Analysis of the histone cluster in Senegalese sole (*Solea senegalensis*): evidence for a divergent evolution of two canonical histone clusters

Journal:	<i>Genome</i>
Manuscript ID	gen-2016-0143.R1
Manuscript Type:	Article
Date Submitted by the Author:	05-Oct-2016
Complete List of Authors:	Merlo, Manuel Alejandro; Universidad de Cadiz, Area de Genetica, Facultad de Ciencias del Mar y Ambientales Iziga, Roger; Universidad de Cadiz, Area de Genetica, Facultad de Ciencias del Mar y Ambientales Portela, Silvia; Universidad de Cadiz, Area de Genetica, Facultad de Ciencias del Mar y Ambientales Cross, Ismael; Universidad de Cadiz, Area de Genetica, Facultad de Ciencias del Mar y Ambientales Kosyakova, Nadezda; Universit��tsklinikum Jena, Institut fur humangenetik, Molekulare Zytogenetik Liehr, Thomas; Universit��tsklinikum Jena, Institut fur humangenetik, Molekulare Zytogenetik Manchado, Manuel; Junta de Andalucia, Laboratorio de Biologia Molecular, Centro IFAPA El Toru��o Rebordinos, Laureana; Universidad de Cadiz, Area de Genetica, Facultad de Ciencias del Mar y Ambientales
Keyword:	<i>Solea senegalensis</i> , chromosome evolution, Pleuronectiformes, BAC clones

SCHOLARONETM
Manuscripts

Analysis of the histone cluster in Senegalese sole (*Solea senegalensis*): evidence for a divergent evolution of two canonical histone clusters

Manuel Alejandro Merlo^a, Roger Iziga^a, Silvia Portela-Bens^a, Ismael Cross^a, Nadezda Kosyakova^b, Thomas Liehr^b, Manuel Manchado^c, and Laureana Rebordinos^a

^a Área de Genética, Facultad de Ciencias del Mar y Ambientales, Universidad de Cádiz, 11510 Cádiz, Spain.

^b Institut für Humangenetik, Universitätsklinikum Jena, 07743 Jena, Germany.

^c Centro IFAPA “El Toruño”, 11500 Puerto de Santa María, Cádiz, Spain.

Corresponding author: Laureana Rebordinos (email: laureana.rebordinos@uca.es).
Tlf.: +34 956016181. Área de Genética, Facultad de Ciencias del Mar y Ambientales, Universidad de Cádiz, 11510 Cádiz, Spain.

ABSTRACT

The Senegalese sole (*Solea senegalensis*) is commercially very important, and is a priority species for aquaculture product diversification. The main histone cluster was found in two BAC clones. However, two replacement histones (H1.0 and H3.3) were found in another BAC clone. Different types of canonical H2A and H2B have been found in one species for the first time. Phylogenetic analysis demonstrated that the different H1, H2A and H2B types were all more similar to each other than to canonical histones from other species. The canonical H3 of *S. senegalensis* differs from subtypes H3.1 and H3.2 in humans in the residue 96, in which a serine residue is found instead of an alanine. This same polymorphism has been found only in *Danio rerio*. The *S. senegalensis* karyotype comprises 21 pairs of chromosomes, distributed in 3 metacentric pairs, 2 submetacentric pairs, 4 subtelocentric pairs and 12 acrocentric pairs. The two BAC clones that contain the clusters of canonical histones have both been mapped in the largest metacentric pair, and mFISH confirmed the co-location with the *dmrt1* gene in that pair. Three chromosome markers have been identified, which in addition to those previously described, account for 18 chromosome pairs with markers in *S. senegalensis*.

Key words

Solea senegalensis, Pleuronectiformes, histone arrangement, BAC clones, mFISH, chromosome evolution

Introduction

The histones are basic structural proteins of the nucleosomes and are essential for DNA packaging in eukaryotic organisms. This group is subdivided into five protein classes, termed H1, H2A, H2B, H3, and H4. The latter four are regarded as core histone genes and they constitute the octameric structure of the nucleosome. Concerning the DNA, a fragment of 146 bp is wrapped around the central protein moiety and the linker DNA connects neighboring nucleosomal cores. H1 histone is termed linker histone, since it interacts with the linker DNA at its entry-exit site at the nucleosomal core. H1 acts at both nucleosomal and supranucleosomal levels, since it also interacts with the wrapped DNA, thus resulting in higher order structures (Doenecke et al. 1997).

The core and linker histones have functional consequences, since trans-acting regulatory factors and the transcription machinery must gain access to their specific binding sites at the DNA level (Doenecke et al. 1997). Thus, the histones play a critical role in the regulation of expression. Because H1 histone binds to scaffold-associated regions and participates in nucleosome positioning, the higher-order organization of chromatin by H1 is the mechanism by which H1 plays its regulatory role (Eirín-López et al. 2002). Moreover, H1 histone has also been proposed as an efficient protein for gene transfer (Wei et al. 2007).

The histone proteins are extremely well-conserved throughout evolution and, in mammals, each of the histone classes, except H4, is subdivided into several non-allelic variants, thus generating histone subtypes which vary structurally and are encoded by single copy genes (Doenecke et al. 1997). Several histone genes possess two types named as canonical or replication dependent (RD) histones on one hand, and replacement or replication independent (RI) histones on the other hand. The former are highly expressed just before the S-phase and are then repressed on the completion of DNA replication, whereas the latter are expressed independently of the cellular cycle. The most divergent class of histones is the H1 protein family (Cole 1987). The five main gene types of H1

histone are H1.1, H1.2, H1.3, H1.4, and H1.5 (Eick et al. 1989; Albig et al. 1991). In addition, there are two subtypes: an H1^o replacement and an H1t germ cell-specific subtype (Doenecke et al. 1997). All five main H1 histones are expressed in dividing cells, while H1.3 and H1.5 are majority expressed in non-dividing cells (Ponte et al. 1998). H1^o is expressed in several differentiated tissues, and H1t is restricted to mammalian testicular cells (Eirín-López et al. 2009). Several subtypes of the core histones H2A, H2B, and H3 have been described. The H2 histone includes the H2A subtypes H2A.1 and H2A.2, and the replacement subtypes H2A.X and H2A.Z; the H2B has subtypes H2B.1, H2B.2 and H2B.3; and the H3 has subtypes H3.1, H3.2 and H3.3 (reviewed in Doenecke et al. 1997). The H4 histone is the only class in mammals which shows invariantly the same amino acid sequence. An interspecies comparison revealed only a few variations among the H4 sequences in species as distant from each other as human and sea urchin (Grunstein et al. 1976).

Solea senegalensis (Kaup 1858), known as Senegalese sole, is a flatfish belonging to the Soleidae family. It is distributed along the north-western coast of Africa to the south-western coast of the Iberian Peninsula, including the Mediterranean Sea (Díaz-Ferguson et al. 2007). It is a species much appreciated for consumption that commands a high market price; this, in addition to its fast rate of growth in captivity, makes it an important species for aquaculture diversification (Padrós et al. 2011). In addition, this species is closely related to other commercially important species such as common sole (*Solea solea*), tongue sole (*Cynoglossus semilaevis*), turbot (*Scophthalmus maximus*) and halibut (*Hippoglossus hippoglossus*). As a result of these characteristics many research studies have been undertaken, from a variety of perspectives, such as biology (Mai et al. 2009), reproduction (Oliveira et al. 2010; Viñas et al. 2013), and aquaculture (Morais et al. 2014), among others. In recent years, an increasing number of genetic-genomic studies of Senegalese sole has been published, particularly cytogenetics (Cross et al. 2006), species identification (Manchado et al. 2006a), gene expression (Guzmán et al. 2009; Manchado et al. 2009; Iziga et al. 2010; Ponce et al. 2011) and genomic approaches (García-Cegarra et al. 2013; Benzekri et al. 2014).

The Pleuronectiformes order, in which all the flatfish species mentioned before are included, has been under discussion for many years in respect of its taxonomic status (Pardo et al. 2005; Betancur et al. 2013). The histone genes have been used as markers for taxonomic differentiation (Armbruster et al. 2005; González-Romero et al. 2008; Oskars et al. 2015; Combosch and Giribet 2016). In addition we recently reported the existence of a possible sex proto-chromosome in *S. senegalensis*, a finding that needs to be confirmed (Portela-Bens et al. 2016). In this work, a complete analysis of the histone cluster has been carried out for several purposes. First, the analysis will contribute to knowledge of the evolutionary process undergone by the histone cluster in the species; second, new clues about the origin of the larger metacentric chromosome pair have been investigated; and finally, a new marker has been obtained that could clarify the taxonomical status of the order in future studies.

Materials and Methods

Samples source

A BAC library of Senegalese sole was obtained, taking as starting material a pool of larvae before mouth opening (3 days after hatching). Larvae were washed with DEPC water and then frozen directly in liquid nitrogen. High molecular weight genomic DNA was isolated from the samples, digested with *Bam*HI, cloned into the Copy Control™ pCC1 Bacterial Artificial Chromosome (BAC)™ (Epicentre Biotechnologies, USA) and transformed into the host cell DH10B™ (Invitrogen, Life Technologies, USA). The library comprised 29,184 positive clones distributed in 384-well plates (76 plates in total). Approximately 99.99% of the clones contained nuclear DNA inserts of Senegalese sole (average size, 285 kb) with 0.01% of empty clones found.

BAC library screening

To identify a BAC clone that contained the full-length histone *H3* and *H4* genes, the BAC library was screened using the method described by Asakawa et al. (1997). The primer pairs SseH3.1/SseH3.2 and SseH4.1/SseH4.2 were used to carry out the 4D-PCR (Table 1). The non-canonical histones were found during the screening for the Toll-like Receptor 8 gene (*TLR8*), whose primer pairs are also detailed in Table 1. PCR was performed in a total volume of 25 μ L containing 5-10 ng of BAC-DNA template, 3 mM of $MgCl_2$, 1 mM of each dNTP, 5 μ mol of each primer, and 1.25 U of MyTaq™ HS Red DNA Polymerase (Bioline, UK). PCR conditions involved denaturing at 95°C for 2 min, followed by 30 cycles of 15 s at 95°C, 15 s at 60°C or 65°C and 30 s at 72°C.

BAC isolation and sequencing

After BAC library screening, positive BAC clones were isolated using two different kits, the first being the Large-Construct Kit (Qiagen, Netherlands), which was used to recover large DNA quantity. The second is the BACMAX™ DNA purification kit (Epicentre, USA), for Fluorescence *in situ* Hybridization (FISH). In both cases, the procedure was carried out following the manufacturer's recommendations.

The BAC DNA was quantified by spectrophotometry using a Nanodrop© (Thermo, USA). A yield of 5-12 μ g was required for sequencing, which was done by Lifesequencing™ using a 454 technology (Roche, Switzerland) powered by the Genome Sequencing FLX System.

Sequence annotation

The functional and structural annotation of the BAC sequences was analyzed in a semi-automatic process. Protein and EST from *S. senegalensis* and other fish species, such as *Danio rerio*, were compared. The homologous sequences obtained were used to get the best predictions. To localize non-transcribed elements in the genome, RNA structure prediction tools were applied too. Finally, all available information was used to create

plausible models and, whenever possible, functional information was added. Using the genomic editor Apollo (Lewis et al. 2002), Signal map software (Roche Applied Science, Switzerland) and Geneious basic 5.6.5 (Biomatters Ltd.), the results were individually tested and adjusted in the final edition process.

Molecular evolutionary analysis

The histone amino acid sequences were obtained from the annotated contigs using the Geneious 8.1.5 program (Biomatters Ltd.). The sequences were aligned in MAFFT (Katoh et al. 2002) using an iterative method. The nucleotide variability (π) within species and types of sequences were estimated in both coding and spacer sequences as the average number of nucleotide differences per site between two sequences, and was performed in the DnaSP v5 program (Librado and Rozas 2009). In a similar way, nucleotide divergence between species (Dxy) was estimated as the average number of nucleotide substitutions per site between species. In both cases the Jukes and Cantor correction was used. The amino acid variability was estimated using the MEGA6 program (Tamura et al. 2013) following the Poisson's model.

A Maximum-Likelihood tree was inferred in the PhyML 3.0 program (Guindon et al. 2010), using a Bio Neighbor-Joining (BioNJ) as starting tree. Previously, a statistical selection of best-fit models of protein evolution was performed using the ProtTest version 2.4 program (Abascal et al. 2005) with the corrected Akaike Information Criterion (AICc) used for model selection. The DCMut model (Kosiol and Goldman 2005) was chosen as the best using the PhyML 3.0 program ($-\ln L = -1.714.173.839$, $AICc = 3.474.391.314$). The branch support of the ML tree was measured by applying the non-parametric version of the approximate likelihood ratio test (SH-aLRT) (Anisimova et al. 2011), and the SPR option was also chosen for tree improvement. Finally, the tree edition was carried out in the MEGA6 program (Tamura et al. 2013).

The analysis of synteny was carried out using the CIRCOS software (Krzywinski et al. 2009). The comparison was performed from the ENSEMBL database of the stickleback (*Gasterosteus aculeatus*), medaka (*Oryzias latipes*) and zebrafish (*D. rerio*).

FISH techniques

Chromosome preparations were made from *S. senegalensis* larvae aged from 1 to 3 days. The specimens were pre-treated with 0.02% colchicine for 3 h, subjected to hypotonic shock and finally fixed in a freshly-made Carnoy solution (absolute ethanol-acetic acid (3:1)). Larvae were homogenized in Carnoy, and the preparations were dropped onto wet slides by splashing on a hot plate with damp paper.

The BAC-DNA was used as probe; the labeling was done with a first amplification by DOP-PCR followed by a conventional PCR for labeling, as described previously in Liehr et al. (2009). Three different fluorochromes were used, i.e. Texas Red (Thermo Fisher Scientific, USA), Fluorescein-Isothiocyanate (FITC) (Enzo, USA) and diethyl-aminocoumarin (DEAC) (Vysis, USA). Finally, the probes were precipitated using a standard protocol with sodium acetate and ethanol.

Chromosome preparations were pre-treated with pepsin solution at 37°C and fixed with paraformaldehyde solution. Finally, the preparations were dehydrated with ethanol series of 70%, 90% and 100%, and air-dried before hybridization. The hybridization was carried out by denaturation of the probes and chromosome preparations in parallel, following the protocol described by Liehr et al. (2009) with minor modifications.

The post-hybridization treatment consisted of serial washes of SSC, Tween20 and PBS. The preparations were then dehydrated with ethanol and counterstained with antifade-DAPI solution (Vector Labs, USA). Hybridization images were obtained with a digital CCD camera (Olympus DP70) coupled to a fluorescence microscope (Olympus BX51 and/or Zeiss Axioplan using software of MetaSystems, Altlussheim, Germany).

Results

Histone sequences and arrangement

During the analysis of the three BAC clones (BAC5K5, BAC12D22 and BAC30J4) with histone genes we found up to 28 histone genes distributed among four different contigs (Table 2). Assembled BAC contigs were deposited in the GenBank database (NCBI) under accession numbers AC275286-AC275288. There are two main BAC clones with all five histones, and the other BAC clone contains only two histone genes. The analysis with the BLAST tool demonstrated that histones *H1* and *H3* from BAC30J4 are variants different from those of the other BAC clones. While BAC30J4 contains the Replication Independent (RI) variant of *H1* (*H1.0*) and the *H3.3* variant, the other two BAC clones have Replication Dependent (RD) variants of *H1* and *H3* (Fig. 1). However, fixed differences were also found between histones from BAC5K5 and BAC12D22, at both nucleotide level and amino acid level (Table 3). Although histone genes are characterized by not having intron sequences, we have found that the *H3.3* histone of the BAC30J4 has two intron sequences of 89 bp and 83 bp respectively, and the *H1.0* another two introns of 79 and 114 bp (data not shown). The most variable histone was *H1*, since three isoforms of RD variants were found (Fig. 1A), whereas two isoforms of *H2A* and *H2B* were found, each obtained from the two main histone-BAC clones (Fig. S1). The *H3* RD variant obtained differs in one amino acid with respect to the *H3.2* variant from other species (Fig. 1B), but the position of this variation is different from the one that differentiates between *H3.1* and *H3.2* in other vertebrates (Fig. 1C).

It was possible to determine two main histone cluster arrangements from contigs 634, 641 and 824. The contig 634 showed an arrangement lacking the *H1* histone (Fig. 2A); contig 641 showed an arrangement with the *H1* histone (Fig. 2B); and contig 824 showed a mixed arrangement with and without the *H1* histone (Fig. 2C). Moreover, we were able

to deduce an inversion event from contig 824. The arrangement lacking *H1* is composed sequentially by *H2B* (direct), *H2A* (reverse), *H3* (direct) and *H4* (direct). In the other arrangement, *H1* is situated between *H2B* and *H2A* in direct sense.

Phylogenetic analysis

The phylogenetic results showed that clades were formed by histone types rather than by species. One clade is characterized by the RI variants, and the second by the RD variants. Each variant of histone *H1* found in *S. senegalensis* robustly grouped in different clades with the rest of vertebrate species (Fig 3). In both cases, *S. senegalensis* was clustered with the most modern fish species. Moreover, the *H1* sequences of the two BAC clones with the RD variant grouped each separately. The invertebrate species were divided in two groups, one associated with the *H1* RI variant and the other with the *H1* RD variant.

The *H3* phylogenetic tree also showed two main clades, each belonging to the two different H3 variants, i.e., *H3* RD and *H3.3* variants (Fig. 4). The *H3* histone found within BAC30J4 grouped within the *H3.3* clade, and the other two BAC clones clustered within the *H3.2* clade. The genes from BAC5K5 and BAC12D22 each again clustered in two robust groups. The main species groups showed a more paraphyletic behavior within the *H3.3* clade than in the *H3.2* clade.

Analysis of synteny

Syntenic comparison showed that the genes of the three BAC clones were distributed in more than three chromosomes in the analyzed species (Fig. 5). The 59 genes found among the three BAC clones (Table S1) were distributed in 5 chromosomes in stickleback (*Gasterosteus aculeatus*) (Fig. 5A), in 6 chromosomes in medaka (*Oryzias latipes*) (Fig. 5B), and in 8 chromosomes in zebrafish (*Danio rerio*) (Fig. 5C). In the latter three cases the main histone cluster was found in one chromosome, but the other histone variants (*H1.0* and *H3.3*) were found in chromosomes different from that of the

main cluster. These variants were found in the same chromosome in stickleback, but in zebrafish were in several different chromosomes, and in medaka the *H3.3* variant did not show any hit when it was compared to the ENSEMBL database.

FISH analysis

The karyotype of Senegalese sole has been described in previous work; it comprises 21 chromosomes pairs, distributed in 3 metacentric pairs, 2 submetacentric pairs, 4 subtelocentric pairs and 12 acrocentric pairs (Vega et al. 2002). The three BAC probes gave four signals distributed among three chromosome pairs (Fig. 6). The BAC5K5 (*H1*, *H2A*, *H2B*, *H3.2*, *H4*) and BAC12D22 (*H1*, *H2A*, *H2B*, *H3*, *H4*) probes both hybridized in the largest metacentric chromosome pair, but in different positions. The first was localized in the telomeric position, and the second in the subcentromeric position of the opposite arm (Fig. 6A). In some cases, additional minor signals could be seen in other chromosomes. Moreover, the BAC30J4 (*H1.0*, *H3.3*) probe showed two hybridization signals in two different chromosome pairs: the stronger signal was found in a submetacentric pair in an interstitial position of the q arm, and the weaker signal was localized in an acrocentric pair in a telomeric site (Fig. 6B). The largest metacentric chromosome has been considered a sex proto-chromosome in *S. senegalensis* because several genes of the *dmrt* gene family, notably the *dmrt1* gene, have been located on that chromosome (Portela-Bens et al. 2016); hence multiple FISH (mFISH) was carried out. Results confirmed the co-localization of BAC5K5, BAC12D22 and BAC11O20 harboring the *dmrt1* gene (Portela-Bens et al. 2016) on the same chromosome (Fig. 6C).

Discussion

In the present work, the main histone cluster was found in two BAC clones (BAC5K5 and BAC12D22) localized in the same chromosome pair, but in different regions. Both BAC clones contain the five canonical histones. However, two replacement histones (*H1.0* and *H3.3*) were found in another BAC clone (BAC30J4) situated in a different

chromosome pair. Different subtypes of histones show a varied expression pattern depending on the tissue and cell line, such as the spermatogenesis cell line, which undergoes a high chromatin reorganization (Doenecke et al. 1997). However, knockout experiments have demonstrated that different histone subtypes are not essential for the organism's survival; instead, their role is more subtle (Shen et al. 1995; Sirotkin et al. 1995; Yu and Gorovsky 1997).

The different localization for the *H1* RD and *H1* RI has already been observed in human and mouse (Wang et al. 1997). Some authors have interpreted this separation as a way to escape from the homogenizing factors acting over the family members, and the separated histones then become “orphan” genes (Eirín-López et al. 2002). This separation could facilitate control of the differential expression between RD and RI *H1* histone genes (Albig et al. 1993). While the *H1* RD genes are expressed in every type of cell from the early developmental stage, the *H1* RI are expressed only in oocytes, testes or cells in terminal stage of development (Khochbin and Wolffe 1994). Furthermore, the *RI H3.3* histone has been demonstrated to have a high expression in mouse pro-spermatogonial cells and it has been associated with the genome-wide epigenetic reformatting that those cells suffer (Tang et al. 2014). The main signal observed for the *H3.3* histone gene is in the same chromosome pair as that of the *vasa* gene (Fig. S2). This last gene is also over-expressed in primordial germ cells (PGC) and spermatogonias, thus being associated with the development of the germinal line in several species, including *S. senegalensis* (Pacchiarini et al. 2013). The co-localization of *vasa* and *H3.3* genes could represent an improvement by synchronizing the expression of the two genes.

The canonical histones have shown some differences in *S. senegalensis*, and these differences have been detected between the histones present within the two BACs localized in the same chromosome pair. Therefore, the histones of each BAC must have evolved in a different way. Phylogenetic analysis demonstrated that the different *H1*, *H2A* and *H2B* types were all more similar to each other than to canonical histones of other species. Therefore, different types of canonical *H2A* and *H2B* have been found in

one species for the first time. The residue 97 of the amino acid sequence of canonical *H3* serves as marker to differentiate between *H3.1* and *H3.2* subtypes in mammals, in which a cysteine residue is present in the *H3.1* subtype, and a serine is present in the *H3.2* subtype (Hake and Allis 2006). The canonical *H3* of *S. senegalensis* differs from these two subtypes in that, in the residue 96, it has a polymorphism in which a serine residue is found instead of an alanine. Meanwhile, the residue 97 of the Senegalese sole (serine) is similar to that of the *H3.2* subtype in humans. The same polymorphism as in *S. senegalensis* has been found only in *D. rerio* (Acc. No. NP_001104686.1), so this could be a new canonical subtype common to these fish species. This amino acid substitution is conservative (alanine-to-serine substitution), so it is probable that the functionality has not been lost.

Attending to the chromosomal localization, the two BACs are in the same chromosome pair, but in different localizations, so the two loci should be genetically isolated. The most conservative situation in fish species is for one signal to be localized on the chromosome pair bearing the histone cluster (Costa et al. 2016). However, in some fish species two or more loci have been mapped (Hashimoto et al. 2011; Lima-Filho et al. 2012; Silva et al. 2014; Daniel et al. 2016), and a dispersed pattern has even been seen in *Rachycentron canadum* (Costa et al. 2014) and in *Ocyurus chrysurus* (Costa et al. 2016).

Duplication or transposable element-mediated mechanisms have been proposed to explain the occurrence of multiple histone loci (Costa et al. 2014). Although two loci at the same chromosome pair have been detected in the clam *Venerupis aurea* (Carrilho et al. 2011), in fish species this situation has not been described until this work. In *S. senegalensis*, an inversion or transposition event could account for the existence of these two clusters, since it has been postulated that the different gene orientations observed within a DNA strand in evolutionary distinct species is commonly caused by transposition and inversion events (Eirín-López et al. 2004). In humans, the main histone cluster is localized in chromosome 6, but two smaller clusters have also been observed in

a different chromosome pair (Pineau et al. 2005), thus indicating that the histone cluster transposition is a common event in vertebrates.

The two BAC clones that contain the clusters of canonical histones have both been mapped in the largest metacentric pair, and mFISH confirmed the co-localization with the *dmrt1* gene in that pair. This chromosome has been studied recently (Portela-Bens et al. 2016), and those researchers suggest that it originated from a Robertsonian fusion of two acrocentric pairs. *S. senegalensis* has $2n=42$ chromosomes with three metacentric pairs, similar to the results obtained in other congeneric species such as *S. solea* and *S. lascaris*, which showed $2n=42$ chromosomes with four metacentric pairs (Pardo et al. 2001). This situation is different from that observed in the two closely-related species *Scophthalmus maximus* and *S. rhombus*, which showed a karyotype with $2n=44$ chromosomes with two metacentric pairs (Bouza et al. 1994; Pardo et al. 2001). A Robertsonian fusion is proposed as a plausible event that could have occurred in the Pleuronectiformes group; indeed it has already been demonstrated that such an event occurred in the Pleuronectiformes species *Trinectes inscriptus*, as revealed by traces of internal telomere sequences (Bitencourt et al. 2014). Taking into account this possibility, the two histone clusters could have come from two acrocentric pairs, in which the histone clusters evolved separately and accumulated polymorphisms. The majority of the fish species analyzed with the histone clusters showed a centromeric localization on acrocentric pairs or a terminal position on subtelocentric pairs (Hashimoto et al. 2011, 2013; Lima-Filho et al. 2012; Costa et al. 2014, 2016; Silva et al. 2014; Daniel et al. 2016).

When the nucleotide or amino acid sequences are compared to those from other fish species, it can be seen that the histones of BAC5K5 present more divergence than those of BAC12D22. However, the differences can be detected at both nucleotide and amino acid levels in histones H1, H2A and H2B, but only at nucleotide level in histones H3 and H4. Of all the histones, H1 is the one that evolves fastest (Eirín-López et al. 2002) and, in vertebrates, it has been estimated that the most conserved H1 subtype (H1e) evolves twice as fast as the subtypes of H2A and H2B (Ponte et al. 1998) and these two evolve up

to 10 times faster than those of H3 and H4 (Thatcher and Gorovsky 1994). The variability data obtained in *S. senegalensis* confirms these wide differences in the rate of evolution.

The cluster organization of the canonical histones has been extensively discussed for many organisms (Hentschel and Birnstiel 1981; Eirín-López et al. 2009). An inversion event was observed in one of the contigs examined, and it has been reported that inversions occur within histone clusters caused by the recombination of two duplexes oriented in opposite directions (Vitelli and Weinberg 1983; D'Andrea et al. 1985). These inversion events could have caused the non-conservative ordering of the histone gene clusters and the occurrence of two histone clusters. As a conclusion, it could be said that there is no consensus with respect to the histone cluster arrangement, so the expression of the histones is positioning-independent (D'Andrea et al. 1985). However, a general rule could be deduced, since *H3* and *H4* gene histones are normally found in neighboring position and in the same direction, whereas *H2A* and *H2B* are normally found joined, but in opposing orientation (D'Andrea et al. 1985). Both rules are present in the histone organization of *S. senegalensis*.

The evolutionary model that drives the histone multigene family has also been debated at length (reviewed by Nei and Rooney (2005)). Currently, the accepted model for the histone family is the birth-and-death model with strong purifying selection. Under this model of evolution, duplication events generate new genes, which may be maintained by selection or may degenerate into pseudogenes (Eirín-López et al. 2004, 2012). Moreover, in *S. senegalensis*, while variability was observed at nucleotide level, scant or no variability was detected at amino acid level, thus indicating that the type of nucleotide variation is mainly synonymous. A higher level of synonymous rather than non-synonymous variations is indicative of a birth-and-death model with strong purifying selection (Nei and Rooney 2005).

As can be deduced from phylogenetic results, a birth-and-death evolution is the most probable model for the histones of Senegalese sole, since the different histone genes are

grouped by types rather than by species. Peretti and Khochbin (1997) suggested that the RD H1 of vertebrate species appeared after invertebrate-vertebrate separation, based on the sequence analyses of the promoter regions, which supports the closer relationship of the RI H1 of vertebrates with the H1 of invertebrates than with the RD H1 of vertebrates. Notwithstanding, in our phylogenetic results, the H1 histone of invertebrate species is grouped either with the RD or the RI H1 of vertebrate species, so the divergence between RD and RI H1 histones might have occurred before the separation of vertebrates and invertebrates 815 MYA (Feng et al. 1997). An ancestor of all H1 subtypes might have undergone evolutionary changes which have led to the current RD subtypes on one hand and the RI subtypes on the other hand (Eirín-López et al. 2002). This argument can also be extended to H3 variants, but the separation between vertebrates and invertebrates in the two variants is not so clear as with those of H1, possibly because the H3 histone has a slower rate of variation than H1, as discussed before.

The analysis of synteny showed that a greater number of chromosome rearrangements have occurred between *S. senegalensis* and *D. rerio* than between the former and the other two species. Between *S. senegalensis* and *D. rerio* the common ancestor is older than that between the former and *O. latipes* or *G. aculeatus*, since *D. rerio* belongs to a different cohort (Otomorpha) than the other three species (Euteleosteiomorpha) (Portela-Bens et al. 2016). In all three cases the RD and RI variants have been found in separated chromosomes, which is the common situation for the species studied to date, thus supporting the idea that the “orphan” RI genes have emerged from canonical histones by transposition to other genome regions (Eirín-López et al. 2009). However a new scenario has recently been opened, in which histone variants also have an ancestral nature as canonical histones (Rivera-Casas et al. 2016).

Three chromosome markers have been identified in this work, which can be added to those obtained with other BAC clones (Ponce et al. 2011; García-Cegarra et al. 2013; Portela-Bens et al. 2016), or with ribosomal families (Cross et al. 2006; Manchado et al. 2006b). All these markers account for 18 chromosome pairs with one or more markers, so

only 3 chromosome pairs remain without markers. In addition, this study has highlighted higher order re-arrangements that must have taken place during Pleuronectiformes evolution to finally configure the karyotype of *S. senegalensis*. The histone cluster could have mediated in such a Robertsonian re-arrangement, so the analysis of this cluster in other closely-related species would be essential to determine this issue definitively.

Acknowledgements

This study has been supported by the Spanish Ministerio de Ciencia e Innovación MICINN-FEDER (AGL 2011-25596 and AGL2014-51860-C2-1-P). The authors express their thanks to the Laboratorio de Cultivos Marinos of the UCA for its support, and to José Ramón Aracama for technical assistance. Roger Iziga received a fellowship from the AECL.

References

- Abascal, F., Zardoya, R., and Posada, D. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**(9): 2104–2105.
- Albig, W., Drabent, B., Kunz, J., Kalff-Suske, M., Grzeschik, K.-H., and Doenecke, D. 1993. All Known Human H1 Histone Genes Except the H10 Gene Are Clustered on Chromosome 6. *Genomics* **16**(3): 649–654. doi:10.1006/geno.1993.1243.
- Albig, W., Kardalidou, E., Drabent, B., Zimmer, A., and Doenecke, D. 1991. Isolation and characterization of two human H1 histone genes within clusters of core histone genes. *Genomics* **10**(4): 940–948. doi:10.1016/0888-7543(91)90183-F.
- Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C., and Gascuel, O. 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* **60**(5): 685–699.
- Armbruster, G.F.J., Böhme, M., Bernhard, D., and Schlegel, M. 2005. The H3/H4 histone gene cluster of land snails (Gastropoda: Stylommatophora): TS/TV Ratio,

GC3 drive and signals in stylommatophoran phylogeny. *J. Molluscan Stud.* **71**(4): 339–348. doi:10.1093/mollus/eyi038.

Asakawa, S., Abe, I., Kudoh, Y., Kishi, N., Wang, Y., Kubota, R., Kudoh, J., Kawasaki, K., Minoshima, S., and Shimizu, N. 1997. Human BAC library: construction and rapid screening. *Gene* **191**(1): 69–79. doi:10.1016/S0378-1119(97)00044-9.

Benzekri, H., Armesto, P., Cousin, X., Rovira, M., Crespo, D., Merlo, M., Mazurais, D., Bautista, R., Guerrero-Fernández, D., Fernandez-Pozo, N., Ponce, M., Infante, C., Zambonino, J., Nidelet, S., Gut, M., Rebordinos, L., Planas, J. V., Bégout, M.-L., Claros, M., and Manchado, M. 2014. De novo assembly, characterization and functional annotation of Senegalese sole (*Solea senegalensis*) and common sole (*Solea solea*) transcriptomes: integration in a database and design of a microarray. *BMC Genomics* **15**(1): 952. doi:10.1186/1471-2164-15-952.

Betancur-R, R., Li, C., Munroe, T.A., Ballesteros, J.A., and Ortí, G. 2013. Addressing gene tree discordance and non-stationarity to resolve a multi-locus phylogeny of the flatfishes (Teleostei: Pleuronectiformes). *Syst. Biol.* **62**(5): 763–85. doi:10.1093/sysbio/syt039.

Bitencourt, J.A., Sampaio, I., Ramos, R.T.C., and Affonso, P.R.A.M. 2014. Chromosomal fusion in Brazilian populations of *Trinectes inscriptus* Gosse, 1851 (Pleuronectiformes; Achiridae) as revealed by internal telomere sequences (ITS). *J. Exp. Mar. Biol. Ecol.* **452**: 101–104. doi:10.1016/j.jembe.2013.12.012.

Bouza, C., Sanchez, L., and Martínez, P. 1994. Karyotypic characterization of turbot (*Scophthalmus maximus*) with conventional, fluorochrome and restriction endonuclease banding techniques. *Mar. Biol.* **120**: 609–613.

Carrilho, J., Pérez-García, C., Leitão, A., Malheiro, I., and Pasantes, J.J. 2011. Cytogenetic characterization and mapping of rDNAs, core histone genes and telomeric sequences in *Venerupis aurea* and *Tapes rhomboides* (Bivalvia: Veneridae). *Genetica* **139**(6): 823–831. doi:10.1007/s10709-011-9587-8.

Cole, R.D. 1987. Microheterogeneity in H1 histones and its consequences. *Int. J. Pept. Protein Res.* **30**: 433–449.

- Combosch, D.J., and Giribet, G. 2016. Clarifying phylogenetic relationships and the evolutionary history of the bivalve order Arcida (Mollusca: Bivalvia: Pteriomorphia). *Mol. Phylogenet. Evol.* **94**: 298–312. doi:10.1016/j.ympev.2015.09.016.
- Costa, G.W.W.F., Cioffi, M.B., Bertollo, L.A.C., and Molina, W.F. 2014. Unusual Dispersion of Histone Repeats on the Whole Chromosomal Complement and Their Colocalization with Ribosomal Genes in *Rachycentron canadum* (Rachycentridae, Perciformes). *Cytogenet. Genome Res.* **144**(1): 62–67. doi:10.1159/000366301.
- Costa, G.W.W.F.D., Cioffi, M.D.B., Bertollo, L.A.C., and Molina, W.F. 2016. The Evolutionary Dynamics of Ribosomal Genes, Histone H3, and Transposable Rex Elements in the Genome of Atlantic Snappers. *J. Hered.* **107**(2): 173–180.
- Cross, I., Merlo, A., Manchado, M., Infante, C., Cañavate, J.P., and Rebordinos, L. 2006. Cytogenetic characterization of the sole *Solea senegalensis* (Teleostei: Pleuronectiformes: Soleidae): Ag-NOR, (GATA) n , (TTAGGG) n and ribosomal genes by one-color and two-color FISH. *Genetica* **128**(1-3): 253–259. doi:10.1007/s10709-005-5928-9.
- D’Andrea, R.J., Coles, L.S., Lesnikowski, C., Tabe, L., and Wells, J.R. 1985. Chromosomal organization of chicken histone genes: preferred associations and inverted duplications. *Mol. Cell. Biol.* **5**(11): 3108–3115. doi:10.1128/MCB.5.11.3108.
- Daniel, S.N., Penitente, M., Silva, D.M.Z.A., Hashimoto, D.T., Ferreira, D.C., Foresti, F., and Porto-Foresti, F. 2015. Organization and Chromosomal Distribution of Histone Genes and Transposable Rex Elements in the Genome of *Astyanax bockmanni* (Teleostei, Characiformes). *Cytogenet. Genome Res.* **146**(4): 311–318. doi:10.1159/000441613.
- Díaz-Ferguson, E., Cross, I., Del Barrios, M.M., and Rebordinos, L. 2007. Genetic Relationships among Populations of the Senegalese Sole *Solea senegalensis* in the Southwestern Iberian Peninsula Detected by Mitochondrial DNA–Restriction Fragment Length Polymorphisms. *Trans. Am. Fish. Soc.* **136**(2): 484–491. doi:10.1577/T06-030.1.

- Doenecke, D., Albig, W., Bode, C., Drabent, B., Franke, K., Gavenis, K., and Witt, O. 1997. Histones: genetic diversity and tissue-specific gene expression. *Histochem. Cell Biol.* **107**(1): 1–10. doi:10.1007/s004180050083.
- Eick, S., Nicolai, M., Mumberg, D., and Doenecke, D. 1989. Human H1 histones: Conserved and varied sequence elements in two H1 subtype genes. *Eur. J. Cell Biol.* **49**: 110–115.
- Eirín-López, J.M., González-Romero, R., Dryhurst, D., Méndez, J., and Ausió, J. 2009. Long-Term Evolution of Histone Families: Old Notions and New Insights into Their Mechanisms of Diversification Across Eukaryotes. *In Evolutionary Biology: Concept, Modeling, and Application. Edited by P. Pontarotti.* Springer Verlag Berlin, Heidelberg. pp. 139–162. doi:10.1007/978-3-642-00952-5_8.
- Eirín-López, J.M., González-Tizón, A.M., Martínez, A., and Méndez, J. 2002. Molecular and Evolutionary Analysis of Mussel Histone Genes (*Mytilus* spp.): Possible Evidence of an “Orphon Origin” for H1 Histone Genes. *J. Mol. Evol.* **55**(3): 272–283. doi:10.1007/s00239-002-2325-1.
- Eirín-López, J.M., González-Tizón, A.M., Martínez, A., and Méndez, J. 2004. Birth-and-death evolution with strong purifying selection in the histone H1 multigene family and the origin of orphon H1 genes. *Mol. Biol. Evol.* **21**(10): 1992–2003. doi:10.1093/molbev/msh213.
- Eirín-López, J.M., Rebordinos, L., Rooney, A.P., Rozas, J. 2012. The Birth- and-Death Evolution of Multigene Families Revisited. In Garrido- Ramos MA (ed): Repetitive DNA. *Genome Dyn.* Basel, Karger, **7**, pp 170–196
- Feng, D.-F., Cho, G., and Doolittle, R.F. 1997. Determining divergence times with a protein clock: Update and reevaluation. *Proc. Natl. Acad. Sci.* **94**(24): 13028–13033. doi:10.1073/pnas.94.24.13028.
- García-Cegarra, A., Merlo, M.A., Ponce, M., Portela-Bens, S., Cross, I., Manchado, M., and Rebordinos, L. 2013. A Preliminary Genetic Map in *Solea senegalensis* (Pleuronectiformes, Soleidae) Using BAC-FISH and Next-Generation Sequencing. *Cytogenet. Genome Res.* **141**(2-3): 227–240. doi:10.1159/000355001.

- González-Romero, R., Ausió, J., Méndez, J., and Eirín-López, J.M. 2008. Early Evolution of Histone Genes: Prevalence of an “Orphon” H1 Lineage in Protostomes and Birth-and-Death Process in the H2A Family. *J. Mol. Evol.* **66**(5): 505–518. doi:10.1007/s00239-008-9109-1.
- Grunstein, M., Schedl, P., and Kedes, L. 1976. Sequence analysis and evolution of sea urchin (*Lytechinus pictus* and *Strongylocentrotus purpuratus*) histone H4 messenger RNAs. *J. Mol. Biol.* **104**(2): 351–369. doi:10.1016/0022-2836(76)90276-X.
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **59**(3): 307–321.
- Guzmán, J.M., Rubio, M., Ortiz-Delgado, J.B., Klenke, U., Kight, K., Cross, I., Sánchez-Ramos, I., Riaza, A., Rebordinos, L., Sarasquete, C., Zohar, Y., and Mañanós, E.L. 2009. Comparative gene expression of gonadotropins (FSH and LH) and peptide levels of gonadotropin-releasing hormones (GnRHs) in the pituitary of wild and cultured Senegalese sole (*Solea senegalensis*) broodstocks. *Comp. Biochem. Physiol. Part A Mol. Integr. Physiol.* **153**(3): 266–277. doi:10.1016/j.cbpa.2009.02.032.
- Hake, S.B., and Allis, C.D. 2006. Histone H3 variants and their potential role in indexing mammalian genomes: The “H3 barcode hypothesis”. *Proc. Natl. Acad. Sci.* **103**(17): 6428–6435. doi:10.1073/pnas.0600803103.
- Hashimoto, D.T., Ferguson-Smith, M.A., Rens, W., Foresti, F., and Porto-Foresti, F. 2011. Chromosome Mapping of H1 Histone and 5S rRNA Gene Clusters in Three Species of *Astyanax* (Teleostei, Characiformes). *Cytogenet. Genome Res.* **134**(1): 64–71. doi:10.1159/000323512.
- Hashimoto, D.T., Ferguson-Smith, M.A., Rens, W., Prado, F.D., Foresti, F., and Porto-Foresti, F. 2013. Cytogenetic Mapping of H1 Histone and Ribosomal RNA Genes in Hybrids between Catfish Species *Pseudoplatystoma corruscans* and *Pseudoplatystoma reticulatum*. *Cytogenet. Genome Res.* **139**(2): 102–106. doi:10.1159/000345299.

- Hentschel, C.C., and Birnstiel, M.L. 1981. The organization and expression of histone gene families. *Cell* **25**(2): 301–313. doi:10.1016/0092-8674(81)90048-9.
- Iziga, R., Ponce, M., Infante, C., Rebordinos, L., Cañavate, J.P., and Manchado, M. 2010. Molecular characterization and gene expression of thyrotropin-releasing hormone in Senegalese sole (*Solea senegalensis*). *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* **157**(2): 167–174. doi:10.1016/j.cbpb.2010.05.013.
- Katoh, K., Misawa, K., Kuma, K.-I., and Miyata, T. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**(14): 3059–3066.
- Khochbin, S., and Wolffe, A.P. 1994. Developmentally regulated expression of linker-histone variants in vertebrates. *Eur. J. Biochem.* **225**(2): 501–510. doi:10.1111/j.1432-1033.1994.00501.x.
- Kosiol, C., and Goldman, N. 2005. Different Versions of the Dayhoff Rate Matrix. *Mol. Biol. Evol.* **22**(2): 193–199.
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. 2009. Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**(9): 1639–1645. doi:10.1101/gr.092759.109.
- Lewis, S.E., Searle, S.M.J., Harris, N., Gibson, M., Iyer, V., Richter, J., Wiel, C., Bayraktaroglu, L., Birney, E., Crosby, M.A., Kaminker, J.S., Matthews, B.B., Prochnik, S.E., Smith, C.D., Tupy, J.L., Rubin, G.M., Misra, S., Mungall C.J., and Clamp, M.E. 2002. Apollo: a sequence annotation editor. *Genome Biol.* **3**(12): RESEARCH0082.
- Librado, P., and Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**(11): 1451–1452.
- Liehr, T. (Editor). 2009. Fluorescence in situ hybridization—application guide. Springer-Verlag, Berlin.
- Lima-Filho, P.A., Cioffi, M.B., Bertollo, L.A.C., and Molina, W.F. 2012. Chromosomal and morphological divergences in Atlantic populations of the frillfin *Bathygobius soporator* (Gobiidae, Perciformes). *J. Exp. Mar. Bio. Ecol.* **434**: 63–70.

- Mai, M.G., Engrola, S., Morais, S., Portella, M.C., Verani, J.R., Dinis, M.T., and Conceição, L.E.C. 2009. Co-feeding of live feed and inert diet from first-feeding affects *Artemia* lipid digestibility and retention in Senegalese sole (*Solea senegalensis*) larvae. *Aquaculture* **296**(3-4): 284–291. doi:10.1016/j.aquaculture.2009.08.024.
- Manchado, M., Infante, C., Rebordinos, L., and Cañavate, J.P. 2009. Molecular characterization, gene expression and transcriptional regulation of thyroid hormone receptors in Senegalese sole. *Gen. Comp. Endocrinol.* **160**(2): 139–147. doi:10.1016/j.ygcen.2008.11.001.
- Manchado, M., Rebordinos, L., and Infante, C. 2006a. U1 and U2 Small Nuclear RNA Genetic Linkage: A Novel Molecular Tool for Identification of Six Sole Species (Soleidae, Pleuronectiformes). *J. Agric. Food Chem.* **54**(11): 3765–3767. doi:10.1021/jf0530594.
- Manchado, M., Zuasti, E., Cross, I., Merlo, A., Infante, C., and Rebordinos, L. 2006b. Molecular characterization and chromosomal mapping of the 5S rRNA gene in *Solea senegalensis*: a new linkage to the U1, U2, and U5 small nuclear RNA genes. *Genome* **49**(1): 79–86. doi:10.1139/g05-068.
- Morais, S., Aragão, C., Cabrita, E., Conceição, L.E., Constenla, M., Costas, B., Dias, J., Duncan, N., Engrola, S., Estevez, A., Gisbert, E., Mañanós, E., Valente, L.M., Yúfera, M., and Dinis, M.T. 2014. New developments and biological insights into the farming of solea senegalensis reinforcing its aquaculture potential. *Rev. Aquac.* **6**(1): 1–37.
- Nei, M., and Rooney, A.P. 2005. Concerted and Birth-and-Death Evolution of Multigene Families. *Annu. Rev. Genet.* **39**(1): 121–152. doi:10.1146/annurev.genet.39.073003.112240.
- Oliveira, C., Duncan, N.J., Pousão-Ferreira, P., Mañanós, E., and Sánchez-Vázquez, F.J. 2010. Influence of the lunar cycle on plasma melatonin, vitellogenin and sex steroids rhythms in Senegal sole, *Solea senegalensis*. *Aquaculture* **306**(1-4): 343–347. doi:10.1016/j.aquaculture.2010.05.003.

- Oskars, T.R., Bouchet, P., and Malaquias, M.A.E. 2015. A new phylogeny of the Cephalaspidea (Gastropoda: Heterobranchia) based on expanded taxon sampling and gene markers. *Mol. Phylogenet. Evol.* **89**: 130–150. doi:10.1016/j.ympev.2015.04.011.
- Pacchiarini, T., Cross, I., Leite, R.B., Gavaia, P., Ortiz-Delgado, J.B., Pousão-Ferreira, P., Rebordinos, L., Sarasquete, C., and Cabrita, E. 2013. *Solea senegalensis* vasa transcripts: molecular characterisation, tissue distribution and developmental expression profiles. *Reprod. Fertil. Dev.* **25**(4): 646. doi:10.1071/RD11240.
- Padrós, F., Villalta, M., Gisbert, E., and Estévez, A. 2011. Morphological and histological study of larval development of the Senegal sole *Solea senegalensis*: an integrative study. *J. Fish Biol.* **79**(1): 3–32. doi:10.1111/j.1095-8649.2011.02942.x.
- Pardo, B.G., Bouza, C., Castro, J., Martínez, P., and Sánchez, L. 2001. Localization of ribosomal genes in Pleuronectiformes using Ag-, CMA₃-banding and in situ hybridization. *Heredity* **86**: 531–536.
- Pardo, B.G., Machordom, A., Foresti, F., Porto-Foresti, F., Azevedo, M.F.C., Bañón, R., Sánchez, L., and Martínez, P. 2005. Phylogenetic analysis of flatfish (order Pleuronectiformes) based on mitochondrial 16S rDNA sequences. *Sci. Mar.* **69**(4): 531–543.
- Peretti, M., and Khochbin, S. 1997. The Evolution of the Differentiation-Specific Histone H1 Gene Basal Promoter. *J. Mol. Evol.* **44**(2): 128–134. doi:10.1007/PL00006129.
- Pineau, P., Henry, M., Suspène, R., Marchio, A., Dettai, A., Debruyne, R., Petit, T., Lécu, A., Moisson, P., Dejean, A., Wain-Hobson, S., and Vartanian, J.-P. 2005. A universal primer set for PCR amplification of nuclear histone H4 genes from all animal species. *Mol. Biol. Evol.* **22**(3): 582–8. doi:10.1093/molbev/msi053.
- Ponce, M., Salas-Leiton, E., Garcia-Cegarra, A., Boglino, A., Coste, O., Infante, C., Gisbert, E., Rebordinos, L., and Manchado, M. 2011. Genomic characterization, phylogeny and gene regulation of g-type lysozyme in sole (*Solea senegalensis*). *Fish Shellfish Immunol.* **31**(6): 925–937. doi:10.1016/j.fsi.2011.08.010.

- Ponte, I., Vidal-Taboada, J.M., and Suau, P. 1998. Evolution of the vertebrate H1 histone class: evidence for the functional differentiation of the subtypes. *Mol. Biol. Evol.* **15**(6): 702–708. doi:10.1093/oxfordjournals.molbev.a025973.
- Portela-Bens, S., Merlo, M.A., Rodríguez, M.E., Cross, I., Manchado, M., Kosyakova, N., Liehr, T., and Rebordinos, L. 2016. Integrated gene mapping and synteny studies give insights into the evolution of a sex proto-chromosome in *Solea senegalensis*. *Chromosoma*. doi:10.1007/s00412-016-0589-2.
- Rivera-Casas, C., Gonzalez-Romero, R., Cheema, M.S., Ausió, J., and Eirín-López, J.M. 2016. The characterization of macroH2A beyond vertebrates supports an ancestral origin and conserved role for histone variants in chromatin. *Epigenetics*: 1–11. doi:10.1080/15592294.2016.1172161.
- Shen, X., Yu, L., Weir, J.W., and Gorovsky, M.A. 1995. Linker histories are not essential and affect chromatin condensation in vivo. *Cell* **82**(1): 47–56. doi:10.1016/0092-8674(95)90051-9.
- Silva, D.M.Z. de A., Pansonato-Alves, J.C., Utsunomia, R., Araya-Jaime, C., Ruiz-Ruano, F.J., Daniel, S.N., Hashimoto, D.T., Oliveira, C., Camacho, J.P.M., Porto-Foresti, F., and Foresti, F. 2014. Delimiting the Origin of a B Chromosome by FISH Mapping, Chromosome Painting and DNA Sequence Analysis in *Astyanax paranae* (Teleostei, Characiformes). *PLoS One* **9**(4): e94896. doi:10.1371/journal.pone.0094896.
- Sirotkin, A.M., Edelman, W., Cheng, G., Klein-Szanto, A., Kucherlapati, R., and Skoultschi, A.I. 1995. Mice develop normally without the H10 linker histone. *Proc. Natl. Acad. Sci. U. S. A.* **92**(14): 6434–6438.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **30**(12): 2725–2729.
- Tang, M.C.W., Binos, S., Ong, E.K., Wong, L.H., and Mann, J.R. 2014. High histone variant H3.3 content in mouse prospermatogonia suggests a role in epigenetic reformatting. *Chromosoma* **123**(6): 587–595. doi:10.1007/s00412-014-0475-8.

- Thatcher, T.H., and Gorovsky, M.A. 1994. Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Res.* **22**(2): 174–179. doi:10.1093/nar/22.2.174.
- Vega, L., Díaz, E., Cross, I., and Rebordinos, L. 2002. Caracterizaciones citogenética e isoenzimática del lenguado *Solea senegalensis* Kaup, 1858. *Bol. Inst. Esp. Oceanogr.* **18**(1-4): 245-250.
- Viñas, J., Asensio, E., Cañavate, J.P., and Piferrer, F. 2013. Gonadal sex differentiation in the Senegalese sole (*Solea senegalensis*) and first data on the experimental manipulation of its sex ratios. *Aquaculture* **384-387**: 74–81. doi:10.1016/j.aquaculture.2012.12.012.
- Vitelli, L., and Weinberg, E.S. 1983. An inverted sea urchin histone gene sequence with breakpoints between TATA boxes and mRNA cap sites. *Nucleic Acids Res.* **11**(7): 2135–2154. doi:10.1093/nar/11.7.2135.
- Wang, Z.-F., Sirotkin, A.M., Buchold, G.M., Skoultchi, A.I., and Marzluff, W.F. 1997. The mouse histone H1 genes: gene organization and differential regulation. *J. Mol. Biol.* **271**(1): 124–138. doi:10.1006/jmbi.1997.1166.
- Wei, Q., Jung, H.J., Hwang, D.S., Hwang, B.H., Gim, Y., and Cha, H.J. 2007. Escherichia coli-based expression of functional novel DNA-binding histone H1 from *Carassius auratus*. *Enzyme Microb. Technol.* **40**(6): 1484–1490. doi:10.1016/j.enzmictec.2006.10.028.
- Yu, L., and Gorovsky, M.A. 1997. Constitutive expression, not a particular primary sequence, is the important feature of the H3 replacement variant hv2 in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **17**(11): 6303–6310. doi:10.1128/MCB.17.11.6303.

Primer pair	Sequence	Tm
SseH3.1	5'-GAGCCGGAGGACTGCAAGCAG-3'	65°C
SseH3.2	5'-GGTTGGTGTCTCGAACAGGCC-3'	
SseH4.1	5'-CGTGGCGGAGTGAAGCGTTCTC-3'	65°C
SseH4.2	5'-GCGTAGACCACGTCCATGGCG-3'	
TLR8SseFw	5'-AGCGAGGCCTTCCTGTTATAC-3'	57°C
TLR8SseRv	5'-GGGAGATGTTTCTCTCCTTCCT-3'	

Table 1. Primer pairs used to screen the *Solea senegalensis* BAC library.

Draft

BAC clone	Contig	Histone genes
BAC5K5	634	H2A, H2B (x2), H3.2, H4 (x2)
	641	H1, H2A, H2B, H3.2, H4
BAC12D22	824	H1 (x2), H2A (x3), H2B (x5), H3.2 (x3), H4 (x2)
BAC30J4	001	H1.0, H3.3

Table 2. Histone genes found in the three different BAC clones analyzed.

Draft

Table 3. Variability data between BAC clones containing the main histone cluster.

Histone	BAC	Intra-BAC variability		Inter-BAC variability		Fixed differences	
		Nucleotide	Amino Acid	Nucleotide	Amino Acid	Nucleotide	Amino Acid
H1	5K5	--	--	0.05571 ± 0.02794	0.062	27	14
	12D22	0.02053 ± 0.01012	0.019				
H2A	5K5	0	0	0.04165 ± 0.02265	0.008	9	1
	12D22	0.03083 ± 0.00970	0				
H2B	5K5	0	0	0.02164 ± 0.01369	0.016	8	2
	12D22	0	0				
H3	5K5	0.00488 ± 0.00243	0	0.01309 ± 0.00672	0	4	0
	12D22	0.00162 ± 0.00076	0				
H4	5K5	0.00214 ± 0.00101	0	0.03387 ± 0.01999	0	10	0
	12D22	0	0				

FIGURE CAPTIONS

Figure 1. Alignment of histone protein sequences of *S. senegalensis* and other species. A. Variant of the H1 in Replication Dependent (RD) histones and Replication Independent (RI) histones in *Solea senegalensis*. RD variants showed three isoforms. B. Variants of the H3.2. of *S. senegalensis* differ in one amino acid from that of other species. C. Different position of this variation from that which differentiates between H3.1 and H3.2 in other vertebrates (Fig. 1C). Sse: *Solea senegalensis*; Gac: *Gasterosteus aculeatus*; Ola: *Oryzias latipes*; Dre: *Danio rerio*.

Figure 2. Schematic representation of histone genes in *S. senegalensis*. A. Structure of the arrangement in the cluster lacking the *H1* histone. B. Cluster arrangement including the *H1*. C. Composed cluster with and without *H1* including one inversion.

Figure 3. Maximum-likelihood phylogenetic tree constructed from a Bio Neighbor-Joining as starting tree for the H1 histone genes. RI variants are clustered in a clade different from that of the RD variants' cluster.

Figure 4. Maximum-likelihood phylogenetic tree constructed from a Bio Neighbor-Joining as starting tree for the H3 histone genes. H3.3-like variants are clustered in a clade different from that of the H3.2-like variants' cluster.

Figure 5. Syntenic comparison by means of a Circos analysis between the 59 genes found from the three BAC clones and A. *Gasterosteus aculeatus*, distributed in 5 chromosomes; B. *Oryzias latipes* in 6 chromosomes; and C. *Danio rerio* in 8 chromosomes. On the left side, the distribution of the BAC clones of *S. senegalensis* can be observed. Indicated within each BAC are the genes found by annotation, and the corresponding localizations

in the genomes by crossing lines. The chromosomes of that genome are represented on the right side of the figure.

Figure 6. FISH of the BACs isolated in the library that contain the genes indicated: A. Double FISH showing co-location of BAC 5K5 (*H1*, *H2A*, *H2B*, *H3.2*, *H4*) (Green) and BAC12D22 (*H1*, *H2A*, *H2B*, *H3*, *H4*) (pink). B. FISH showing the location of BAC30J4 (*H1.0*, *H3.3*) in two pairs of chromosomes. C. mFISH showing co-location in the largest metacentric chromosomes of BAC 5K5 (cyan), BAC12D22 (pink) and BAC11O20 (*dmrt1*) (red). A diagrammatic representation of the location in the chromosomes is included.

Draft

Figure 1
Alignment of histone protein s
488x292mm (72 x 72 DPI)

Figure 2
Schematic representation of hi

Figure 3

Figure 4

Figure 5

Draft

Figure 6
FISH of the BACs isolated in t

Table S1. Genes found in the BACs under study.

BAC Name	Abbreviation genes name	Genes name
BAC5K5	<i>tpm4</i>	<i>Tropomyosin alpha-4 chain-like</i>
	<i>klf2</i>	<i>Krueppel-like factor 2</i>
	<i>eps15l1</i>	<i>epidermal growth factor receptor substrate like 1like</i>
	<i>rab8a</i>	<i>Ras-related protein Rab-8A</i>
	<i>cib3</i>	<i>Calcium and integrin-binding family member 3-like</i>
	<i>slc1a3</i>	<i>Excitatory amino acid transporter 1</i>
	<i>H2a</i>	<i>Histone H2A</i>
	<i>H3.2</i>	<i>Histone H3</i>
	<i>H4</i>	<i>Histone H4</i>
	<i>H2b</i>	<i>Histone H2B</i>
	<i>H1</i>	<i>Histone H1</i>
	<i>calr</i>	<i>Calreticulin</i>
	<i>rx2</i>	<i>Retinal homeobox protein Rx-2</i>
	<i>tpm1</i>	<i>Tropomyosin alpha-1 chain like isoform 2</i>
	<i>ap1m1</i>	<i>AP-1 complex subunit mu-1</i>
BAC12D22	<i>H2b</i>	<i>Histone H2B</i>
	<i>H1</i>	<i>Histone H1</i>
	<i>H2a</i>	<i>Histone H2A</i>
	<i>H3</i>	<i>Histone H3</i>
	<i>H4</i>	<i>Histone H4</i>
	<i>ankrd45</i>	<i>Ankyrin repeat domain-containing protein 45</i>
	<i>tmem 70</i>	<i>transmembrane protein 70 mitochondrial like</i>
BAC30J4	<i>rhot2</i>	<i>Mitochondrial Rho GTPase 2</i>
	<i>H1f0</i>	<i>Histone H1.0</i>
	<i>rhbd11</i>	<i>Rhomboid-related protein 1-like</i>
	<i>wdr24</i>	<i>WD repeat-containing protein 24</i>
	<i>anks3</i>	<i>Ankyrin repeat and SAM domain-containing protein 3</i>
	<i>upfpc8orf33</i>	<i>UPF0488 protein C8orf33 homolog</i>
	<i>frmpd4</i>	<i>FERM and PDZ domain-containing protein 4-like</i>
	<i>nlrc3</i>	<i>NLR family, CARD domain containing 3</i>
	<i>gcgr</i>	<i>Glucagon receptor-like</i>
	<i>pcdh8</i>	<i>Protocadherin-8-like</i>
	<i>wdr90</i>	<i>WD repeat-containing protein 90</i>
	<i>tyb12</i>	<i>Thymosin beta-12</i>
	<i>tlr8</i>	<i>Toll-like receptor 8</i>
	<i>tlr7</i>	<i>Toll-like receptor 7</i>
	<i>mid1</i>	<i>Midline-1-like</i>
	<i>cog3</i>	<i>Conserved oligomeric Golgi complex subunit 3</i>
	<i>arhgap6</i>	<i>Rho GTPase-activating protein 6</i>
	<i>ednrb</i>	<i>Endothelin B receptor</i>
	<i>egfl6</i>	<i>Epidermal growth factor-like protein 6</i>
	<i>gper1</i>	<i>G-protein coupled estrogen receptor 1</i>
	<i>H3.3</i>	<i>Histone 3.3</i>

Figure S1. Intra-specific alignment of canonical histones of *Solea senegalensis*. The red line separates sequences from the two BAC under study.

H1

H2A

H2B

H3

H4

raft