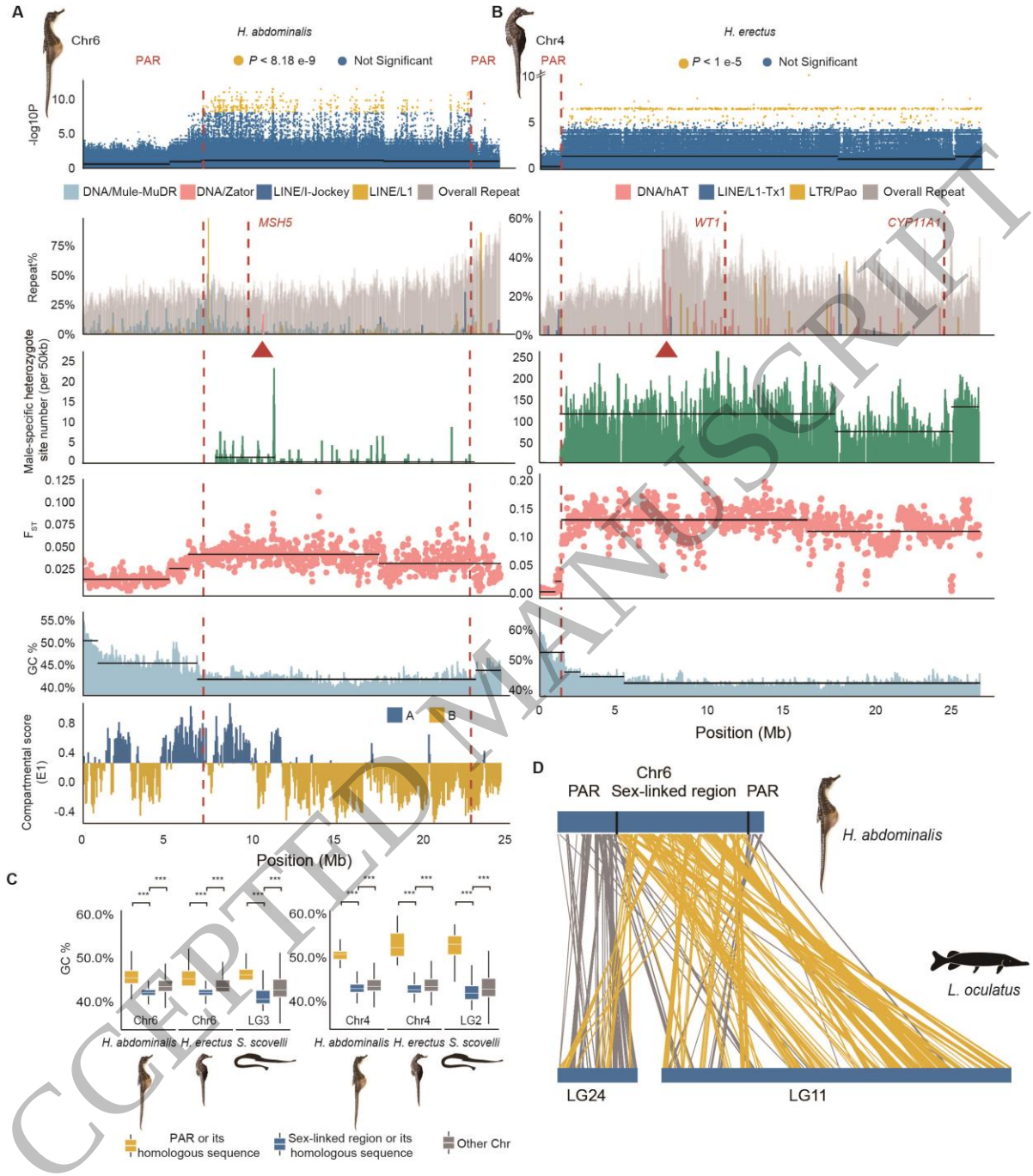

Figure 3
186x210 mm (x DPI)

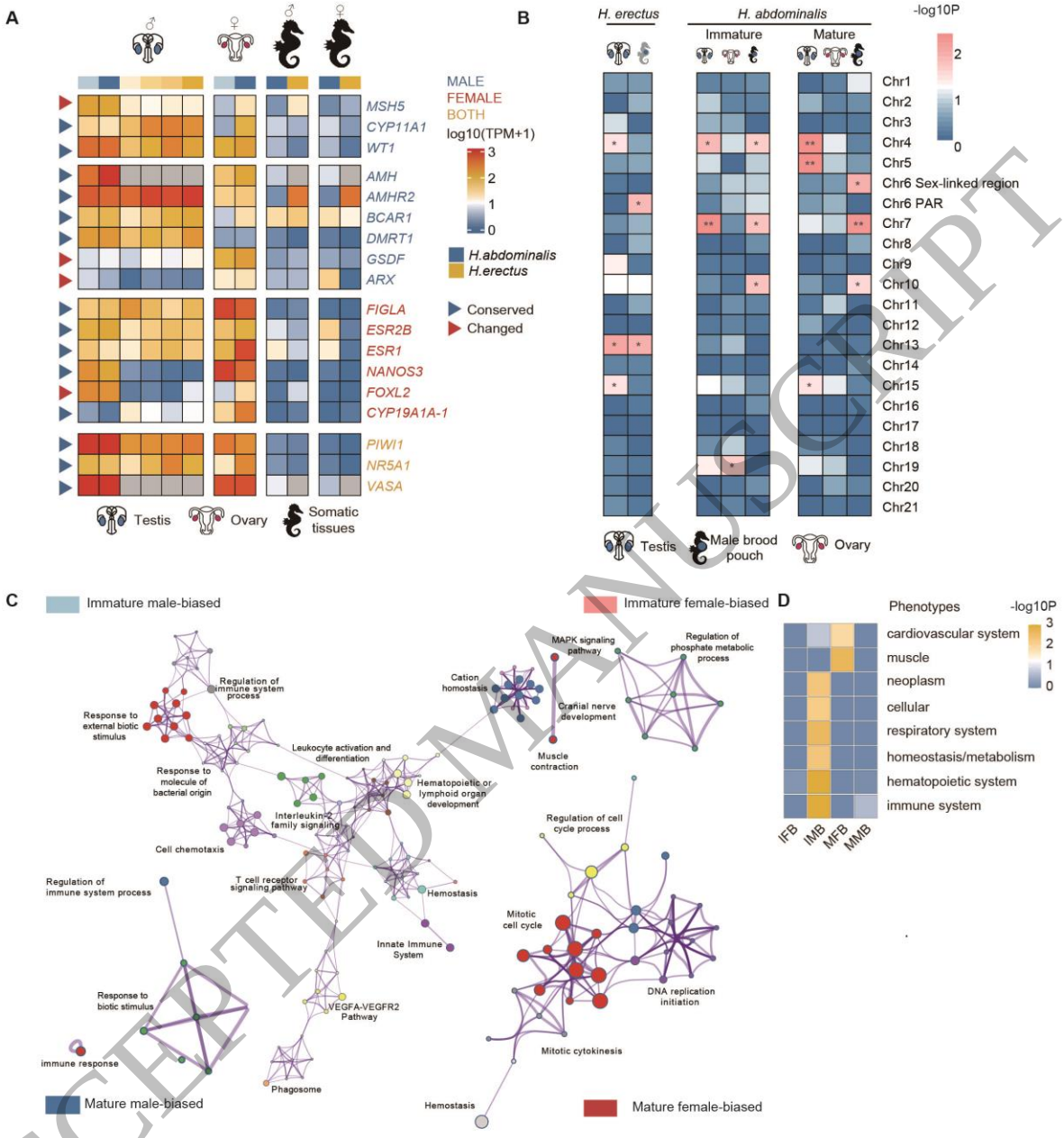

Figure 4
177x184 mm (x DPI)