Gene 503 (2012) 254–259

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Short communication

A *swimy* locus on Y chromosome of the platyfish (*Xiphophorus maculatus*) is derived from a novel DNA transposon *Zisupton*

Qingchun Zhou *

Physiologische Chemie I, Biozentrum, University of Würzburg, Würzburg 97074, Germany
Hubei Key Laboratory of Genetic Regulation and Integrative Biology, College of Life Science, Huazhong Normal University, Wuhan 430079, China

ARTICLE INFO

Article history:
Accepted 18 April 2012
Available online 2 May 2012

Keywords:
Xiphophorus maculatus
Swimy
Sex chromosome marker
Zisupton

ABSTRACT

A *swimy* locus derived from a novel DNA transposon *Zisupton* was located on the sex determination region (SD) of *Xiphophorus maculatus*. The analysis of expression pattern showed that *swimy* was exclusively expressed in adult testis in *X. maculatus*. The putative 939 aa sequence contains four Zn-finger domains, such as two C2H2 type, one NFX type and one SWIM type Zn-finger domain, and one SAP DNA-binding domain. *Swimy* has about 7 copies per haploid *X. maculatus* genome with Y-specific copies located in the SD region, and become the second new W-linked marker of platyfish. Analysis of the structure and distribution of this sex-linked marker is benefit to shed new light on the evolutionary dynamics of sex chromosomes in fish.

1. Introduction

Many repetitive sequences are located in the sex determination (SD) region of sex chromosome in animal. Most of them are transposable elements. Transposable elements have been divided into two major classes based on their mode of transposition: retrotransposons and DNA transposons. Retrotransposons are further divided into 2 major types: long terminal repeat (LTR) and non-LTR (Craig et al., 2002). DNA transposons usually move in the genome by the self-replication mechanism with the encoded transposase, which acts as an endonuclease and triggers the transferring.

There is another group of DNA transposon *Helitrons* which were identified recently, and replicate through the so-called "rolling-circle" mechanism (Kapitonov and Jurka, 2001). They have no TIRs, but rather short conserved terminal motifs (Kapitonov and Jurka, 2001). The third group of DNA transposons named *Mavericks* or *Polinton* that is recently identified in eukaryotic genomes encoded multiple proteins and possessed long terminal inverted repeats (TIRs) with a 6-bp target site duplication (TSD) (Kapitonov and Jurka, 2006; Pritham et al., 2007). They may have originated from ancient relatives of Mavirus (Fischer and Suttle, 2011). Autonomous *Polintons* are typically 15–20 kb in length and encode as many as 5–10 different proteins, including retroviral-like integrase (INT), adenovirus protease (PRO), DNA polymerase B (POLB) and putative ATPase (ATP) (Feschotte and Pritham, 2005; Kapitonov and Jurka, 2006; Pritham et al., 2007). A new family of DNA transposon called *Zisupton* was recently found in fish and various divergent organisms (Böhne et al., 2012). This element is 11 kb in length and carries a single coding sequence of 24 exons. Typical copies are delimited by 99-bp conserved TIRs and flanked by 8-bp TSDs.

In several flies, the nematode *Caenorhabditis elegans* and most mammals, the *SD* genes and mechanisms are known and reasonably well understood (Capel, 1998, cited references). It appears to be a large gap between the mechanisms of SD in invertebrates (nematodes and flies) and mammals (Hornung et al., 2004). Fishes are one of the attractive vertebrates for the study of the evolution of SD because they present various types of sexuality from hermaphroditism to gonochorism and from environmental to genetic SD (Baroiller et al., 1999; Devlin and Nagahama, 2002). Almost all different forms of genetic SD are found in fish, including XX/XY system, ZW/ZZ system, multiple sex chromosomes (e.g. X, Y and W chromosomes in the platyfish), polygenic *SD* factor (e.g. the *SD* factors are distributed over several chromosomes) and autosomal influence (e.g. XX males in the medaka, Nanda et al., 2003). Even in model fish, such as zebrafish (*Danio rerio*) and smooth pufferfish (*Tetraodon nigroviridis*), there is no sex-linked markers or no sex chromosomes identified so far, although the *SD* locus is located in the linkage group 19 of *Takifugu rubripes* (Kikuchi et al., 2007). Hence, alternative fish models like medaka, platyfish, stickleback, tilapia and salmonids are used to study SD and sex chromosome evolution. In the medaka fish *Oryzias latipes*, the *dmrt1bY* (or *DMY*) was identified as candidate of the master male *SD* gene (Matsuda et al., 2002; Nanda et al., 2002).

Abbreviations: SD, sex determination; LTR, long terminal repeat; LTRs, long terminal repeats; TSD, target site duplication; INT, retroviral-like integrase; PRO, adenovirus protease; POLB, DNA polymerase B; ATP, ATPase; BAC, bacterial artificial chromosome; RFLP, restriction fragment length polymorphism; DIG, digoxigenin; AR, androgen receptor; SRCs, steroid receptor coactivators.

* Hubei Key Laboratory of Genetic Regulation and Integrative Biology, College of Life Science, Huazhong Normal University, Wuhan 430079, China. Tel.: +86 027 6786 7714.
E-mail address: qczhou@mail.ccnu.edu.cn.

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2012.04.062

However, *dmrt1bY* was not found in two closely related species of the genus *Oryzias* (*O. celebensis* and *O. mekongensis*) (Kondo et al., 2003), let alone in other more divergent species, such as tilapia, guppy, platyfish, and zebrafish (Kondo et al., 2003; Veith et al., 2003).

Xiphophorus are Poeciliid fish that inhabit freshwater drainages in eastern Mexico, Guatemala, Belize and Honduras. There are 23 species in this genus, which largely fall into three taxonomic groups: the northern and southern swordtails and the platyfishes (Meyer and Schartl, 2002). They are one of the well studied groups of fish for the molecular genetics of SD (Kallman and Schreibman, 1973; Schartl, 2004; Volff and Schartl, 2001). The various tapes of genetic SD observed in fish are well demonstrated in *Xiphophorus* fishes. The platyfish, *Xiphophorus maculatus*, has three types of sex chromosomes: X, Y and W (Rio Jamapa strain has XX/XY, and Usumacinta strain has WY/YY sex chromosome system). Females have WY, WX or XX chromosomes. Males are either YY or XY chromosomes. Therefore the platyfish is a good model to investigate the relationship between XX/XY and ZW/ZZ system. W, X and Y chromosomes are homologous in the platyfish (Kallman and Schreibman, 1973), however, the W chromosome is apparently divergent. None of the X- and Y-chromosomal loci studied so far has alleles on the W chromosome, except for *egfrb* and *Polinton* (Gutbrod and Schartl, 1999; Kallman and Schreibman, 1973). Importantly, the *SD* locus in the *X. maculatus* population (Rio Jamapa) is located between the *Xmrk* and *egfrb* receptor tyrosine kinase genes, which delimit a relatively small region (below 1 Mb) (Gomez et al., 2004; Gutbrod and Schartl, 1999; Volff et al., 2003). Thus it becomes practicable to search the master *SD* gene of *X. maculatus* through a positional cloning method. Over 1400 kb-sized bacterial artificial chromosome (BAC) contigs covering the SD region of the X and Y chromosomes of *X. maculatus* have been assembled and partially sequenced (Froschauer et al., 2002; Schultheis et al., 2006). Several gene candidates, transposons and repeats were located at the SD region on the X or Y chromosome, but no one was found on the W chromosome except *Polinton* (Böhne et al., 2008; Zhou et al., 2006, 2010).

In this study, we reported a *swimy* locus recently identified as partial sequence of a novel fish DNA transposon *Zisupton*, exclusively expressed in adult testis of platyfish. It has one Y-specific copy located in the SD region, and becomes the second new W-linked marker of platyfish. Through further analysis of the structure and distribution of this sex-linked marker, we will provide new insights into the relationship between XX/XY and ZW/ZZ systems.

2. Materials and methods

2.1. Fishes and genomic library

Fishes were maintained in the aquarium facilities of the Biozentrum at the University of Würzburg, such as *X. maculatus* (Rio Jamapa strain WLC1274, Rio Usumacinta strain WLC1372), *Xiphophorus hellerii* (Rio Lantecilla strain hIII and Rio Santecomapan), *Xiphophorus milleri* (Catemaco 96), and the origin of strains has been previously published (Schartl et al., 1999; Volff et al., 2000). Bacterial artificial chromosome clones were isolated from a genomic library of *X. maculatus* XY males (Froschauer et al., 2002).

2.2. DNA extraction and analysis

Genomic DNA extraction and Southern blot analysis were performed according to standard protocols. Southern blots were carried out with random primed ^{32}P-labeled probes as previously described (Zhou et al., 2006). The restriction fragment length polymorphism (RFLP) on genomic DNA was analyzed from Southern blot results. Genomic library screening and contig construction were performed as previously described (Froschauer et al., 2002). DNA sequencing reactions were done using the CEQ DTCS dye terminator cycle sequencing kit and run on a CEQ 2000XL DNA sequencing system (Beckman Coulter, USA). Primers swimyF1 (5′-CTGAGGAATGCA-TTGCAGGT-3′) and swimyR1 (5′-AAGACCTACCTGGCATGCTG-3′) were used to amplify the segment encoding the SWIM, NFX and C2H2 type Zn_finger domains at the N-terminal region (probe I in Fig. 1).

2.3. The whole-mount in situ hybridization

RNA whole-mount in situ hybridization is used to detect specific mRNA sequences in tissue or cell by hybridizing the complementary strand of a nucleotide probe. The procedure in this work is as follows: 1) the antisense RNA probe was prepared by digoxigenin (DIG RNA Labeling Kit, Roche), and stored at −20 °C. 2) Adult testes were fixed, dehydrated, rehydrated, treated with proteinase K, prehybridized and hybridized overnight at 65 °C in routine. 3) After the probe was removed, the testes were subsequently washed twice in 50% formamide in 2× SSCT, once in 2× SSCT and twice in 0.2% SSCT, treated with blocking solution (5% sheep serum in PBST) for at least 1 h, and incubated with sheep anti-digoxigenin (DIG) Fb antibody fragments conjugated with alkaline phosphatase overnight at 4 °C with gentle shaking. 4) The testes were washed 6 times for 20 min each in PBST, washed 5 min in SB staining buffer and then treated with the NBT/BCIP staining solution. The staining reaction was stopped by washing the testes 3 times in PBST. 4) The stained testes were fixed, dehydrated, embedded and cut into sections, which could be viewed under the microscope.

2.4. Sequence analysis

Multiple sequence alignments were set up using ClustalX (Thompson et al., 1997). Phylogenetic analyses were performed on amino-acid alignments using the neighbor-joining method (Saitou and Nei, 1987; 1000 pseudosamples) as implemented in MEGA4 (Tamura et al., 2007). BLAST analysis was essentially performed using sequence databases accessible from NCBI (http://www.ncbi.nlm.nih.gov/BLAST/, including position-specific iterated BLAST analysis), Ensembl (http://www.ensembl.org/) and the medaka genome server (http://dolphin.lab.nig.ac.jp/medaka/).

3. Results

3.1. Swimy locus encodes a putative DNA binding protein

During sequence analysis of BACs in the SD region of platyfish, a predicted exon (645 nt in length) was found to encode a putative translation product with a SWIM Zn_finger domain. This domain has a $CXCX_nCXH$ motif (n = 6–16 residues) with predicted zinc-chelating residues. The SWIM Zn_finger domain is found in a variety of prokaryotic and eukaryotic proteins, including: 1) bacterial ATPases of the SWI2/SNF2 family; 2) plant MuDR transposases and transposase-derived Far1 nuclear proteins; 3) vertebrate mitogen-activated protein kinase kinase kinase-1 (MAPKKK-1) and 4) several hypothetical proteins. The SWIM domain is named after *SWI*2/SNF2 and *M*uDR, and is predicted to have DNA-binding and protein–protein interaction functions (Makarova et al., 2002).

An approximately11 kb large genomic consensus sequence containing the above predicted exon was assembled by “shotgun” sequencing. This consensus sequence was designated under the name *swimy*, for encoding the SWIM Zn_finger domain.

After comparing the cDNA sequence with *swimy*, the result indicated that *swimy* has 7 exons at least. This 11 kb genomic sequence has no stop codon or frame shift in its coding region and has 99.1% nucleotide identity with the cDNA sequence.

The putative 939 aa sequence encoded by exon 1 was predicted to contain four Zn-finger domains such as two C2H2 type (accession