Asa Langefors · Jakob Lohm · Torbjörn von Schantz

Allelic polymorphism in MHC class II B in four populations of Atlantic salmon (Salmo salar)

Received: 18 October 2000 / Accepted: 2 April 2001 / Published online: 20 June 2001
© Springer-Verlag 2001

Abstract We sequenced exon 2 of the MHC class II B gene in Atlantic salmon from the Baltic Sea and identified 17 different exon 2 alleles among 22 different restriction fragment length polymorphism haplotypes. The sequences differed at between 1 and 34 bases. Two different tests were used to estimate the importance of recombination in the generation of new alleles. Recombination events appear to have occurred between three and nine times. Only two pairs of sequences differed by less than five nucleotides, minimizing the importance of point mutations for generating new alleles. Phylogenetic analysis showed that sequences did not cluster according to populations, and genetic distances between populations were small compared to those obtained by allele frequency data. These results, together with the similarity found between exon 2 sequences from Baltic salmon and Norwegian salmon, indicate that all of the identified alleles were present in the ancient salmon population colonizing the Baltic rivers after the last glaciation.

Keywords MHC class II B1 · Atlantic salmon · Recombination · Population · Genetic variation

Introduction

The major histocompatibility complex (MHC) is known to be a highly polymorphic region in the vertebrate genome (Klein 1990). Several different mechanisms have been suggested to explain the generation and maintenance of this extreme polymorphism. New genetic variation is normally generated through single point mutations. However, in MHC genes, microrecombination, the shuffling of short segments between alleles or genes, appears to generate new alleles (Hedrick 1999; Pease et al. 1983), probably at a higher rate than point mutations. The extent of recombination, and its importance for MHC polymorphism are widely debated. Evidence of microrecombination comes from (1) sperm counts in mice and humans (Högstrand and Böhme 1994; Zangenberg et al. 1995), (2) observations in, e.g., South American Indians, of new alleles obviously resulting from gene conversion between two old alleles (Belich et al. 1992; Parham and Ohta 1996), and (3) sequence alignments suggesting that small segments have been shuffled between alleles (Wittzell et al. 1999). Most evidence for recombination in the MHC comes from studies on humans and mice (reviewed in Marthinson et al. 1999).

Mechanisms suggested to maintain MHC variation are different selection pressures operating at different levels, such as parasite-driven selection versus sexual selection (Penn and Potts 1998; Potts et al. 1991), and overdominance selection versus frequency-dependent selection (Hughes and Nei 1989; Takahata and Nei 1990). The phenomenon of shared motifs between distantly related species has recently been attributed to convergent evolution at the molecular level rather than shared ancestry (Kriener et al. 2000; Yeager and Hughes 1999). In any of these cases, intragenic variation originates from either point mutations and/or recombination, while intergenic variation originates from either of these or a “birth and death process” (gene duplication and extinction; Gu and Nei 1999).

Structurally, the first domain of the class II B molecule, encoded by exon 2, is divided into a β sheet and an α helix (Brown et al. 1993). In mammals, these two subregions appear to be of different evolutionary origin, with contrasting polymorphic patterns. The α-helix subregion is more conserved, has a more even distribution of synonymous and nonsynonymous substitutions, and shows a different phylogenetic pattern to the β sheet (Sigurdadóttir et al. 1992). These differ-
ences have been explained by gene conversion-like events both between loci and between alleles within loci (Erlich and Gyllensten 1991; Sigurdadóttir et al. 1992). Recombination between alleles within loci seems to happen more frequently than interlocus recombination.

We have characterized and analyzed exon 2 of 22 MHC class II B restriction fragment length polymorphism (RFLP) haplotypes identified in four hatchery populations of Baltic salmon (Atlantic salmon feeding in the Baltic Sea, *Salmo salar*; Langefors et al. 1998, 2000). The hatcheries are located in the rivers Umeälven and Luleälven in northern Sweden, Dalälven in central Sweden, and Mörrumsån in southern Sweden. Atlantic salmon colonized the rivers after the last glaciation, and commercial culture started in the 1950s. In these populations, sexually mature adult Atlantic salmon are captured in the rivers on their return from the sea and are used in the hatcheries for artificial culture where random mating is performed. At 2 years of age, the artificially hatched salmon are released and allowed to migrate to the sea. After some years in the Baltic Sea, sexual maturation starts and the salmon migrate back to their respective natal river to spawn at an age of 4–10 years (Erikkson and Eriksson 1993). Hence, the populations can be termed “seminatural” in that they reproduce artificially in captivity but spend their adult life in the wild, exposed to a natural set of pathogens in the marine environment. Population mixing is avoided but during a few years in the early 1970s, individuals from, e.g., Dalälven were used in the Luleälven hatchery. At allozyme loci, Atlantic salmon show a more limited level of variation and differentiation than other salmon species (Ståhl 1987). However, modern DNA methods, such as microsatellite and mtDNA analyses, reveal higher levels of variation as well as differentiation between populations (Nielsen et al. 1996; Nilsson 1997). The four hatchery populations studied here are highly differentiated in their MHC class II B allele frequencies, with large genetic distances and significant divergences between populations (Langefors et al. 1998).

Our present aims were to analyze the amount of sequence variation in MHC class II B in Baltic salmon, to evaluate the importance of microrecombination in the generation of MHC variation, and to estimate the level and strength of balancing selection. By studying several populations, we draw some conclusions about population differentiation in salmon and also about the age of MHC alleles.

**Materials and methods**

**Samples**

Juvenile Atlantic salmon were collected at the hatcheries in Mörrumsån, Dalälven, Umeälven, and Luleälven along the Swedish Baltic sea coast in 1993–1995 (Langefors et al. 1998). The fish were deep-frozen until DNA extraction. Blood and/or milt samples were also collected from a number of adult fish from the River Mörrumsån. Immediately after collection, blood and milt were placed in SET-buffer (0.15 m NaCl, 0.05 m Tris, pH 7.5 and 0.001 m EDTA, pH 8.0) and frozen at −80°C. DNA was extracted using phenol-chloroform. We sequenced two copies from each of the formerly identified 22 RFLP haplotypes, either by cloning procedures or by demanding gradient gel electrophoresis (DGGE) and direct sequencing (Langefors et al. 2000).

**PCR procedures**

For amplification of exon 2, we used the primer pair MG 7 (5'-TAATCTGACAAAAACATGA-3') and AL 1002 (3'-CAC CTTGTTCTAGTTGTATG-5') for cloning and the primer pair TVS 4501 (5'-TCTTATTATGTCTTTCTCC-3') and MG 14 (5'-AL 1002 with a 40-bp long GC-clamp in the 3' end; Sheffield et al. 1989) for DGGE (Langefors et al. 2000). MG 7 together with AL 1002 amplifies 24 bp of intron 1 and the complete exon 2 while TVS 4501 is located at the 21 last nucleotides of intron 1 and together with MG 14 amplifies the complete exon 2 (Langefors et al. 2000). The amplification was performed in a Perkin-Elmer-Cetus 9600, in 25 µl PCR reaction mixture containing 100 ng of DNA template, 25 pmol of each of the primers MG 7 and AL 1002 or 20 pmol each of TVS 4501 and MG 14, 0.9 µg bovine serum albumin, 1X PCR reaction buffer, 37.5 nmol MgCl2, with MG 7 and AL 1002 or 25 nmol MgCl2 with TVS 4501 and MG 14, 125 µM dNTP and 0.2 units Taq Polymerase (Perkin Elmer, New Jersey). PCR conditions were as follows: denaturation for 2 min at 94°C, followed by 30 cycles of denaturation for 30 s at 94°C, annealing for 30 s at 53°C for MG 7 and AL 1002 and 56°C for TVS 4501 and MG 14, and primer extension for 30 s at 72°C. The final extension was for 10 min at 72°C. When run on an agarose gel, the PCR products showed one distinct band with the approximate length of 300 bp. PCR products resulting from amplification with TVS 4501 and MG 14 were run on 5% 19:1 acrylamide: bisacrylamide parallel DGGE gels containing 1× TAE buffer and a 30–50% gradient of urea and formamide at 90 V for 15 h (Langefors et al. 2000; Myers et al. 1987).

**Sequencing procedures**

PCR fragments resulting from amplification with primers MG 7 and AL 1002 from seven individuals selected for typing of the known RFLP haplotypes (see above) were ligated into the pCR 2.1 vector (Invitrogen, The Netherlands) and transferred into the bacterial strain DH5α. Eight to ten selected positive clones from each individual (in total 60 clones) were used as templates in DNA sequencing reactions that were set up using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer Corporation, California) or ABI PRISM dRhodamine Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer).

Desired DGGE fragments (two or occasionally one from each of 25 individuals selected for characterization of all haplotypes; 41 fragments in total) were identified, cut out from the gel, mixed with 160 µl ddH2O, extracted by keeping the samples for 1 h at −80°C, 1 h at 25°C, 1 h at −80°C, 2–3 h at 4°C, and used as template in a PCR reaction with TVS 4501 and AL 1002 as above but in a 50-µl reaction volume. The PCR products were diluted with 0.3 M NaAc and 2.5 vol 99% EtOH, dissolved in 12–16 µl ddH2O and used as templates in DNA sequencing reactions in the ABI PRISM 310 as described above. For each DGGE fragment or clone, we sequenced both directions using TVS 4501 as primer for the 5' strand and primer AL 1002 for the 3' strand. Sequencing reactions were run in the ABI PRISM 310 Genetic Analyzer and edited with the Gene Scan Analysis Software (Perkin Elmer).