Marine Genomics 22 (2015) 45–54

Contents lists available at ScienceDirect

Marine Genomics

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

Evolutionary history and adaptive significance of the polymorphic Pan I in migratory and stationary populations of Atlantic cod (*Gadus morhua*)

Øivind Andersen [a,b,*], Hanne Johnsen [c], Maria Cristina De Rosa [d], Kim Præbel [e], Suzana Stjelja [f], Tina Graceline Kirubakaran [b], Davide Pirolli [g], Sissel Jentoft [h], Svein-Erik Fevolden [e]

[a] Nofima, P.O. Box 5010, N-1430 Ås, Norway
[b] Department of Animal and Aquaculture Sciences, University of Life Sciences, P.O. Box 5003, N-1430 Ås, Norway
[c] Nofima, Muninbakken 9-13, N-9291 Tromsø, Norway
[d] Institute of Chemistry of Molecular Recognition—CNR and Institute of Biochemistry and Clinical Biochemistry, Catholic University of Rome, 00168 Rome, Italy
[e] Department of Arctic and Marine Biology, UiT The Arctic University of Norway, N-9037 Tromsø, Norway
[f] Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7023, S-75007 Uppsala, Sweden
[g] Institute of Biochemistry and Clinical Biochemistry, Catholic University of Rome, 00168 Rome, Italy
[h] Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, P.O. Box 1066 Blindern, N-0316 Oslo, Norway

ARTICLE INFO

Article history:
Received 2 February 2015
Received in revised form 19 March 2015
Accepted 19 March 2015
Available online 31 March 2015

Keywords:
Pan I
Rhodopsin
Synaptic vesicle
Atlantic cod
Polymorphism
Ecotypes

ABSTRACT

The synaptophysin (SYP) family comprises integral membrane proteins involved in vesicle-trafficking events, but the physiological function of several members has been enigmatic for decades. The presynaptic SYP protein controls neurotransmitter release, while SYP-like 2 (SYPL2) contributes to maintain normal Ca^{2+}-signaling in the skeletal muscles. The polymorphic pantophysin (*Pan* I) of Atlantic cod shows strong genetic divergence between stationary and migratory populations, which seem to be adapted to local environmental conditions. We have investigated the functional involvement of Pan I in the different ecotypes by analyzing the 1) phylogeny, 2) spatio-temporal gene expression, 3) structure–function relationship of the Pan I^A and I^B protein variants, and 4) linkage to rhodopsin (*rho*) recently proposed to be associated with different light sensitivities in Icelandic populations of Atlantic cod. We searched for *SYP* family genes in phylogenetic key species and identified a single *syp*-related gene in three invertebrate chordates, while four members, *Syp*, *Sypl1*, *Sypl2* and *synaptoporin* (*Synpr*), were found in tetrapods, Comoran coelacanth and spotted gar. Teleost fish were shown to possess duplicated *syp*, *sypl2* and *synpr* genes of which the *sypl2b* paralog is identical to *Pan* I. The ubiquitously expressed cod *Pan* I codes for a tetra-spanning membrane protein possessing five amino acid substitutions in the first intravesicular loop, but only minor structural differences were shown between the allelic variants. Despite sizable genomic distance (>2.5 Mb) between *Pan* I and *rho*, highly significant linkage disequilibrium was found by genotyping shallow and deep water juvenile settlers predominated by the *Pan* I^A–*rho*^{A} and *Pan* I^B–*rho*^{B} haplotypes, respectively. However, the predicted rhodopsin protein showed no amino acid changes, while multiple polymorphic sites in the upstream region might affect the gene expression and pigment levels in stationary and migratory cod. Alternatively, other strongly linked genes might be responsible for the sharp settling stratification of juveniles and the different vertical behavior patterns of adult Atlantic cod.

1. Introduction

Synaptic vesicles contain two classes of obligatory components: transport proteins involved in neurotransmitter uptake, and trafficking proteins that participate in synaptic vesicle exocytosis, endocytosis, and recycling. The latter proteins include the synaptophysin (SYP) family of MARVEL-domain containing membrane proteins featuring a four transmembrane-helix structure (Arthur and Stowell, 2007). SYP is the most abundant protein in presynaptic neurotransmitter vesicles and synaptic-like microvesicles of neuroendocrine cells, but its physiological function has been enigmatic for decades (Jahn et al., 1985; Wiedenmann and Franke, 1985; Navone et al., 1986). Several studies have provided evidence that SYP is required for the exocytotic release of neurotransmitters (Alder et al., 1992a,b,1995; Shibaguchi et al., 2000), but *SYP* knock-out mice form typical synaptic vesicles with normal synaptic transmission that is probably due to redundancy with related proteins (Eshkind and Leube, 1995; McMahon et al., 1996). Accordingly, retinal rod photoreceptor cells, which do not synthesize the SYP-related synaptoporin (SYNPR), showed a considerable

* Corresponding author at: Nofima, P.O. Box 5010, N-1430 Ås, Norway. Tel.: +47 930 60 248.
E-mail addresses: oivind.andersen@nofima.no (Ø. Andersen), hanne.johnsen@nofima.no (H. Johnsen), mariacristina.derosa@icrm.cnr.it (M.C. De Rosa), kim.praebel@uit.no (K. Præbel), suzana.stjelja@gmail.com (S. Stjelja), tina.graceline@nmbu.no (T.G. Kirubakaran), davide.pirolli@rm.unicatt.it (D. Pirolli), sissel.jentoft@ibv.uio.no (S. Jentoft), svein-erik.fevolden@uit.no (S.-E. Fevolden).

http://dx.doi.org/10.1016/j.margen.2015.03.009
1874-7787/

reduction in the number of synaptic vesicles in SYP-deficient mice suggesting an important role of SYP in synaptic vesicle recycling and formation (Spiwoks-Becker et al., 2001). SYP was consistently shown to be required for kinetically efficient endocytosis to ensure vesicle availability during and after sustained activity, but the lack of SYP did not affect exocytosis of synaptic vesicles (Kwon and Chapman, 2011).

The involvement of SYP in modulating behavior was demonstrated in mice lacking SYP, which displayed increased exploratory behavior, but reduced learning ability and memory (Schmitt et al., 2009). Intriguingly, *SYP* knock-out mice were recently shown to be defective in delaying the clock phase implying that SYP plays a role in resetting of the circadian clock (Aramendy et al., 2013). It should be noted that human SYP is mapped to the Xp11.23–p11.22 interval implicated in the eye disorders retinitis pigmentosa 2, congenital stationary night blindness and Åland Island disease (Fisher et al., 1997). While SYP and SYNPR are restricted to synaptic vesicles (Knaus et al., 1990; Marqukze-Pouey et al., 1991), SYP-like 1 (SYPL1) is a broadly distributed membrane component of small cytoplasmic transport vesicles and associated with GLUT4-containing vesicles in adipocytes (Leube, 1994; Haass et al., 1996; Brooks et al., 2000). The physin members of the SYP family also include SYPL2 designated mitsugumin29, which seems to be essential to maintain triad junction structural integrity and normal Ca^{2+}-handling in the skeletal muscles, but is also involved in Ca^{2+}-dependent exocytosis in the brain (Takeshima et al., 1998; Zhao et al., 2011; Satoh et al., 2012). The functional role of the physins is largely unknown in non-mammalian species, and the phylogenetic relationship of the duplicated teleost physins has not been resolved (Lagman et al., 2013). Brain transcriptome variation among behavioral distinct strains of zebrafish was shown to include differentially expressed *sypb* and *synpr* (Drew et al., 2012).

Two decades ago, genetic analysis of trans-Atlantic populations of Atlantic cod revealed highly variable allele frequencies of the GM798 locus coding for a SYP-like protein designated synaptophysin (Syp I), later renamed pantophysin or Pan I (Pogson et al., 1995; Fevolden and Pogson, 1997). The highly polymorphic marker has been widely used to characterize the genetic diversity of Atlantic cod populations and the association with growth properties (Imsland and Jónsdóttir, 2003; Nordeide et al., 2011), but the physiological function of the protein and the allelic variants still remains unknown. Atlantic cod is divided into multiple populations widely distributed in Arctic and temperate waters across the North Atlantic Ocean, including the Barents Sea and the brackish Baltic Sea, from the shoreline down to the continental shelf at depths down to 600 m. The *Pan* I polymorphism has been associated with different spawning and feeding behaviors in Canadian and Icelandic cod populations (Jónsdóttir et al., 1999; Pampoulie et al., 2008; Tamdrari et al., 2012; Thorsteinsson et al., 2012), but the genetic divergence is much more pronounced in offshore migratory and coastal stationary populations in the Northeast Atlantic Ocean. The large population of migratory Northeast Arctic cod (NEAC) is predominated by the *Pan* I^{BB} genotype, while *Pan* I^{AA} is mainly found in the relative stationary Norwegian coastal cod (NCC) living in shallow waters and fjords (Fevolden and Pogson, 1997; Pogson and Fevolden, 2003; Sarvas and Fevolden, 2005). The NEAC and NCC populations occur sympatrically during the breeding season in the coastal region of Lofoton and Vesterålen, and eggs and larvae drift northwards with the coastal current. Profound difference in *Pan* I allele frequencies was recently reported between Atlantic cod 0-group juveniles settling at shallow versus deep water with *Pan* I frequencies similar to adult NCC and NEAC fish, respectively, whereas more intermediate frequencies were found in yet unsettled juveniles caught in the pelagic habitat (Fevolden et al., 2012).

The *Pan* I polymorphism has been proposed to be caused by divergent selection pressure acting on the two alleles at different environmental conditions, such as water temperature, salinity and depth (Fevolden and Pogson, 1997; Pogson, 2001; Pogson and Mesa, 2004; Case et al., 2005; Sarvas and Fevolden, 2005; Pogson and Fevolden, 2003). However, strong linkage disequilibrium in a 5.7-kb region encompassing the *Pan* I locus could indicate that selection is instead targeting a linked gene (Fevolden and Pogson, 1997; Pogson and Fevolden, 2003). Localized genomic differentiation was recently identified within three linkage groups (LG1, 2, 7) of the Atlantic cod genome, and the divergence of stationary and migratory ecotypes seemed to be uniquely associated with a large genomic region comprising multiple genes, including *Pan* I and the flanking *sortilin* (*sort*)*1* and *ataxin* (*atxn*)*7* genes on LG1 (Hemmer-Hansen et al., 2013; Karlsen et al., 2013; Therkildsen et al., 2013). Interestingly, the *rhodopsin* (*rho*) gene coding for the highly light sensitive rod photoreceptor was identified as one of the eight outlier loci likely to be subject to directional selection in local demes, or closely linked to loci under selection (Nielsen et al., 2009). Consistently, *rho* polymorphisms in Icelandic coastal and frontal populations were recently proposed to be connected with different light sensitivities, although no non-synonymous sites were identified (Pampoulie et al., 2015). In diverse fish species, amino acid substitutions at specific spectral tuning sites of rhodopsin have been found to provide local adaptation to spectral regimes (Sugawara et al., 2005; Nickle and Robinson, 2007; Wang et al., 2009; Sivasundar and Palumbi, 2010; Larmuseau et al., 2010; Hofmann et al., 2012).

Although the intracellular loops of Pan I display signatures of positive selection (Pogson, 2001; Pogson and Fevolden, 2003; Canino and Bentzen, 2004; Pogson and Mesa, 2004), the functional relevance of these variable regions needs to be assessed. As a first step towards understanding the physiological role played by Pan I in the different ecotypes of Atlantic cod, we investigated its evolutionary origin, spatio-temporal gene expression patterns and structure–function relationship of the two main variants. The proposed involvement of rhodopsin in the divergence of cod populations was further examined by comparing potential regulatory *rho* sequences in NEAC and NCC fish and by studying patterns of linkage disequilibrium between *Pan* I and *rho* in juvenile fish displaying different settling regimes in the Norwegian Varangerfjord.

2. Materials and methods

2.1. Identification of physin genes

The four SYP family members SYP, SYNPR, SYPL1 and SYPL2 were retrieved from the following databases: http://www.ncbi.nlm.nih.gov, http://www.ensembl.org (release 73–76), http://www.uniprot.org/taxonomy/7739 (Amphioxus, *Branchiostoma lanceolatum*), http://www.uniprot.org/taxonomy/10224 (acorn worm, *Saccoglossus kowalevskii*), http://esharkgenome.imcb.a-star.edu.sg/ (elephant shark, *Callorhinchus milii*), http://codgenome.no (Atlantic cod), and http://www.icisb.org/sequence (Atlantic salmon, *Salmo salar*). Unannotated genes were identified using known orthologs as query sequences to BLAST searching the databases. Downloaded sequences were controlled by examination of sequence homology and conserved motifs, including MARVEL domain. Predicted proteins with extended N- or C-terminal ends or lacking conserved motifs (e.g. ground tit SYP) were manually examined for exon–intron organization and undetermined regions. Missing genes were verified by blasting the genomic sequences against available orthologs and paralogs. Conserved flanking genes were identified using Genomicus (v. 73.01–75.01) to further search for missing genes.

2.2. Phylogenetic analyses

A multiple sequence alignment of 82 SYP, SYPL1, SYPL2, and SYNPR peptide sequences from 20 species (Supplementary Table S1) was obtained using MUSCLE (www.ebi.ac.uk/Tools/msa/muscle/). The alignment was manually edited in BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html) to remove poorly aligned positions and divergent regions. In particular, gaps over four residues found in at least half of the sequences were removed from the alignment. The resulting trimmed aligned sequences were then used for Bayesian inference of phylogeny