Chromosome-level assembly reveals a putative Y-autosomal fusion in the sex determination system of the Greenland Halibut (*Reinhardtius hippoglossoides*)

Anne-Laure Ferchaud [ID] ,[1],* Claire Mérot,[1] Eric Normandeau,[1] Jiannis Ragoussis,[2] Charles Babin,[1] Haig Djambazian,[2]
Pierre Bérubé,[2] Céline Audet [ID] ,[3] Margaret Treble [ID] ,[4] Wojciech Walkusz,[4] and Louis Bernatchez [ID] [1]

[1]Département de Biologie, Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC G1V 0A6, Canada,
[2]McGill Genome Centre and Department for Human Genetics, McGill University, Montreal, Quebec, QC H3A 0G1, Canada,
[3]Département de l'Institut des Sciences de la Mer de Rimouski, Université du Québec à Rimouski, Rimouski, QC G5L 3A1, Canada, and
[4]Fisheries and Oceans Canada, Arctic Aquatic Research Division, Winnipeg Department, Freshwater Institute Winnipeg, Manitoba, MB R3T2N6, Canada

*Corresponding author: Département de Biologie, Laboratoire du Dr Louis Bernatchez, Institut de Biologie Intégrative et des Systèmes (IBIS), Pavillon Charles-Eugène
Marchand 1030 rue de la Médecine, Université Laval, Québec, QC G1V 0A6, Canada. Email: annelaureferchaud@gmail.co

Abstract

Despite the commercial importance of Greenland Halibut (*Reinhardtius hippoglossoides*), important gaps still persist in our knowledge of this species, including its reproductive biology and sex determination mechanism. Here, we combined single-molecule sequencing of long reads (Pacific Sciences) with chromatin conformation capture sequencing (Hi-C) data to assemble the first chromosome-level reference genome for this species. The high-quality assembly encompassed more than 598 Megabases (Mb) assigned to 1594 scaffolds (scaffold N50 = 25 Mb) with 96% of its total length distributed among 24 chromosomes. Investigation of the syntenic relationship with other economically important flatfish species revealed a high conservation of synteny blocks among members of this phylogenetic clade. Sex determination analysis revealed that similar to other teleost fishes, flatfishes also exhibit a high level of plasticity and turnover in sex determination mechanisms. A low-coverage whole-genome sequence analysis of 198 individuals revealed that Greenland Halibut possesses a male heterogametic XY system and several putative candidate genes implied in the sex determination of this species. Our study also suggests for the first time in flatfishes that a putative Y-autosomal fusion could be associated with a reduction of recombination typical of the early steps of sex chromosome evolution.

Keywords: long-reads; PacBio; optical mapping; low-coverage whole-genome sequencing; sex determination; Y-autosomal fusion; flatfish

Introduction

The mechanisms underlying sex determination (SD) are known to affect developmental processes, the evolution of genomes, and are responsible for determining sex ratio, a key demographic parameter for population viability and stability (Bull 1983; Uller *et al.* 2007). In recent years, genetic analyses in nonmodel species have revealed a greater degree of sex chromosome diversity and turnover than previously appreciated (Abbott *et al.* 2016). In fact, many case studies in different taxonomic groups, including lizards, fish, amphibians, insects, and plants show frequent changes in the location of SD genes and high rates of turnover of sex chromosomes, suggesting that exceptions are the rules in sex determination (Yano *et al.* 2013; McGrath 2020, for review). Therefore, a more cyclical than linear conceptual framework could be used to explain the evolution of sex chromosomes (Bachtrog 2006; Furman *et al.* 2020). In such a cycle, a new master sex-determining locus arises on an autosome, leading to sex chromosome formation, generally characterized by recombination suppression between the Y and X, or the Z and W

chromosomes. As under a classic sex chromosome evolution model, this situation may progress toward increased divergence between male and female chromosomes, leading to chromosome heteromorphism characterized by genomic degradation, including rearrangements and loss of genetic material (Charlesworth *et al.* 2005; Bachtrog 2013). However, sex chromosomes and sex-determining regions can also undergo turnover with either the evolution of a new sex-determining gene or the transposition of the sex-determining gene to a new location in the genome, leading to another cycle of sex chromosome evolution (Bachtrog 2006). The increasing availability of high-quality genome assemblies in nonmodel taxonomic groups is now providing exciting opportunities to revisit the evolution of genetic SD (Whibley *et al.* 2021).

Fish exhibit a remarkable plasticity in gonad development and a high evolutionary SD turnover, making them a particularly relevant group to investigate the evolutionary dynamics of sex chromosomes (Kitano and Peichel 2012). Fish display highly diverse chromosome SD systems, including models analogous to the

XX/XY (male heterogametic) and ZZ/ZW (female heterogametic) systems of mammals and birds, respectively, but also other systems involving multiple sex chromosomes (Capel 2017). In accordance with this rapid turnover, chromosome heteromorphism is rare in fishes (*e.g.*, Pan *et al.* 2019; Martinez *et al.* 2021), which generally show limited sex chromosome differentiation (Chen *et al.* 2014), characteristic of early SD mechanisms or frequent sex determination transitions than can be identified even between populations of the same species (*e.g.*, Pan *et al.* 2021). Until recently, the DM-domain gene on the Y chromosome (*dmy*) was the only known gene coding for a functional protein in the Y-specific chromosome region (Nanda *et al.* 2002; Kondo *et al.* 2006) and was described as the controlling SD gene in a wide range of invertebrates and vertebrates, including fish. However, thanks to new genomics tools, the genomic features underlying SD have been more intensively studied and several other strong candidates of controlling SD genes have been found in fish, including *amhy*, *gsdf*, *amhr2*, and *sdY* (Guerrero-Estévez and Moreno-Mendoza 2010; Yano *et al.* 2013; Rondeau *et al.* 2016; Hattori *et al.* 2019; Pan *et al.* 2019). In fact, three of these genes (*amhy*, *amhr2*, and *gsdf*) are involved in the transforming growth factor beta (*TGF-β*) signaling pathway, while the *sdY*, *dmrt1*, and *dmy* genes code for transcription factors (Rastetter *et al.* 2015). Unsuspected SD mechanisms are also being unraveled. For example, a recent study in Channel Catfish (*Ictalurus punctatus*) identified the breast cancer anti-resistance 1 (BCAR1) gene as the controlling SD candidate (Bao *et al.* 2019). Insights into SD genomics go beyond controlling genes in fish, as several studies have documented, and include mixed determination involving both genetic and environmental factors, polygenic SD systems, extensive within-species variation in SD system, and complex rearrangements of sex chromosomes favoring species diversification (Baroiller *et al.* 2009; Martinez *et al.* 2014; Capel 2017). With such dynamism and variability in fish sex chromosome evolution, we need improved genomic resources in nonmodel species, including chromosome-scale reference genomes and whole-genome sequences from both sexes, to improve our understanding of the subject.

In exploited species, improved genomic resources and the understanding of the genetic basis for SD is also relevant to improve fisheries management in the face of increasing demand for seafood worldwide and to secure the economies and livelihoods that depend on them (Bernatchez *et al.* 2017). Genomic approaches allow investigating genome-wide diversity and accurately identifying genetically distinct fish stocks (Ovenden *et al.* 2013). Moreover, accounting for sex-differentiation and sex-linked loci is necessary to avoid unexpected biases due to sex-associated markers (Benestan *et al.* 2017) and to consider variation in sex-ratio in management practices. This is particularly true for species harboring a size-related sexual dimorphism, like flatfish species, that generally favors larger females. Example species are the European Plaice (*Pleuronectes platessa*; Rljnsdorp and Ibelings 1989), the Turbot (*Scophtalmus maximus*; Imsland *et al.* 1997), and the Atlantic Halibut (*Hippoglossus hippoglossus*; Hagen *et al.* 2006). Sex control identification is of high importance for both flatfish production in aquaculture and sustainable fisheries. The selection of large individuals is a well-established concept in fisheries management that increases yield per recruit. However, in the case of sexual dimorphism, fisheries targeting large individuals may reduce the stock's reproductive potential, produce biases in the assessment of the number of reproductive females, and feed erroneous data to fisheries management decision-makers (Keyl *et al.* 2015). Overall, both applied and fundamental evolutionary biology thus benefit from moving beyond genetic markers toward

contiguous reference genomes annotated for functional elements and documented for SD.

Greenland Halibut (*Reinhardtius hippoglossoides*) is one of the main exploited demersal fish in Eastern Canada (Gulf of St. Lawrence, Newfoundland, and the Arctic), representing around 4000 tons of landings per year across the two last decades (DFO 2018). However, the management of Greenland Halibut fisheries would benefit from improved genomic resources. Indeed, no karyotype, linkage map nor genome assembly has been developed for this species so far. Given that juvenile growth potential and size are sexually dimorphic in this species (Ghinter *et al.* 2019), it is valuable to better understand the sex-determining mechanisms to control sex imbalance in catches. Moreover, the recent genome assemblies of other flatfish species (Figure 1, Table 1; Chen *et al.* 2014; Figueras *et al.* 2016; Robledo *et al.* 2017; Shao *et al.* 2017; Einfeldt *et al.* 2021) together with studies on sex-determining loci in this taxon (Chen *et al.* 2014; Drinan et al. 2018; Liang *et al.* 2018; Einfeldt *et al.* 2021) provide an opportunity to investigate the evolutionary turnover of sex chromosomes and SD systems. Flatfishes studied so far are characterized by a variable heterogametic system and different sex-determining genes. For instance, a similar female heterogametic system (ZZ/ZW) is found both in the Tongue Sole (*Cynoglossus semilaevis*) and the Turbot but the candidate SD drivers differ, being *sox2* in the latter (Martinez *et al.* 2021; Table 1), and *dmrt1* in the former species (Chen *et al.* 2014). Japanese Flounder, *Paralichthys olivaceus* (also called the Olive Flounder) has an XX/XY system (Liang *et al.* 2018) but no sex determining gene has been found (Shao *et al.* 2017) and the environment is known to alter sex differentiation in this species (Ospina-Alvarez and Piferrer 2008). Within the righteye flounder family (*Pleuronectidae*), the *Hippoglossinae* subfamiliy harbors variable genetic SD systems. The Pacific Halibut (*Hippoglossus stenolepis*) has a ZZ/ZW system and several sex-linked loci aligned in three different linkage groups (Drinan et al. 2018) while its sister-species, the Atlantic Halibut (*H. hippoglossus*), has a XX/XY system and a putative SD factor called *gsdf* on chromosome 12 (Palaiokostas *et al.* 2013; Einfeldt *et al.* 2021).

In this study, we used an array of genomic technologies to provide high-quality genomic resources and to investigate the genetic basis of sex determination in Greenland Halibut. We produced the first chromosome-level genome assembly for this species by combining single-molecule sequencing of long reads (Pacific Sciences) with chromatin conformation capture sequencing (Hi-C) data. Furthermore, we also identified and characterized the molecular SD system by mapping whole genome sequencing data of 198 individuals onto the assembly to screen for genetic markers that diverged between sexes. Finally, in order to provide an estimate of the extent of rearrangements between species, we analyzed our results in a comparative genomic framework by assessing synteny (*i.e.*, blocks of genes found in the same order in the genome) with other species of flatfish and by targeting candidate sex-determining genes in this group.

Methods
Reference genome assembly
We built a reference genome assembly for the Greenland Halibut in two steps. First, we created a draft assembly using the PacBio long-reads sequencing approach (Pacific Biosciences, Menlo Park, CA, USA) and we then applied Dovetail™ Hi-C (Lieberman-Aiden *et al.* 2009), from Dovetail Genomics to increase contiguity, break up mis-joins, and orient and join scaffolds into chromosomes.
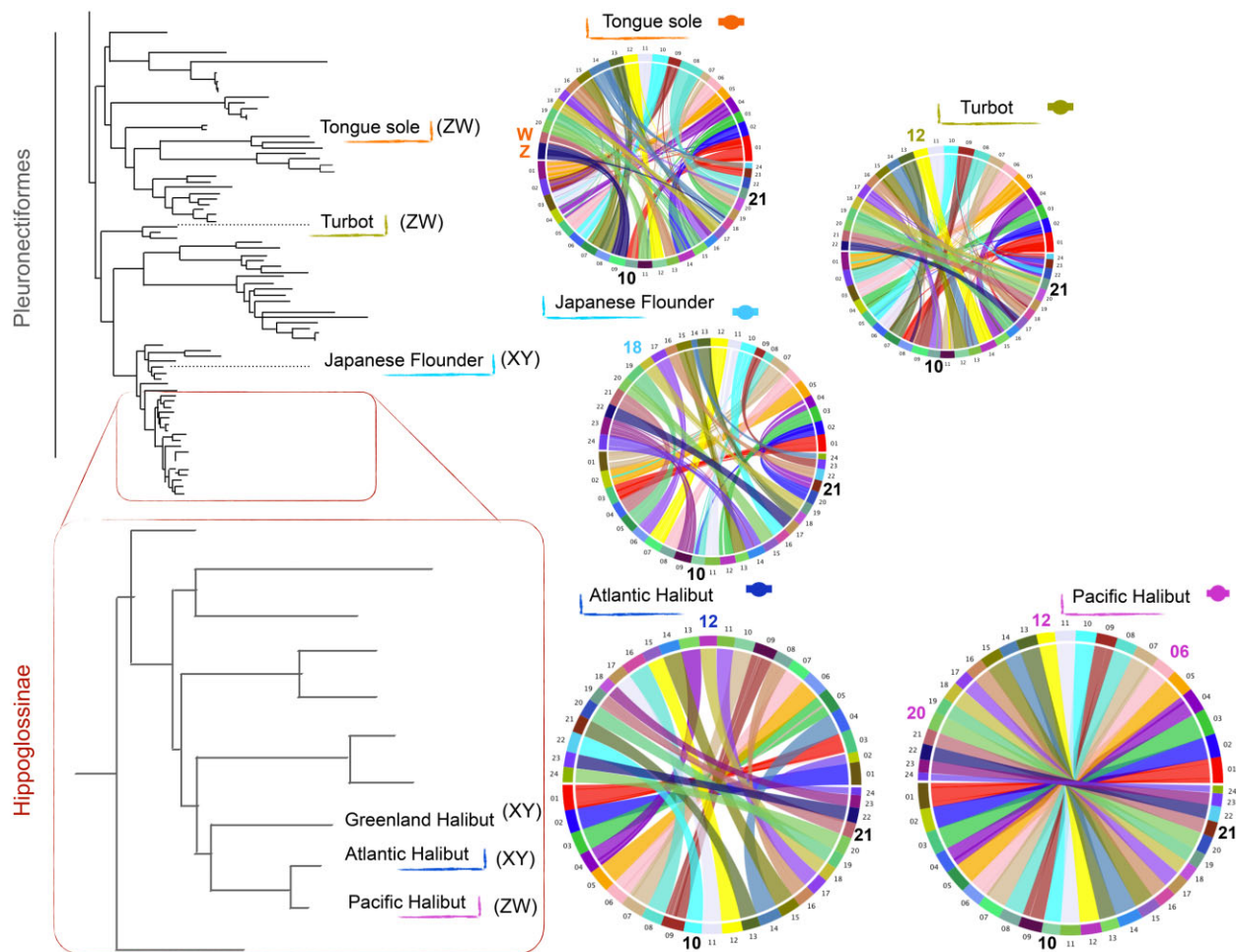
Figure 1 Phylogenetic relationships and syntenic distances between important flatfish species in fisheries and aquaculture. Phylogenetic trees were retraced according to previous studies, Betancur-R and Orti (2014) for the top one and Vinnikov *et al.* (2018) for the zoom into the sub-family *Hippoglossinae* framed in red. Only the names of relevant species compared in our studies are reported as well as the SD system, either "XY" for a male heterogametic system (XX/XY) or "ZW" for a female heterogametic system (ZZ/ZW). The XY system for Greenland Halibut has been assessed in the present study. Circos plots showing synteny between the Greenland Halibut (displayed in low half circle) and the Tongue Sole, Turbot, Japenese Flounder, Atlantic, and Pacific Halibut (displayed in upper half circle). Sex chrosomosomes or the chromosomes in which sex-related markers have been described in the literature have been increased in front size and adequately colored according to the relative species. See Table 1 and its caption for references about other genome assemblies and SD systems in related flatfish species. The same color code as the one displayed in Supplementary Figure S2 is used.

Sampling

The individual whose DNA was used to assemble the reference was a juvenile female caught in the St. Lawrence River Estuary (48°39′11″N, 68°28′37″W) with a Commando-type bottom trawl aboard the CCGS Leim [Fisheries and Oceans Canada (DFO) survey] at the end of May 2016 and kept alive in controlled environmental conditions at the Maurice Lamontagne Institute (IML), Department of Fisheries and Oceans Canada (Ghinter *et al.* 2019). The female was anesthetized for 5 min in a 0.18 g L^{-1} MS-222 solution (tricaine methane sulfonate; Sigma - Aldrich Co, MI, USA). Blood was then sampled from the caudal artery using a 23-gauge needle in a 1 ml TB syringe (Becton, Dickinson & Co., NJ, USA), both of which were previously heparinized in a 100 U ml^{-1} heparin solution. Blood was immediately flash-frozen in liquid nitrogen and then stored at −80°C until DNA extraction. Multiple tissues including liver, muscle, heart, brain, eye, gonads, gall bladder, and kidney were then excised from the fish, flash-frozen, and stored at −80°C until further transcriptomic analyses.

DNA extraction, sequencing, and assembly

High molecular weight DNA was extracted from flash-frozen blood using the Qiagen MagAttract kit, (Valencia, CA, USA). Then a library was created using the SMRTbell Template Prep kit 1.0 (Pacific Biosciences, Menlo Park, CA, USA), which incorporates size selection and DNA damage repair steps. We obtained fragments of >10 kb using size selection with a Blue Pippin instrument (Sage Science, Beverly, MA, USA) according to the manufacturer's recommended protocol for 20 kb template preparation. Five micrograms of concentrated DNA was then used as input for the library preparation reaction. Library quality and quantity were assessed using the Pippin Pulse field inversion gel electrophoresis system (Sage Science), as well as with the dsDNA Broad Range Assay kit and Qubit Fluorometer (Thermo Fisher). Single-Molecule Real Time sequencing was performed on the Pacific Biosciences RS II instrument at McGill University using an on-plate concentration of 125 pM, P6-C4 sequencing chemistry, with magnetic bead loading and 360-min movies. A total of 14 SMRT Cells were run, generating a mean sequence coverage of

Table 1 Statistics of genome assemblies of the important flatfish species in fisheries and aquaculture

	Greenland Halibut	Pacific Halibut	Atlantic Halibut	Japanese flounder	Turbot	Tongue sole
Scientific name	Reinhardtius hippoglossoides	Hippoglossus stenolepis	Hippoglossus hippoglossus	Paralichthys olivaceus	Scophthalmus maximus	Cynoglossus semilaevis
Total sequence length	598,525,902	594,145,200	596,792,615	545,775,252	524,979,463	470,199,494
Number of scaffolds	1,594	120	57	9,525	22	31,181
Max scaffolds length	31,982,681	32,413,155	30,753,830	27,810,759	31,901,587	34,528,841
Scaffold N50	25,015,410	24,985,257	26,312,018	3,817,360	24,811,384	509,861
Scaffold L50	11	11	11	40	10	289
Number of chromosomes	24	24	24	24	22	21
GenBank accession number	GCA_006182925.3	GCA_013339905.1	GCA_009819705.1	GCA_001904815.2	GCA_003186165.1	GCA_000523025.1
Reference	This study	International Pacific Halibut Commission	Einfeldt et al. (2021)	Shao et al. (2017)	Martinez et al. (2021)	Chen et al. (2014)
Name of the assembly	UL_REHI_2.0	IPHC_HiSten_1.0	fHipHip1	ParOli_1.1	ASM318616v1	Cse_v1.0
Complete BUSCO (%)	98.04	na	96.90	na	na	na
Repeat elements (%)	8.20	7.28	7.14	4.92	6.35	6.08
Genetic sex-determinism	XX/XY	ZZ/ZW	XX/XY	XX/XY	ZZ/ZW	ZZ/ZW
SD genes or putative SD genes	sox2, sox9a2, gdf6 (1)	no gene has been specifically identified (2)	gsdf (3)	Cyp11a (4)	sox2 (5)	dmrt1 (6)
Homolog of sex chromosome in Greenland Halibut	—	GH06, GH12, G20	GH18	GH19	GH14	GH09

na, data not available.
(1) This study; (2) Drinan et al. (2018); (3) Einfeldt et al. (2021); (4) Liang et al. (2018); (5) Martinez et al. (2021); (6) Chen et al. (2014).

~70-fold, with half of the sequences contained in reads longer than 10.6 kb. An initial assembly was carried out on the PacBio long reads using Wtdbg v2.5 (-x sq gg 0.7 g; Ruan and Li 2020), resulting in 2037 contigs (scaffold N50 = 3.17 Mb), for a total assembly size of 598.5 Mb. One Hi-C proximity ligation library was generated from a flash-frozen piece of muscle of the same individual using the Dovetail's™ Hi-C Kit v.1.0. This kit employs the restriction enzyme *Dpn*II which recognizes the sequence 5′/GATC 3′. The long-range library was then sent to Dovetail Genomics to be sequenced and run with the initial PacBio assembly through Dovetail's HiRise™ scaffolding pipeline. The completeness of the final assembly was assessed with BUSCO v 4.0.5 (Simão et al. 2015), the eukaryota_odb10 database (created in October 2020) and the -m genome option. The occurrence of repeat elements in our assembly as well as in Atlantic and Pacific Halibut was assessed using RepeatMasker v.4.0.6 (Smit et al. 2013–2015).

Synteny with other flatfish species

Synteny blocks between the genomes of Greenland Halibut and other flatfish species for which a reference genome is available (Tongue Sole, Turbot, and Japanese Flounder as well as Atlantic and Pacific Halibut) were computed using SyMAP v4.2 (Soderlund et al. 2011). Genomic sequences were first aligned using promer/ MUMmer (Kurtz et al. 2004). Raw anchors resulting from MUMmer were clustered into (putative) gene anchors, filtered using a reciprocal top-2 filter and used as input to the synteny algorithm (Soderlund et al. 2006). The algorithm constructs maximal-scoring anchor chains based on a given gap penalty. It also searches a range of gap penalties to generate the longest chains subject to several quality criteria, which are based on the Pearson correlation coefficient applied to the anchors in the chain as well as the anchors in its bounding box. The chains are not required to be entirely colinear and may incorporate local inversions relative to the overall chain orientation. Results stemming from the synteny analysis were visualized using the chromosome explorer and the circle view options in SyMAP v4.2. Finally, we named scaffolds by mapping our assembly to the Pacific Halibut assembly (IPHC_HiSten_1.0, GenBank accession: GCA_013339905.1).

Transcriptome assembly and annotation of the assembly

In order to annotate the reference assembly, total RNA from eight different tissues (liver, muscle, heart, brain, eye, gonads, gall bladder, and kidney) from the same individual was extracted using the RNeasy Mini kit from Qiagen then treated using DNase I, Amplification Grade (1 unit per µl; Invitrogen, Carlsbad, CA, USA) following the manufacturer's protocol. RNA concentration of each tissue was quantified with Quant-iT™ RiboGreen® RNA Assay Kit (Invitrogen) and fragment size distribution was estimated with an Agilent BioAnalyzer before being shipped to the Génome Québec Centre d'Expertise et de Services (Montréal, Canada) to prepare libraries and perform sequencing. One RNA-seq library was constructed for each tissue using the NEBNext® Ultra™ II Directional RNA Library Prep Kit (Illumina, San Diego, CA, USA) and sequenced with a 2 × 100bp (paired-end) read module using the 6000, generating 50M reads.

For each of the eight paired-end tissue-specific RNAseq datasets, raw reads were trimmed with trimmomatic (v0.36, ILLUMINACLIP: univec.fasta: 2:20:7, LEADING: 20, TRAILING: 20, SLIDINGWINDOW: 30:30, MINLEN: 80). The tissue-specific paired-end datasets were then each assembled separately using Trinity (v2.1.1, –seqType fq, –min_contig_length 200) to produce 8 assembled transcriptomes that contained between 68,481 and 276,168, with a mean of 131,038, assembled expressed sequences.

The 1,048,307 assembled expressed sequences were used to annotate the assembled genome using the GAWN pipeline v0.3.4 (https://github.com/enormandeau/gawn). The annotations were then simplified by using the simplify_genome_annotation_table.py script (included in the GAWN pipeline), which removes consecutive almost-identical annotations from the genome annotation table.

Genomic sex determination

Sampling of individuals, DNA extraction, libraries, and sequencing

A total of 432 Greenland Halibut were sampled aboard the R/V Paamiut, from October to November 2016, as part of the annual multispecies survey conducted by the Department of Fisheries and Oceans in Northwest Atlantic Fisheries Organization Subarea 0 in Davis Strait, Canada. Fish were lethally sampled in order to collect length, weight, phenotypic sex, and maturity information. A fin-clip was taken from each individual for genetic analysis. From those individuals, 99 females and 99 males have been selected for sex determination analysis (see sample selection details in Supplementary Material).

Toward this end, we performed the low coverage whole genome sequencing approach proposed by Therkildsen and Palumbi (2017). Briefly, genomic DNA was extracted from a fin-clip using a salt-extraction protocol (Aljanabi and Martinez 1997) with a RNase A treatment (Qiagen). DNA quality of each extract was evaluated with Nanodrop 2000 (ThermoFisher scientific) and migration on a 1% agarose gel electrophoresis. Following Therkildsen and Palumbi (2017), we removed DNA fragments shorter than 1 kb by treating each extract with Axygen magnetic beads in a 0.4:1 ratio, and eluted the DNA in 10 mM Tris-Cl, pH 8.5. We measured DNA concentrations with the Accuclear ultra high sensitivity dsDNA quantification kit (Biotium) and normalized all samples at a concentration of 5 ng/μl. Then, sample DNA extracts were randomized, distributed in plates (96 -well see details about randomization and number of plates used in Supplementary Material) and re-normalized at 2 ng/μl.

Whole-genome high-quality libraries were prepared for each sample according to the protocol described in (Baym *et al.* 2015; Therkildsen and Palumbi 2017). Briefly, a tagmentation reaction using enzyme from the Nextera DNA sample preparation kit (Illumina), which simultaneously fragments the DNA and incorporates partial adapters, was carried out in a 2.5 μl volume with approximately 2 ng of input DNA. Then, we used a two-step PCR procedure with a total of 12 cycles (8 + 4) to add the remaining Illumina adapter sequence with dual index barcodes and to amplify the libraries. The PCR was conducted with the KAPA Library Amplification Kit and custom primers derived from Nextera XT barcode sets A, B, C, and D (total of 384 possible combinations). Amplification products were purified from primers and size-selected with a two-steps Axygen magnetic beads cleaning protocol, first with a ratio 0.5:1, keeping the supernatant (medium and short DNA fragments), second with a ratio 0.75:1, keeping the beads (medium fragments). Final concentrations of the libraries were quantified with the Accuclear ultra high sensitivity dsDNA quantification kit (Biotium) and fragment size distribution was estimated with an Agilent BioAnalyzer for a subset of 10–20 samples per plate. Finally, libraries were pooled (see details in Supplementary Material) for sequence lanes of paired-end 150 bp reads on an Illumina HiSeq 4000 at the Norwegian Sequencing Center at the University of Oslo. Given our multiplexing (up to 96 individuals per lane) and genome size ($\sim$ 600 Mb), we targeted a low sequencing coverage around 1.26 X.

Sequencing filtering and processing

Raw reads were trimmed, filtered for quality, mapped to the reference genome (obtained above), cleaned for duplicate reads and mapping quality and then re-aligned using the pipeline available at https://github.com/enormandeau/wgs_sample_preparation, inspired by Therkildsen and Palumbi (2017) and fully described in Mérot *et al.* (2021). Given the high percentage of the assembly comprised into the anchored 24 chromosomes (>96%, see results section), raw reads were mapped against a reduced version of the genome including only the 24 chromosomes and excluding the unassembled scaffolds. A preliminary analysis, in which individual reads were mapped to the full genome did not change the outcome of the results.

For low coverage whole genome sequencing (lcWGS) data, the recommended practice is to avoid basing downstream analysis on called genotypes (Nielsen *et al.* 2011) and to instead use a probabilistic approach based on Genotype Likelihoods. Several models for computing genotype-likelihood-based on read data have been implemented in the program ANGSD (Korneliussen et al. 2014) and it is currently the most widely used and versatile software package for the analysis of lcWGS (Lou *et al.* 2021). Therefore, ANGSD v0.931 was used for most of our subsequent analyses according to the pipeline documentation available at https://github.com/clairemerot/angsd_pipeline. For all analyses, input reads were filtered to remove reads with a samtools flag above 255 (not primary, failure and duplicate reads, tag -remove_bads = 1), with mapping quality below 30 (-minMapQ 30) and to remove bases with quality below 20 (-minQ 20). We also filtered in order to keep only SNPs covered by at least one read in at least 30% of individuals (-minInd 60) and remove SNPs in putative repeated regions allowing a maximum depth of 3 times the number of individuals (-setMaxDepth 594). Finally, for most of the subsequent analyses (unless mentioned otherwise) we kept SNPs with minor allele frequency above 5%.

First, we ran ANGSD to estimate genotype likelihoods (GL) with the GATK model (-doGlf 2 -GL 2 -doCounts 1), the spectrum of allele frequency (-doSaf 1) and the minor allele frequency (-doMaf 1) options. The major allele was based on the genotype likelihood and was the most frequent allele (-doMajorMinor 1). From this first analysis, we generated (1) a beagle file with GL estimates and (2) a list of variants passing those filters and their respective major and minor alleles that were used for most subsequent analysis. The R program (R Core Team 2020) was employed for graphic output in subsequent analyses, *via* the package ggplot (Wickham 2016).

Genome-wide coverage difference between sexes

Total depth along the entire genome was calculated using the function -doDepth 1 in ANGSD, excluding high coverage (-setMaxDepth 594 and -maxDepth 1000) and low coverage regions (-minInd 60), along with the following arguments for depth: -dumpCounts 2 -doCounts 1. Depth was then averaged across sliding windows of 25 Kb with 5 Kb steps using the winScan function of the R package windowscanr (Tavares 2021). The sliding approach allowed us to visualize individual coverage along the genome, while a ratio of total depth of females above total depth of males was also plotted using ggplot. This ratio is expected to deviate from 1 in genomic regions in which a sex-bias in coverage exists.

Genome wide variation between sexes

Genome-wide variation across samples was explored using PCAangsd (Meisner and Albrechtsen 2018) on the genotype likelihoods. This program extracts a covariance matrix that is then decomposed into the principal component analysis (PCA) with R, using a scaling 2 transformation adding an eigenvalues correction, to obtain the individuals PC scores (Legendre and Legendre 1998).

Identifying genomic regions implied in sex determination

In order to identify the genomic regions involved in sex differentiation, two analyses were performed. First, we ran local PCAs (Li and Ralph 2018) with PCAangsd on genotype likelihoods in nonoverlapping windows of 1000 SNPs in each chromosome to extract local covariance matrices and obtained local PCAs of genomic variation (as detailed above). For each local PCA, we analyzed the correlation between the PC1 score and PC scores from the global PCA performed on the entire genome. The resulting correlation indices were then plotted across their relative position in each chromosome. Second, genome-wide Fst was estimated between males and females. To do this, allele frequency spectrum (-doSaf 1) and minor allele frequencies were calculated for each sex with the previous list of variant positions (-sites) and their polarization as major or minor alleles (-doMajorMinor 3). Genome-wide Fst was estimated using the realSFS function in ANGSD between males and females and then summarized across sliding-windows of 25 Kb with a step of 5 Kb. Observed heterozygosity was computed for each sex and each SNP using the function -doHWE in ANGSD and then averaged across sliding-windows of 25 Kb with a step of 5 Kb using the winScan function of the R package windowscanr.

LD estimation

For the two chromosomes associated with sex determination (see results below), intra-chromosomal as well as inter-chromosomal linkage disequilibrium (LD) was calculated within each sex for a reduced number of SNPs, filtered with more stringent criteria (MAF > 10%, at least one read in 75% of the samples). Pairwise R^2 were calculated using ngsLD (Fox _et al._ 2019) based on genotype likelihood reported in the beagle file obtained previously. The 10th percentile of pairwise R^2 values was averaged across windows of 250 Kb and plotted for each sex using custom scripts available at https://github.com/clairemerot/angsd_pipeline. Then changes in LD along the sex identified chromosomes were analyzed by calculating mean R^2 for males and females separately.

Mapping known sex locus

To localize orthologous genes of previously identified sex-associated genes in our assembly, we aligned all sex-associated markers from Drinan et al. (2018) as well as others previously described to have a SD effect in flatfish and some other teleost species, to our reference assembly using LAST v.1179 (Kiełbasa _et al._ 2011; see Supplementary Table S1 for details about those genes and their accession number).

Results and discussion
Chromosome-level reference genome assembly

We generated a final genome assembly of 598.5 Mb with high contiguity (contig N50 = 3.1 Mb, scaffold N50 = 25 Mb, scaffold N90 = 20 Mb), of which 96.35% is comprised in 24 chromosome-length scaffolds (GenBank accession: GCA_006182925.3). These 24 scaffolds are consistent with the number of diploid chromosomes (2n = 48) for Atlantic Halibut (Einfeldt _et al._ 2021) and Pacific Halibut (GCA_013339905.1). The Hi-C data supported a high degree of accuracy in the overall assembly into these 24 scaffolds, as indicated by the strong concentration of data points along the diagonal rather than elsewhere in the contact maps (Supplementary Figure S1). Altogether, this suggests that these 24 scaffolds can be considered as chromosomes. The completeness of the assembly based on single-copy orthologs across eukaryotes was 98.04%, with 96.47% complete and single copy orthologs and 1.57% duplication.

There were 60,313 transcripts detected from multi-tissue transcriptomics data, with 97.7% of them (58,908 transcripts) within the 24 anchored chromosomes. A total of 19,252 genes and pseudogenes were annotated, of which 19,053 (99%) were located within the 24 chromosomes (Supplementary Table S2). Repeated elements accounted for 8.20% of chromosome lengths (49.1 Mb). These measures of completeness and assessments of repeat elements in Greenland Halibut were highly concordant with the statistics obtained from Atlantic and Pacific Halibut as well as the three other flatfish species with reference genome and important for fisheries and/or aquaculture production (Table 1). SyMap revealed shared syntenic blocks with other flatfish species (Table 2 and Figure 1). Most of the chromosomes harbor a single syntenic block compared with the most closely related _Hippoglossinae_, the Pacific Halibut (16 out of 24 chromosomes), and the Atlantic Halibut (15/24). In contrast, mapping with the three other flatfish species displayed several syntenic blocks for all chromosomes (Table 2). As a result, syntenic blocks were larger with Pacific and Atlantic Halibut than with the three other flatfishes (Table 2 and Supplementary Figure S2). Supplementary Figure S2 displays this negative relationship between the size of syntenic blocks and phylogenetic distance.

By generating this chromosome-level genome assembly for Greenland Halibut, among the most contiguous of flatfish assemblies to date, we provide a genomic resource for this unique species of the genus _Reinhardtius_. This reference genome will now become a fundamental tool to understand the genomic variability of Greenland Halibut in the context of comparative genomics and fisheries applications.

XY sex determination system with a putative autosomal-Y fusion

After mapping the reads of 99 male and 99 female Greenland Halibut samples to the reference genome, cleaning for quality and processing for SNP identification with genotype likelihoods, we identified 7,301,470 single nucleotide polymorphisms (SNPs) with a minor allelic frequency (MAF) above 5%. These SNPs were then used for further analyses pertaining to sex differentiation.

Genome-wide variation analyzed by a global PCA displayed two clusters corresponding to males and females on PC1, which explained 5.7% of the genetic variance (Supplementary Figure S4). Local PCAs on windows of 1000 SNPs along the genome revealed that this sex differentiation was mainly explained by genetic variation on two chromosomes (Chr10 and Chr21, Figure 2A). The correlation between the 1st PC of the global PCA and each local PCA was very low at the very beginning of each of those two chromosomes but increased rapidly along both chromosomes. A large portion of each of these two chromosomes explains the total genetic variation observed between the two sexes in PC1. Particularly, the last 66.5% of Chr10 (16.5 Mb out of

Table 2 Syntenic relations between Greenland Halibut and related commercial flatfish species

|	Greenland Halibut/ Pacific Halibut	Greenland Halibut/ Atlantic Halibut	Greenland Halibut/ Japanese Flounder	Greenland Halibut/ Turbot	Greenland Halibut/ Tongue sole
Number of blocks	40	45	191	333	745
Number of chromo-somes with single block	16	15	0	0	0
Minimum size	560,659	537,880	39,610	23,190	25,486
Maximum size	28,459,961	30,361,793	21,363,777	24,314,634	16,925,444
Median size	24,561,162	11,813,728	1,877,480	507,912	482,179
Number of blocks >10 Mb	24	24	11	23	6
Percentage of blocks >10 Mb (%)	60	53	6	7	1
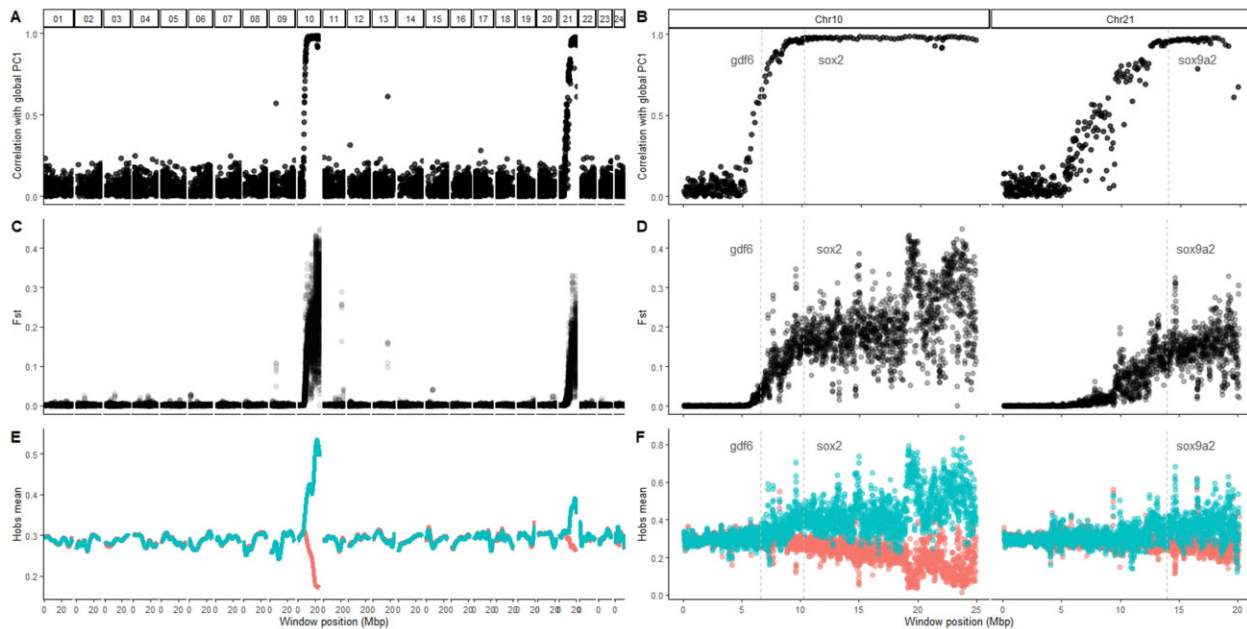
Figure 2 Two genomic regions implied in sex differentiation. (A) Correlation between PC1 scores of the local PCAs performed on windows of 1000 SNPs and PC1 scores of the global PCA along with their genomic position in each chromosome. (C) Fst differentiation estimated between males and females in sliding-windows of 25 Kb within each chromosome. (E) Mean observed heterozygosity in each sex smoothed for visualization. Panels (B), (D), and (F) are zooms on the two differentiated chromosomes (Chrs 10 and 21) of the previous values, respectively (A), (C), and (E). Vertical dashed gray lines indicate the positions of detected sex-related candidate genes.

24.8 Mb) and 52.2% of Chr21 (10.5 Mb out of 20.1 Mb) explained more than 90% of the sex differentiation observed on global PC1 (Figure 2B). Genome-wide Fst estimations between males and females corroborate this result revealing high differentiation on Chr10 and Chr21 (Figure 2C). While the mean genome-wide Fst between the sexes is as low as 0.01, Chr10 and Chr21 harbored a large region with higher Fst values with a mean Fst of 0.21 and 0.15, respectively (Figure 2D). Few localized peaks of sex differentiation were also observed on Chr9, Chr11, Chr12, Chr13, and Chr15 (Figure 2B).

The observed proportion of heterozygotes followed the same pattern, with similar values between males and females across most of the genome (Figure 2E), but with a much higher fraction of heterozygotes in males than females in the two major regions previously described. This difference between males and females was more pronounced in Chr10 than in Chr21. Both Chr10 and Chr21 were also characterized by strong LD, which was markedly higher in males than in females (Figure 3), suggesting lower

recombination rate in males than females for these genomic regions.

In contrast with differentiation and heterozygosity, coverage did not differ between males and females, neither genome-wide nor in regions of sex differentiation (Supplementary Figure S5), suggesting that Greenland Halibut is not harboring sex chromosome heteromorphism. Moreover, the putative sex chromosomes Chr10 and Chr21 have apparently not undergone degradation, which could be either typical of early SD systems or the result of an older system that has sex chromosome turnover as found in other fish species (see below).

Altogether, these results point toward a major role of Chr10 and Chr21 in genetic sex determination. The higher heterozygosity and LD in males suggest that Greenland Halibut is characterized by a XX/XY sex determination mechanism. The putative male Y chromosome exhibits two characteristics of sex chromosome evolution: it is divergent from the X chromosome and exhibits a reduction of recombination. However, it does not
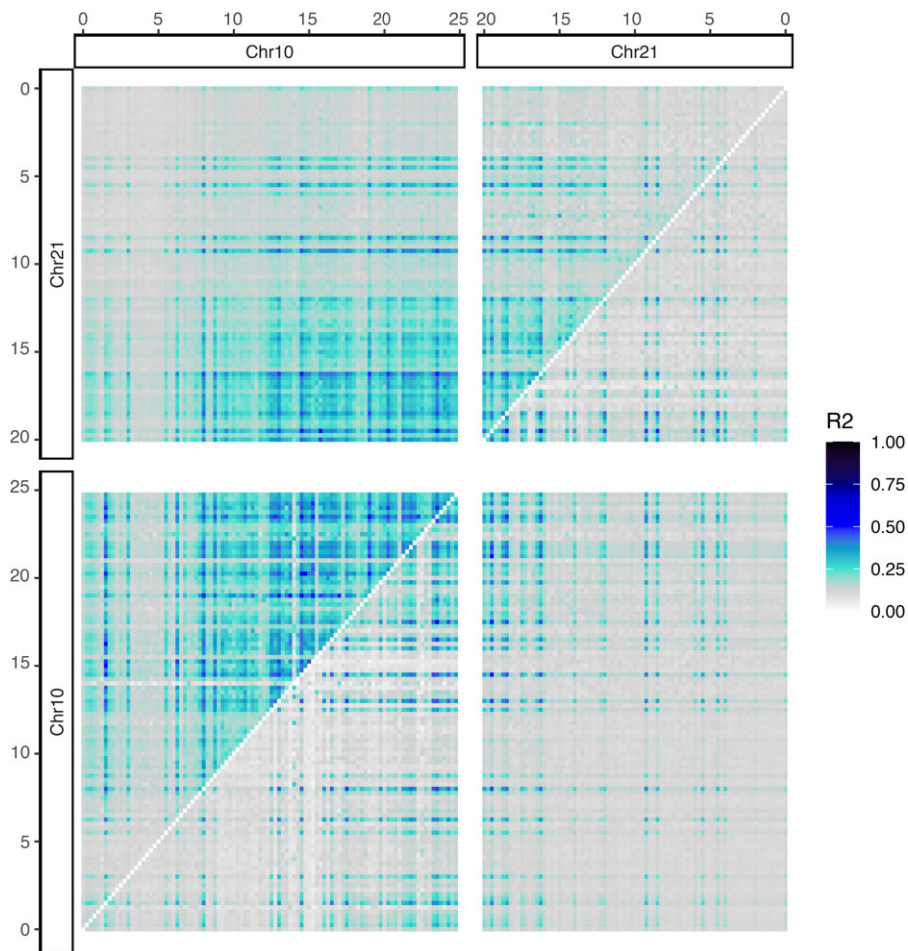
Figure 3 Intra- and Interchromosomal LD among males and females, revealing a putative Y-autosomal fusion. The color scale corresponds to the 10th highest percentile of the R^2 value of each SNP pairwise summarized by windows of 250 Kb. Male values are displayed above the diagonal and female values are displayed below the diagonal. Positions are represented in megabases.

appear degraded, suggesting either very nascent sex chromosome at an early evolutionary stage or resulting from the high evolutionary lability of SD in teleost fish. This is corroborated by the syntenic relationships with closely related flatfish species which indicate that homologs of Chr10 and Chr21 are autosome chromosomes. Chr10 and Chr21 neither correspond to locations where sex-determining regions have been defined in other flatfish species (Figure 1, Table 1, and Supplementary Figure S3). Overall, the resolved sex-determining system in Greenland Halibut is highly consistent with the rapid turnover of sex chromosomes in fish, particularly in flatfish, given that the Atlantic Halibut has also undergone a recent evolution of an XY SD mechanism (Einfeldt *et al.* 2021).

Observing high sex differentiation and characteristics of sex chromosome evolution on two different chromosomes was unexpected and puzzling. The Hi-C contact map very strongly supported that Chr10 and Chr21 are distinct chromosomes in the specimen used for the assembly, a juvenile female (Supplementary Figure S1). Inter-chromosomal LD was also low in females, confirming the independent segregation of Chr10 and Chr21. Conversely, LD between Chr10 and Chr21 was very high in males (Figure 3), suggesting that the male haplotypes of those two chromosomes segregate together. One possible mechanism explaining such pattern would be a male-restricted fusion between the Y chromosome and one autosome. Chromosome fission and fusion events are frequent in the evolution of

vertebrates, particularly in fish, in which they could explain the enormous species richness and rapid evolution of karyotypes (Cheng *et al.* 2020; Sutherland *et al.* 2017). Fusions may also be polymorphic in fish and play an adaptive role, as reported in Atlantic salmon (*Salmo salar*), in which fused and unfused karyotypes have been shown to coexist (Lehnert *et al.* 2019; Wellband *et al.* 2019). Moreover, fusions between a Y chromosome and an autosome are known to occur more frequently than autosomal fusions or W-autosome fusions, particularly in fishes (Pennell *et al.* 2015). The most probable explanation for the preponderance of Y-autosome fusions involves both selection and sex biases, as the smallest effective size (Ne) of the Y makes it more likely to fix rearrangements, including those with partially deleterious effects. This process may be even more prevalent in species with polygynous mating system like fish, which increases the variance of male reproductive success, as well as in species with skewed sex-ratios, such as the Greenland Halibut (Ghinter *et al.* 2019). Because of its significant impact on recombination, chromosomal fusion may be one of the key factors underlying the evolution of differentiated sex chromosomes. Fusions are a peculiar class of rearrangements that not only link chromosomes that previously segregated independently but also reduce recombination rates within each arm and prevent recombination between fused and unfused homologous chromosomes in polymorphic populations (Dumas and Britton-Davidian 2002). This typically leads to gradual differentiation from the fusion point as observed

here for Greenland Halibut or for the young sex chromosomes found in three-spined stickleback (*Gasterosteus aculeatus*) (Kitano et al. 2009). This represents a typical step in the birth and establishment of new sex chromosomes (McAllyster 2003; Kirkpatrick 2010), often mediated *via* large chromosomal rearrangements such as inversions or fusions. One of the main determinants of this process is selection against recombination, because linkage between SD loci and alleles beneficial to one sex has the potential to resolve sexual conflict (Charlesworth et al. 2005). As discussed below, we identified three sex-associated loci distributed on two chromosomes and one may speculate whether epistatic interactions may occur between them and putative other candidates not identified in this study. Further work would be needed to better understand gene interactions in this region and more formally confirm the putative male-limited fusion (Guerrero and Kirkpatrick 2014; Wright et al. 2016).

Several gene candidate drivers of Greenland Halibut SD

Alignment of DNA or mRNA sequences of previously described sex-associated markers revealed that several putative gene candidate drivers of SD were distributed throughout Greenland Halibut assembly (on seven chromosomes, Supplementary Table S1). Most of those genes are located in regions of low differentiation between males and females: *amhr2*, *cyp19a1A* on chr05; *cyp11b2* and *rspo1* on chr08; *dmrt1* on chr09, *foxl2* on chr10 and *gsdf* on chr18, making them unlikely to underlie sex differentiation in Greenland Halibut. However, three candidate genes mapped to the putative sex chromosomes: *gdf6* and *sox2* on Chr10, and *sox9a2* on Chr21 (Figure 2 and Supplementary Table S1) and are thus candidate drivers of SD in Greenland Halibut.

Growth differentiation factor 6 (*Gdf6*) is a member of the TGF-β family involved in vertebral segmentation and cell differentiation. While no gonadal function was reported previously, *gdf6Y* emerged recently as a putative SD gene in the Turquoise Killifish (*Nothobranchius furzeri*) (Reichwald et al. 2015) because sex-linked allelic variations were found to affect the *Gdf6* protein function. The Y chromosome allele, *gdf6Y*, also exhibits sex-linked expression during testicular differentiation in the Turquoise Killifish, supporting its SD role but functional evidence of its role as a SD gene is yet not available.

Sox2 is a member of the *SOX* (SRY-related HMG-box) family with a high-mobility-group (HMG) DNA-binding domain (Zhang and Cui 2014). It is an important transcription factor, which is involved in developmental processes of vertebrates, such as embryogenesis, maintenance of stem cells, and neurogenesis. In the Rohu Carp (*Labeo rohita*), *sox2* mRNA is expressed in various organs, as well as in the culture proliferating spermatogonial stem cells (Patra et al. 2015). *Sox2* is also a candidate SD gene in two aquaculture species, the mollusk Zhikong Scallop (*Chlamys varia*) (Liang et al. 2019) and the crustacean Black Tiger Shrimp (*Penaeus monodon*) (Guo et al. 2019). In flatfishes, *sox2* is also a strong candidate for SD. For instance, in Japanese Flounder *sox2* is expressed in gonadal tissues and the transcript abundance in ovaries is higher than in testis (Gao et al. 2014). In Turbot, *sox2* is defined as a consistent candidate gene putatively driving SD (Martinez et al. 2021).

The *Sox9* gene has been related to sexual differentiation in several teleost species: Rainbow Trout, (*Oncorhynchus mykiss*, Takamatsu et al. 1997; Vizziano et al. 2007), Tilapia (*Oreochromis niloticus*, Ijiri et al. 2008), Zebrafish (*Danio rerio*; Chiang et al., 2001; von Hofsten and Olsson 2005), Medaka, *Oryzias latipes* (Yokoi et al. 2002), Guppy (*Poecilia reticul*, Shen et al. 2007), the Common Carp (*Cyprinus carpio*, Du et al. 2007), and the Asian Swamp Eel

(*Monopterus albus*, Zhou et al. 2003). In the Rainbow Trout, two *sox9* genes have been identified: *sox9a1* and *sox9a2*. *Sox9a1* shows higher expression levels in males than in females before sexual differentiation (Vizziano et al. 2007). In Tilapia, *sox9* expression is similar in XX and XY gonads, but, after sexual differentiation, it is positively regulated in XY gonads (Ijiri et al. 2008). In Medaka, testicular *sox9* (*sox9a2*) expression in somatic cells is similar in both sexes during early gonadal differentiation and becomes positively regulated in males during the testicular lobe formation stage (Nakamoto et al. 2005). Overall, these observations support the hypothesis of a significant role for *sox9* in either testicular differentiation or testicular development. This gene may also be performing a similar function in flatfish. In Atlantic Halibut, *sox9a2* is included in a cascade of gene activation starting with the *dmrt1* gene which is activated by the putative SD factor, the *gsdf* gene (Norris and Carr 2013; Einfeldt et al. 2021).

In Greenland Halibut, *sox2*, *gdf6*, and *sox9a2* represent three possible candidate genes for sex determination. One can think that these three genes may be interacting in shaping the determination of sex and/or the subsequent development of sex organs. In addition, it is not excluded that some genes located in the highest peaks of differentiation but currently not included in the search neither in the annotation, and thus not identified by our study, could also have an effect on sex differentiation in Greenland Halibut and/or interact with current identified candidates.

SD genes identified in the two most closely related flatfish species, Atlantic and Pacific Halibut (Table 1), do not appear to differentiate male and female Greenland Halibut. This is somewhat unsurprising given that even within the same genus (*Hyppoglossus*), not only do different species have different genes driving SD, but they also harbor a different heterogametic sex chromosome system (XY for Atlantic Halibut and ZW for Pacific Halibut, Drinan et al. 2018; Einfeldt et al. 2021).

The identified SD system with low differentiation between sex chromosomes revealed in Greenland Halibut might reflect the observed complex and still enigmatic sex maturation. For example, in some Greenland Halibut individuals the maturation cycle can be interrupted by a sudden degeneration of the oocytes (Fedorov 1971). More recently, it has been suggested that this species utilizes a very unusual oocyte development pattern containing two simultaneously developing cohorts of vitellogenic oocytes, not seen in any related species (Rideout et al. 2012; Dominguez-Petit et al. 2018). Also, Greenland Halibut present high densities and regular occurrence of rodlet cells in male and female gonads, which is unusual and still unexplained (Rideout et al. 2015). Furthermore, environmental factors such as temperature, pH, population density, and social interactions have all been found to influence the sex-ratio in fish (Crews 1976; Nakamura et al. 1998; Baroiller et al. 2009) and we cannot rule out a possible role of such factors. Clearly, more research is needed to better disentangle the relative contributions of environmental and genetic factors in SD in Greenland Halibut.

Conclusion

Despite the commercial importance of Greenland Halibut, major knowledge gaps still persist for this species, namely pertaining to its reproduction cycle and sex determination mechanisms. In this study, we used a compendium of recent technologies [single-molecule sequencing of long reads (Pacific Sciences) with chromatin conformation capture sequencing (Hi-C) data and whole-genome sequencing of hundreds of individuals] to provide the

very first chromosome-level genome reference and explore the sex determination mechanism of this demersal fish. The chromosome-level assembly and the investigation of its syntenic relationships with other economically important flatfish species highlight the high conservation of synteny blocks within the flat-fish phylogenetic clade. In fish, the early evolutionary stage of most sex chromosomes and the subsequent small differences between those chromosomes make it extremely difficult to find diagnostic differences between them. Our sex determination analysis revealed that flatfishes do not escape this rule and exhibit a high level of lability and turnover in SD mechanisms. A whole-genome sequencing approach of 198 individuals allowed us to unravel the molecular SD system in Greenland Halibut. Specifically, our results support the view that Greenland Halibut has a male heterogametic XY system, with several putative candidate genes, also described as SD drivers in other flatfish species like the Turbot (*S. maximus*) and the Japanese Flounder. Interestingly, our study also suggested for the first time in flatfish the presence of a putative Y-autosomal fusion that could be associated with a reduction of recombination typical of the early stages of sex chromosome evolution.

Data availability

Genome assembly is available on GenBank with accession number GCA_006182925.3. Biosamples for the 8 RNA-seq lanes are available on NCBI with the following numbers: SAMN18221283 to SAMN18221290, and 198 individual whole-genome sequences are submitted under the BioProject ID PRJNA737354 (SAMN19692655 to SAMN19692852) on NCBI SRA and will be available upon publication.

Supplementary material is available at G3 online.

Acknowledgments

Funding

Conflicts of interest

The authors declare that there is no conflict of interest.

Literature cited

Abbott JK, Nordén AK, Hansson B. 2016. Sex chromosome evolution: historical insights and future. Proc Biol Sci. 284:20162806.

Aljanabi SM, Martinez I. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Res. 25:4692–4693.

Bachtrog D. 2006. A dynamic view of sex chromosome evolution. Curr Opin Genet Dev. 16:578–585.

Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat Rev Genet. 14: 113–124.

Bao L, Tian C, Liu S, Zhang Y, Elaswad A, *et al.* 2019. The Y chromosome sequence of the channel catfish suggests novel sex determination mechanisms in teleost fish. BMC Biol. 17:6.

Baroiller JF, D'Cotta H, Saillant E. 2009. Environmental effects on fish sex determination and differentiation. Sex Dev. 3:118–135.

Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, *et al.* 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. PLoS One. 10:e0128036.

Benestan L, Moore J-S, Sutherland BJG, Le Luyer J, Maaroufi H, *et al.* 2017. Sex matters in massive parallel sequencing: evidence for biases in genetic parameter estimation and investigation of sex determination systems. Mol Ecol. 26:6767–6783.

Bernatchez L, Wellenreuther M, Araneda C, Ashton DT, Barth JMI, *et al.* 2017. Harnessing the power of genomics to secure the future of seafood. Trends Ecol Evol. 32:665–680.

Betancur R, Ortí G. 2014. Molecular evidence for the monophyly of flatfishes (Carangimorpharia: Pleuronectiformes). Mol Phylogenet Evol. 73:18–22.

Bull JJ. 1983. Evolution of Sex Determining Mechanisms. Benjamin/Cummings Publishing Co, the University of California

Capel B. 2017. Vertebrate sex determination: evolutionary plasticity of a fundamental switch. Nat Rev Genet. 18:675–689.

Charlesworth D, Charlesworth B, Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. Heredity (Edinb). 95: 118–128.

Chen S, Zhang G, Shao C, Huang Q, Liu G, *et al.* 2014. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. Nat Genet. 46: 253–260.

Cheng Y, Shang D, Luo M, Huang C, Lai F, *et al.* 2020. Whole genome-wide chromosome fusion and new gene birth in the *Monopterus albus* genome. Cell Biosci. 10:67.

Chiang EF, Pai CI, Wyatt M, Yan YL, Postlethwait J, *et al.* 2001. Two sox9 genes on duplicated zebrafish chromosomes: expression of similar transcription activators in distinct sites. Developmental biology. 231:149–163. Doi: 10.1006/dbio.2000.0129.

Crews D. 1976. Temperature-dependent sex determination: the interplay of steroid hormones and temperature. Zoolog Sci. 3:1–13.

DFO (Fisheries and Oceans Canada). 2018. Assessment of the Greenland Halibut Stock in the Gulf of St. Lawrence (4RST) in 2017. Canadian Science Advisory Secretariat Science Advisory Report 2018/035. Fisheries and Oceans Canada.

Dominguez-Petit R, Rideout RM, Garabana D, Lambert Y, Hermida M, *et al.* 2018. Original article evaluating the use of the autodiametric method for estimating fecundity of *Reinhardtius hippoglossoides*, a species with an unusual oocyte development strategy. ICES J Mar Sci. 75:831–839.

Drinan DP, Loher T, Hauser L. 2018. Identification of genomic regions associated with sex in Pacific Halibut (S Grant, Ed). J Hered. 109: 326–332.

Du Q-Y, Wang F-Y, Hua H-Y, Chang Z-J. 2007. Cloning and study of adult-tissue-specific expression of *Sox9* in *Cyprinus carpio*. J Genet. 86:85–91.

Dumas D, Britton-Davidian J. 2002. Chromosomal rearrangements and evolution of recombination: comparison of chiasma distribution patterns in standard and Robertsonian populations of the house mouse. Genetics. 162:1355–1366.

Einfeldt AL, Kess T, Messmer A. 2021. Chromosome level reference of Atlantic Halibut *Hippoglossus hippoglossus* provides insight into the evolution of sexual determination systems. Mol Ecol Resources. 21:1686–1696.

Fedorov KE. 1971. The state of the gonads of the Barents sea Greenland Halibut, *Reinhardtius hippoglossoides* (Walb.) in connection with failure to spawn. Vop Ichtiol. 11:673–682.

Figueras A, Robledo D, Corvelo A, Hermida M, Pereiro P, *et al.* 2016. Whole genome sequencing of turbot (*Scophthalmus maximus*; Pleuronectiformes): a fish adapted to demersal life. DNA Res. 23: 181–192.

Fox EA, Wright AE, Fumagalli M, Vieira FG. 2019. ngsLD: evaluating linkage disequilibrium using genotype likelihoods. Bioinformatics. 35:3855–3856.

Furman BLS, Metzger DCH, Darolti I, Wright AE, Sandkam BA, *et al.* 2020. Sex chromosome evolution: so many exceptions to the rules. Genome Biol Evol. 12:750–763.

Gao J, Wang Z, Shao K, Fan L, Yang L, *et al.* 2014. Identification and characterization of a *Sox2* homolog in the Japanese flounder *Paralichthys olivaceus*. Gene. 544:165–176.

Ghinter L, Lambert Y, Audet C. 2019. Influence of sex on growth of juvenile Greenland Halibut (*Reinhardtius hippoglossoides*) in the Estuary and Gulf of St. Lawrence. Fisheries Res. 219:105332.

Guerrero RF, Kirkpatrick M. 2014. Local adaptation and the evolution of chromosome fusions. Evolution. 68:2747–2756.

Guerrero-Estévez S, Moreno-Mendoza N. 2010. Sexual determination and differentiation in teleost fish. Rev Fish Biol Fisheries. 20: 101–121.

Guo L, Xu Y-H, Zhang N, Zhou F-L, Huang J-H, *et al.* 2019. A high-density genetic linkage map and QTL mapping for sex in black tiger shrimp (*Penaeus monodon*). Front Genet. 10:326.

Hagen O, Solberg C, Johnston IA. 2006. Sexual dimorphism of fast muscle fibre recruitment in farmed Atlantic Halibut (*Hippoglossus hippoglossus* L.). Aquaculture. 261:1222–1229.

Hattori RS, Somoza GM, Fernandino JI, Colautti DC, Miyoshi K, *et al.* 2019. The duplicated Y-specific *amhy* gene is conserved and linked to maleness in silversides of the genus *Odontesthes*. Genes. 10:679.

Ijiri S, Kaneko H, Kobayashi T, Wang D-S, Sakai F, *et al.* 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.

Imsland AK, Folkvord A, Grung GL, Stefansson SO, Taranger GL. 1997. Sexual dimorphism in growth and maturation of turbot, *Scophthalmus maximus* (Rafinesque, 1810). Aquaculture Res. 28: 101–114.

Keyl F, Kempf AJ, Sell AF. 2015. Sexual size dimorphism in three North Sea gadoids. J Fish Biol. 86:261–275.

Kiełbasa SM, Wan R, Sato K, Horton P, Frith MC. 2011. Adaptive seeds tame genomic sequence comparison. Genome Res. 21:487–493.

Kirkpatrick M. 2010. How and why chromosome inversions evolve. PLoS Biol. 8:e1000501.

Kitano J, Peichel CL. 2012. Turnover of sex chromosomes and speciation in fishes. Environ Biol Fishes. 94:549–558.

Kitano J, Ross JA, Mori S, Kume M, Jones FC, *et al.* 2009. A role for a neo-sex chromosome in stickleback speciation. Nature. 461: 1079–1083.

Kondo M, Hornung U, Nanda I, Imai S, Sasaki T, *et al.* 2006. Genomic organization of the sex-determining and adjacent regions of the sex chromosomes of medaka. Genome Res. 16:815–826.

Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of Next Generation Sequencing Data. BMC Bioinformatics. 15:356.

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, *et al.* 2004. Versatile and open software for comparing large genomes. Genome Biol. 5:R12.

Legendre P, Legendre L. 1998. Numerical Ecology. Amsterdam: Elsevier Science.

Lehnert SJ, Bentzen P, Kess T, Lien S, Horne JB, *et al.* 2019. Chromosome polymorphisms track trans-Atlantic divergence and secondary contact in Atlantic salmon. Mol Ecol. 28: 2074–2087.

Li H, Ralph P. 2018. Local PCA shows how the effect of population structure differs along the genome. Genetics. 211:289–304.

Liang D, Fan Z, Zou Y, Tan X, Wu Z, *et al.* 2018. Characteristics of *Cyp11a* during Gonad Differentiation of the Olive Flounder *Paralichthys olivaceus*. Int J Mol Sci. 19:2641.

Liang S, Liu D, Li X, Wei M, Yu X, *et al.* 2019. SOX2 participates in spermatogenesis of Zhikong scallop *Chlamys farreri*. Sci Rep. 9:76.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, *et al.* 2009. Comprehensive mapping of long-range-interactions reveals folding principles of the human genome. Science. 326:289–293.

Lou RN, Jacobs A, Wilder AP, Therkildsen NO. 2021. A beginner's guide to low-coverage whole genome sequencing for population genomics. Mol Ecol. 1–68. 10.1111/mec.16077.

Martinez P, Robledo D, Taboada X, *et al.* 2021. A genome-wide association study, supported by a new chromosome-level genome assembly, suggests sox2 as a main driver of the undifferentiatiated ZZ/ZW sex determination of turbot *(Scophthalmus maximus)*. Genomics. 113:1–54.

Martinez P, Viñas AM, Sánchez L, *et al.* 2014. Genetic architecture of sex determination in fish: applications to sex ratio control in aquaculture. Front Genet. 5:340.

McAllyster BF. 2003. sequence differentiation associated with an inversion on the neo-x chromosome of *Drosophila americana*. Genetics. 165:1317–1328.

McGrath C. 2020. Highlight: exceptions are the rule in sex determination. Genome Biol Evol. 12:764–765.

Meisner J, Albrechtsen A. 2018. Inferring population structure and admixture proportions in low-depth NGS data. Genetics. 210: 719–731.

Mérot C, Berdan E, Cayuela H, *et al.* 2021. Locally adaptive inversions modulate genetic variation at different geographic scales in a seaweed fly. Mol Biol Evol. 38:msab143.

Nakamoto M, Suzuki A, Matsuda M, Nagahama Y, Shibata N. 2005. Testicular type *Sox9* is not involved in sex determination but might be in the development of testicular structures in the medaka, *Oryzias latipes*. Biochem Biophys Res Commun. 333: 729–736.

Nakamura M, Kobayashi T, Chang XT, Nagahama Y. 1998. Gonadal sex differentiation in teleost fish. J Exp Zool. 281:362–372.

Nanda I, Kondo M, Hornung U, Asakawa S, Winkler C, *et al.* 2002. A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. Proc Natl Acad Sci USA. 99:11778–11783.

Nielsen R, Paul JS, Albrechtsen A, Song YS. 2011. Genotype and SNP Calling from Next-Generation Sequencing Data. Nature Review Genetics. 12:443–451.

Norris DO, Carr JA. 2013. Comparative aspects of vertebrate reproduction. In: DO Norris, JA Carr, editors. Vertebrate Endocrinology. Springer US. 5th ed. p. 375–441.

Ospina-Alvarez N, Piferrer F. 2008. Temperature-dependent sex determination in fish revisited: prevalence, a single sex ratio response pattern, and possible effects of climate change. PLoS One. 3:e2837.

Ovenden JR, Berry O, Welch DJ, Buckworth RC, Dichmont CM. 2013. Ocean's eleven: a critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries. Fish Fish. 16:125–159.

Pan Q, Feron R, Yano A, Guyomard R, Jouanno E, *et al.* 2019. Identification of the master sex determining gene in Northern pike (*Esox lucius*) reveals restricted sex chromosome differentiation. PLoS Genet. 15:e1008013.

Pan Q, Feron R, Jouanno E Darras H, Herpin A, *et al.* 2021. The rise and fall of the ancient Northern pike master sex-determining gene. eLife. 10:e62858.

Palaiokostas C, Bekaert M, Davie A, Cowan ME, Oral M, *et al.* 2013. Mapping the sex determination locus in the Atlantic Halibut (*Hippoglossus hippoglossus*) using RAD sequencing. BMC Genomics. 14:566.

Patra SK, Chakrapani V, Panda RP, Mohapatra C, Jayasankar P, *et al.* 2015. First evidence of molecular characterization of rohu carp *Sox2* gene being expressed in proliferating spermatogonial cells. Theriogenology. 84:268–276.e1.

Pennell MW, Kirkpatrick M, Otto SP, Vamosi JC, Peichel CL, *et al.* 2015. Y fuse? sex chromosome fusions in fishes and reptiles. PLoS Genet. 11:e1005237.

Rastetter RH, Smith CA, Wilhelm D. 2015. The role of non-coding RNAs in male sex determination and differentiation. Reproduction. 150:R93–R107.

Core Team R. 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical computing. https://www.R-project.org/.

Reichwald K, Petzold A, Koch P, Downie BR, Hartmann N, *et al.* 2015. Insights into sex chromosome evolution and aging from the genome of a short-lived fish. Cell. 163:1527–1538.

Rideout RM, Morgan MJ, Lambert Y Cohen AM, *et al.* 2012. Oocyte development and vitellogenin production in Northwest Atlantic Greenland Halibut *Reinhardtius hippoglossoides*. J Northwest Atlantic Fish Sci. 44:15–29.

Rideout RM, Smith SA, Morgan MJ. 2015. High-density aggregations of rodlet cells in the gonads of Greenland Halibut *Reinhardtius hippoglossoides*, a deep-water marine flatfish. J Fish Biol. 86: 1630–1637.

Rljnsdorp AD, Ibelings B. 1989. Sexual dimorphism in the energetics of reproduction and growth of the North Sea plaice, *Pleuronectes platessa* L. J Fish Biol. 35:401–415.

Robledo D, Hermida M, Rubiolo JA, Fernández C, Blanco A, *et al.* 2017. Integrating genomic resources of flatfish (Pleuronectiformes) to boost aquaculture production. Comp Biochem Physiol Part D Genomics Proteomics. 21:41–55.

Rondeau EB, Laurie CV, Johnson SC, Ben F. Koop 2016. A PCR assay detects a male-specific duplicated copy of Anti-Müllerian hormone (*amh*) in the lingcod (*Ophiodon elongatus*). BMC Res Notes. 9: 230.

Ruan J, Li H. 2020. Fast and accurate long-read assembly with wtdbg2. Nat Methods. 17:155–158.

Shao C, Bao B, Xie Z, Chen X, Li B, *et al.* 2017. The genome and transcriptome of Japanese flounder provide insights into flatfish asymmetry. Nat Genet. 49:119–128.

Shen XY, Cui JZ, Gong QL, Nagahama Y. 2007. Cloning of the full-length coding sequence and expression analysis of *Sox9b* in guppy (*Poecilia reticulata*). Fish Physiol Biochem. 33:195–202.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 31: 3210–3212.

Smit AFA, Hubley R, Green P. 2013–2015. RepeatMasker Open-4.0. http://www.repeatmasker.org

Soderlund C, Bomhoff M, Nelson W. 2011. SyMAP v3.4: a turnkey synteny system with application to plant genomes. Nucleic Acids Res. 39:e68.

Soderlund C, Nelson W, Shoemaker A, Paterson A. 2006. SyMAP: a system for discovering and viewing syntenic regions of FPC maps. Genome Res. 16:1159–1168.

Sutherland BJG, Rico C, Audet C, Bernatchez L. 2017. Sex chromosome evolution, Heterochiasmy, and Physiological QTL in the salmonid brook charr *Salvelinus fontinalis*. G3 (Bethesda). 7: 2749–2762.

Takamatsu N, Kanda H, Ito M, Yamashita A, Yamashita S, *et al.* 1997. Rainbow trout *SOX9*: cDNA cloning, gene structure and expression. Gene. 202:167–170.

Tavares H. 2021. windowscanr: Apply Functions Using Sliding Windows. R package version 0.1.

Therkildsen NO, Palumbi SR. 2017. Practical low-coverage genome-wide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in nonmodel species. Mol Ecol Resour. 17:194–115.

Uller T, Pen I, Wapstra E, Beukeboom LW, Komdeur J. 2007. The evolution of sex ratios and sex-determining systems. Trends Ecol Evol. 22:292–297.

Vinnikov KA, Thomson RC, Munroe TA. 2018. Revised classification of the righteye flounders (Teleostei: Pleuronectidae) based on multilocus phylogeny with complete taxon sampling. Mol Phylogenet Evol. 125:147–162.

Vizziano D, Randuineau G, Baron D, Cauty C, Guiguen Y. 2007. Characterization of early molecular sex differentiation in rainbow trout, *Oncorhynchus mykiss*. Dev Dyn. 236:2198–2206.

VonHofsten J, Olsson PE. 2005. Zebrafish sex determination and differentiation: involvement of *FTZ-F1* genes. Rep Biol Endocrinol. 3: 63.

Wellband K, Mérot C, Linnansaari T, Elliott JAK, Curry RA, *et al.* 2019. Chromosomal fusion and life history-associated genomic variation contribute to within-river local adaptation of Atlantic salmon. Mol Ecol. 28:1439–1459.

Whibley A, Joanna K, Narum SR. 2021. The changing face of genome assemblies: guidance on achieving high-quality reference genomes. Mol Ecol Resour. 21:641–652.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag.

Wright A, Dean R, Zimmer F, Mank JE. 2016. How to make a sex chromosome. Nat Commun. 7:12087.

Yokoi H, Kobayashi T, Tanaka M, Nagahama Y, Wakamatsu Y, *et al.* 2002. *Sox9* in a teleost fish, medaka (*Oryzias latipes*): evidence for diversified function of Sox9 in gonad differentiation. Mol Reprod Dev. 63:5–16.

Yano A, Nicol B, Jouanno E, Quillet E, Fostier A, *et al.* 2013. The sexually dimorphic on the Y chromosome gene (*sdY*) is a conserved

male specific Y chromosome sequence in many salmonids. Evol Appl. 6:486–496.

Zhang S, Cui W. 2014. Sox2, a key factor in the regulation of pluripotency and neural differentiation. World journal of stem cells. 6: 305–311. Doi: 10.4252/wjsc.v6.i3.305.

Zhou R, Liu L, Guo Y, Yu H, Cheng H, *et al.* 2003. Similar gene structure of two *Sox9a* genes and their expression patterns during gonadal differentiation in a teleost fish, rice field eel (*Monopterus albus*). Mol Reprod Dev. 66:211–217.

Communicating editor: M. Arbeitman