GEOMAR
Helmholtz Centre for Ocean Research Kiel

Characterization and evolution of MHC II genes in the European eel (Anguilla anguilla)

M.Sc. Thesis

Seraina Emilia Bracamonte

8th October 2013

Supervisor: Dr. Christophe Eizaguirre

Abstract

Natural populations are exposed to a diverse community of parasites. As a consequence, hosts have evolved defence strategies. In jawed vertebrates, the major histocompatibility complex (MHC) is an essential part of the immune system. These genes show an exceptional polymorphism which is assumed to be maintained mainly by parasite-mediated selection but also by natural and demographic events. At the sequence level, this usually results in gene duplications and positive selection at sites involved in binding of parasite derived antigens. At the population level, balancing selection is assumed to maintain of a large and diverse allele pool. However, genetic drift can have a major impact on the evolution and population structure at these genes and selection might become undetectable in strongly differentiated populations. In the European eel on the other hand, there is extensive gene flow due to its near-panmictic life history, wherefore it might be a particularly suitable system to study the effects of natural selection on MHC evolution in the absence of confounding neutral processes. In this species, up to four expressed MHC IIA and MHC IIB loci were isolated, suggesting the presence of at least two potentially functional loci. Both genes show signs of positive selection and recombination, indicating that natural selection was an important factor for their evolution. However, in contrast to the current view, MHC IIA seemed to experience stronger selection than MHC IIB. Nevertheless, an extraordinary variability of MHC IIB alleles was found in a large population survey. The different populations did not show signs of divergent selection on the MHC pool. However, the differences did not correlate with neutral distances, suggesting some form of selection on MHC but its drivers remain elusive.

Zusammenfassung

Natürliche Populationen sind einer diversen Parasitengemeinschaft ausgesetzt und haben folglich effiziente Abwehrmechanismen entwickelt. In Kiefermäulern ist der Haupthistokompatibilitätskomplex (MHC) ein wichtiger Bestandteil dieser Immunabwehr. Die Gene des MHC zeigen eine ausserordentliche Variabilität, die im Allgemeinen von der Selektion durch Parasiten aufrecht erhalten wird. Allerdings beeinflussen auch neutrale und demographische Prozesse die Zusammensetzung des Allelepools in den jeweiligen Populationen. Als Folge dieser Selektion sind die Gene meist dupliziert und mehrere Kopien im Genom vorhanden. Ausserdem zeigen Codons, die direkt an der Erkennung von parasitären Antigenen beteiligt sind, häufig Anzeichen starker positiver Selektion. In Populationen und Metapopulationen führt die Selektion durch Parasiten zur Erhaltung eines diversen Allelpools und lokaler Anpassung. Die Zusammensetzung und Divergenz der Allelpools in und zwischen den Populationen hängt allerding auch von Faktoren wie genetischer Drift ab, die starke Auswirkungen auf die Evolution der Gene haben kann, wodurch Selektion nicht mehr nachweisbar ist. Der panmiktische Lebenszyklus des Europäischen Aals führt zu starkem Genfluss in der Gesamtpopulation, weshalb die Auswirkungen natürlicher Selektion nicht von neutralen Faktoren überlagert werden sollten. Bis zu vier exprimierte MHC IIA und MHC IIB Allele wurden in einzelnen Individuen des Europäischen Aals isoliert, was darauf hin deutet, dass mindestens zwei Loci pro Gen vorhanden sind. Beide Gene zeigen klare Anzeichen von positiver Selektion und Rekombination. Im Gegensatz zur generellen Auffassung, scheint in Europäischen Aal MHC IIA stärker selektiert zu werden als MHC IIB. Trotzdem zeigt MHC IIB eine hohe Variabilität in der Gesamtpopulation, die zwischen den einzelnen Lokalpopulationen keine signifikanten Unterschiede in der Zusammensetzung des Allelepools aufweist. Die Differenzen korrelieren jedoch nicht mit den Unterschieden neutraler Marker, was darauf hinweist, dass MHC IIB selektivem Druck ausgesetzt ist, dessen Faktoren aber nicht ersichtlich sind.

Table of content

General introduction

Natural populations are exposed to a large diversity of parasites (Wegner *et al.* 2003; Evans & Neff 2009; Eizaguirre *et al.* 2011; Raeymaekers *et al.* 2013). Their exploiting life history can cause severe reductions in host fitness, therefore, hosts have evolved mechanisms to overcome this challenge. In vertebrates, one of the key features of the immune system is the major histocompatibility complex (MHC) (Janeway *et al.* 2001). Its genes code for cell surface proteins that bind foreign antigens and present them to T cells, initiating an adaptive immune response. There are two classes of MHC molecules. MHC class I presents antigens of intracellular parasites, mainly viruses, to cytotoxic T cells, triggering the cell-mediated immune response whereas MHC class II presents antigens derived from extracellular parasites, such as bacteria, metazoans, and protozoans, that have been endo- or phagocytosed, to helper T cells, initiating the humoral immune response. MHC class I molecules are expressed on most nucleated cells and are composed of a membrane-spanning α-chain that folds into three extracellular domains and a non-covalently associated β_2-microglobulin. MHC class II molecules comprise an α-chain and a β-chain each of which folds into two extracellular domains and spans the membrane. They are expressed exclusively on antigen-presenting cells, such as macrophages, B-lymphocytes, and dendritic cells.

While MHC genes are tightly linked in a gene dense region in all other vertebrates, fish MHC class II genes have been translocated and, as a consequence, the two classes of MHC are located in different genomic regions (Sato *et al.* 2000; Kuroda *et al.* 2002).

The MHC genes are the most polymorphic genes in the vertebrate genome (Wegner 2008; Eizaguirre & Lenz 2010). They show exceptionally high sequence variation among alleles and a large number of alleles within and among populations. Additionally, the genes are often duplicated. The number of loci, however, varies considerably among and within species ranging from one expressed MHC class I and MHC class II locus in salmonids to >10 MHC class IIB loci in cichlids and about 100 MHC class I loci in the cod genome (Klein *et al.* 1998b; Malaga-Trillo *et al.* 1998; Shum *et al.* 2001; Star *et al.* 2011). Despite this large variation, there seems to be an optimal number of alleles per individuals which might be induced by the necessity of an appropriate T cell repertoire (Woelfing *et al.* 2009).

The polymorphism is created by positive selection and recombination (Wegner 2008). Positive selection is driven by parasites and mainly acts on antigen-binding sites (ABS) which show the highest rate of non-synonymous over synonymous substitutions (Bahr & Wilson 2011; Sin *et al.* 2012). Recombination will then shuffle the genetic variation into new combinations, thereby increasing the allele repertoire (Shum *et al.* 2001; Spurgin *et al.* 2011).

Balancing selection induced by the co-evolution of hosts and parasites is thought to maintain the large MHC diversity. Additionally, it is invoked to explain the maintenance of ancient allelic lineages through speciation events resulting in trans-species polymorphism (Klein *et al.* 1998a; Lenz 2011) which has been observed in many vertebrate taxa (Glaberman & Caccone 2008; Kikkawa *et al.* 2009; Kiemnec-Tyburczy *et al.* 2010; Lenz *et al.* 2013a).

Three main hypotheses have been proposed to maintain the polymorphism at MHC genes: negative frequency-dependent selection, heterozygote advantage, and spatiotemporal variation in parasite fauna (Bernatchez & Landry 2003; Eizaguirre & Lenz 2010; Spurgin & Richardson 2010).

Rare alleles are assumed to increase the fitness of their carrier, because common parasites will adapt to exploit the most common host genotype giving rare or unfamiliar alleles an advantage, a phenomenon termed negative frequency-dependent selection (van Valen 1973; Kubinak *et al.* 2012). This mechanism is expected to lead to cyclic fluctuations of host genotypes due to selection and deselection of specific alleles by parasites (Decaestecker *et al.* 2007). The hypothesis relies on the assumption of tight interactions between host and parasite genotypes in determining resistance and susceptibility and a resistance allele might increase in frequency regardless of its initial frequency (Dionne *et al.* 2009; Eizaguirre *et al.* 2012b; Luijckx *et al.* 2013).

However, rare alleles might not confer resistance per se, but rather because they usually occur in heterozygotes. Under the hypothesis of heterozygote advantage, heterozygous individuals are fitter than homozygous when exposed to a diverse parasite community because they can resist a wider variety of parasites (Doherty & Zinkernagel 1975). Increased resistance can be due to a dominant or overdominant effect and in nature the latter is observed more commonly (Evans & Neff 2009; Kekalainen *et al.* 2009; Oliver *et al.* 2009). A special case of heterozygote advantage is the diverse allele advantage which states that heterozygous individuals carrying divergent alleles have an increased fitness over individuals carrying similar alleles (Wakeland *et al.* 1990; Lenz 2011). The reasoning behind this assumption is that dissimilar alleles will bind dissimilar antigens from different parasites, whereas similar alleles will bind antigens derived from similar parasites. Divergent alleles within individuals have indeed been found to be associated with better body conditions and survival (Lenz *et al.* 2009; Lenz *et al.* 2013b).

At the metapopulation level, a diverse allele pool might be maintained by spatial and temporal variation in parasite fauna (Hedrick 2002). Different parasite communities will most likely exert different selective pressure on host populations which will result in local adaptation of both, parasites and hosts. Divergence of MHC gene pools among populations of different habitats are common and, in populations for which parasite diversity was assessed, differences in MHC pools and parasite communities seem to be correlated (Dionne *et al.* 2007; Fraser & Neff 2010; Cammen *et al.* 2011; Eizaguirre *et al.* 2012a). However, in species with limited gene flow among populations, the effects of selection on MHC are inferred from comparisons with differentiation at neutral markers and at larger geographic scales or in fragmented populations neutral processes seem to be strong enough to mask signs of selection (Landry & Bernatchez 2001; Aguilar & Garza 2006; Alcaide 2010; Spurgin & Richardson 2010).

The European eel on the other hand, does not show strong genetic differentiation at neutral markers across its distribution range due to its panmictic or near-panmictic life history (Wirth & Bernatchez 2001; Palm *et al.* 2009; Als *et al.* 2011; Pujolar *et al.* 2011a) and might

prove to be an excellent system to study the sole effects of natural selection on the differentiation of MHC alleles among individuals and populations.

The European eel is a catadromous species that spends most of its life in European and North African freshwater systems and estuaries (Tesch 2003). Its distribution ranges from as far north as Iceland, the North Sea, and Baltic Sea to the southern limits of the Canary Islands and the Azores. In the east, the Black Sea forms the boundary of its distribution. The life cycle of the European eel includes two extended migrations across the Atlantic Ocean. After several years of foraging in the freshwater systems, adult eels silver and commence their spawning migration towards the Sargasso Sea where they spawn once and then die (Schmidt 1923; Tesch 2003; Aarestrup *et al.* 2009). The leptocephali larvae are then drifted with the ocean currents towards the European continent. After reaching the shelf, they metamorphose into glass eels, enter freshwater, and migrate upstream

The European eel is a commercially important species that is declining rapidly since the early 1980s and does not show signs of potential recovery (Dekker 2003; ICES 2012) wherefore it is listed as critically endangered on the IUCN red list (www.iucnredlist.org). Several causes are assumed to be responsible for the decline and act synergistically (Wirth & Bernatchez 2003). These causes include anthropogenic factors, such as overfishing and river constructions that impede the migration of glass eels and silver eels (Dekker 2003; ICES 2012), climatic and oceanic changes, especially in the Sargasso Sea (Knights 2003; Friedland *et al.* 2007), and the occurrence of a new parasite, the invasive swim bladder nematode *Anguillicola crassus* (Kirk 2003).

Anguillicola crassus is a parasite originally endemic to the Japanese eel, *Anguilla japonica*. It was detected in Europe for the first time in Germany in 1982 and since then it has expanded to the entire distribution range of the European eel (Kirk 2003). This rapid spread was likely mediated by trafficking and restocking of infected eels. The European *A. crassus* populations are introduced from one source population in Taiwan and differentiated into three geographically distinct clusters, which are consistent with the boundaries of the invertebrate intermediate hosts (Wielgoss *et al.* 2008). Furthermore, they have adapted to its new host, the European eel (Weclawski *et al.* 2013).

Unlike its native host, the Japanese eel (*A. japonica*), the European eel is highly susceptible to *A. crassus* infections and the prevalence in wild populations stabilized at 60 – 70 % (Lefebvre & Crivelli 2004). Eels get infected with *A. crassus* from the glass eels stage onwards by feeding on the parasite´s intermediate host (De Charleroy *et al.* 1990). The nematode then migrates to the swim bladder where it attaches to the wall and feeds on the blood of its host. The swim bladder of infected eels is characterized by a reduced lumen and thickened wall, which can lead to complete damage of the organ (Kirk 2003).

Although *A. crassus* does not generally seem to have severe impacts on the continental phase of the European eel (Lefebvre *et al.* 2013), in the presence of additional stressors such as hypoxia or increased water temperatures it might cause mass mortalities within the population (Kirk 2003; Gollock *et al.* 2005). Furthermore, infected eels show reduced swimming speed and higher energy consumption than healthy individuals and as a

consequence, they are assumed to be unable to reach the spawning grounds to reproduce (Palstra *et al.* 2007).

The European eel´s susceptibility can be partly explained by its incapability of mounting an adequate immune response (Knopf *et al.* 2000). The onset of a specific immune response is late and it is inefficient, however the variability of the response is large among individuals and might reflect genetic differences in immune genes such as the MHC.

This study aimed at characterizing MHC class II genes in the European eel to provide insights into their evolutionary history and then assessing differentiation across populations at small and large geographic scales.

Chapter 1: Characterization of MHC II genes

1. Introduction

The major histocompatibility complex (MHC) is a polymorphic multigene family that encodes proteins crucial for the immune system of vertebrates. MHC molecules present foreign antigens to T cells and thereby initiate the adaptive immune response (Janeway *et al.* 2001). MHC molecules are divided into class I and class II proteins with different structures and functions. MHC class I molecules are expressed on most nucleated cells and consist of a large membrane-spanning α-chain that folds into three distinct domains and a non-covalently associated β_2-microglobulin. They present antigens derived from cytosolic pathogens such as viruses. MHC class II molecules bind antigens from extracellular or vesicular parasites, such as metazoans and bacteria, and present them to helper T cells. In contrast to MHC I, MHC class II proteins are only expressed on antigen-presenting cells such as macrophages, B-lymphocytes, and dendritic cells. MHC II molecules are heterodimers consisting of two membrane-spanning chains, the α- and the β-chain, each folding into two extracellular domains. The α_1- and β_1-domains form the antigen-binding groove, whereas the β_2-domain is involved in the interaction with the helper T cell receptor. The antigen-binding groove of MHC II molecules is open at both ends and, therefore, the bound antigens are not restricted in size. However, the usual length of the presented antigens is 13-17 amino acids (Brown *et al.* 1993; Janeway *et al.* 2001). They are bound along the backbone by conserved residues of the antigen-binding groove and additional anchoring is provided by polymorphic pockets holding side chains of the antigen.

The teleostean MHC class I and MHC class II genes, in contrast to other vertebrates, are not linked in one genomic cluster, but spread throughout the entire genome (Sato *et al.* 2000; Wegner 2008), probably as a result of translocation of the MHC class II genes (Kuroda *et al.* 2002). Despite the translocation of loci to different chromosomes, the general structure of the MHC II genes is conserved. MHC IIA consists of four exons separated by three introns (Stet *et al.* 2002; Xu *et al.* 2009; Xu *et al.* 2011), whereas MHC IIB is composed of five (e.g half-smooth tongue sole Xu *et al.* 2009; orange-spotted grouper Lu *et al.* 2012) or six exons (e.g. cichlids Ono *et al.* 1993; miiuy croaker Xu *et al.* 2011).

The MHC genes are the most polymorphic genes in the vertebrate genome with respect to sequence variation among alleles as well as differentiation of alleles among individuals and populations (reviewed in Bernatchez & Landry 2003; Eizaguirre & Lenz 2010; Spurgin & Richardson 2010). The diverse MHC pool is thought to be maintained by balancing selection via negative frequency-dependent selection (van Valen 1973), heterozygote advantage (Doherty & Zinkernagel 1975) and as a special case divergent allele advantage (Wakeland *et al.* 1990; Lenz 2011), or spatial and temporal variation in parasite fauna (Hedrick 2002).

One common characteristic of the MHC is that several loci are present within the genome, which are thought to arise from gene duplications in a birth-and-death process (Nei *et al.* 1997; Klein *et al.* 1998b; Reusch *et al.* 2004; Sato *et al.* 2012). The number of loci, however, varies among species for both, MHC class I and MHC class II genes. While cod seems to have

entirely lost MHC II genes (Star *et al.* 2011), salmonids have only one expressed locus for both, MHC IIA and MHC IIB genes (Stet *et al.* 2002). In contrast, cyprinids seem to express two MHC IIA and MHC IIB loci (van Erp *et al.* 1996; Kruiswijk *et al.* 2004), sticklebacks 2-4 MHC IIB loci (Reusch & Langefors 2005), and cichlids possess up to 13 MHC IIB and 8 MHC IIA loci per haplotype (Malaga-Trillo *et al.* 1998; Murray *et al.* 2000), however, not all of them seem to be expressed. The number of loci may also vary among individuals within species (Malaga-Trillo *et al.* 1998; Murray *et al.* 2000; Lenz *et al.* 2009; Eizaguirre *et al.* 2011), probably as a response to parasite-mediated selection.

Generally, the extreme polymorphism observed at MHC genes seems to be maintained by parasite selection (Milinski 2006; Spurgin & Richardson 2010). Consequently, exons encoding the antigen-binding groove show the highest polymorphism and usually strong signs of positive selection at particular antigen-binding sites (ABS). On the other hand, exons encoding domains that are essential for the molecule´s structure and stability and sites involved in the interaction with T cells are more conserved (Xu *et al.* 2011; Bahr & Wilson 2012). For MHC II, the ABS is encoded by exon 2 of both, MHC IIA and MHC IIB, with MHC IIB generally being more variable than MHC IIA (Wegner 2008). Within exon 2, polymorphic sites are concentrated in codons of the ABS (Reusch & Langefors 2005; Wegner 2008; Bahr & Wilson 2011).

The second mechanism generating the high diversity found at MHC is recombination and gene conversion within and between loci (Reusch & Langefors 2005; Wegner 2008; Bahr & Wilson 2012). These processes may shuffle exons into new combinations and will create alleles that share short sequence motives. Furthermore, they seem to be powerful mechanisms to increase allelic diversity after a bottleneck or founder event (Spurgin *et al.* 2011).

However, assessing the effects of natural selection is not trivial due to neutral and demographic processes that influence the population structure (Bernatchez & Landry 2003; Spurgin & Richardson 2010). The European eel is a good system to study the sole effect of natural selection on the evolution of MHC genes without confounding effects due to its panmictic life history (Palm *et al.* 2009).

The European eel (*Anguilla anguilla*) is a catadromous species which is widely distributed across Europe and North Africa (Tesch 2003). After completing the continental phase adult eels migrate across the Atlantic Ocean to spawn in the Sargasso Sea. The larvae are then drifted back to the European continent where they enter the freshwater systems resulting in a panmictic or nearly panmictic population across the species´ distribution range (Wirth & Bernatchez 2001; Palm *et al.* 2009; Als *et al.* 2011). During the last decades the recruitment of the European eel has been declining to less than 1 % of the pre-1980 level (ICES 2012). Coinciding with this decline was the introduction and rapid spread of the swim bladder nematode *Anguillicola crassus* in the early 1980s (Kirk 2003). Despite its supposedly strong impact on the spawning migration (Palstra *et al.* 2007) the European eel is incapable of mounting an adequate immune response (Knopf *et al.* 2000).

In order to study the role of parasite-mediated selection on the evolution of MHC genes in eels and offer tools to understand the impact of the invasive parasite on the species viability, the genetic structure of MHC II genes was characterized in this species.

2. Material & Methods

2.1. Nucleic acid extraction and cDNA synthesis

DNA was extracted from muscle tissue of four adult eels and six glass eels using the Qiagen DNeasy® Blood and Tissue kit (Hilden, Germany) following the manufacturer´s instructions, except that DNA was eluted twice in the same 200 µl of AE buffer.
RNA was extracted from liver tissue of the same four adult eels used for DNA extraction. Extractions were performed with the InviTrap® Spin Tissue RNA Mini kit (STRATEC Molecular, Berlin, Germany) using the protocol for spleen, kidney and lung tissue. The Lysis solution TR was adjusted with 1/100 volume of β-mercaptoethanol instead of DTT. Tissue was disrupted on a homogenizer and RNA was eluted twice with 30 µl of Elution buffer R to a final volume of 60 µl.
Reverse transcription of RNA into cDNA was performed using the QuantiTect® Reverse Transcription kit (Qiagen, Hilden, Germany), following the manufacturer´s instructions but using twice the amount for a final volume of 40 µl and 1.4-1.9 µg of RNA.

2.2. Primer design

MHC II initiates the adaptive immune response against extracellular parasites. It is a heterodimer consisting of an α- and a β-chain encoded by the MHC IIA and MHC IIB genes, respectively. Exons 2 and 3 of both genes encode the extracellular domains of which the α_1- and β_1-domains form the ABS.

MHC class IIA

Primers that would amplify exon 2 and exon 3 of MHC IIA were designed in conserved regions of an alignment containing eleven sequences from EeelBase (Coppe *et al.* 2010) (contig IDs: eeel2_rep_c5554, eeel2_s5274, eeel2_rep_c6479, eeel2_s8066, eeel2_s5630, eeel2_s8370, eeel2_rep_c6276, eeel_rep_c58577, eeel2_s8017, eeel2_s8564, eeel2_rep_c5060) and ten teleost sequences obtained from NCBI (*Oncorhynchus mykiss* [gi|8920254], *Salmo salar* [gb|L77086.1], *Cynoglossus semilaevis* [gi|349502471], *Epinephelus coioides* [gi|326632480], *Sparus aurata* [gi|66096123], *Pagrus major* [gb|AY698064.1], *Barbus intermedius* [gi|45433850], *Gasterosteus aculeatus* [gi|51449916], *Stizostedion vitreum* [gb|AY158870.1]). Four of the sequences from Eeelbase had shifts in their reading frame and were separated into two sequences, one containing the upstream part of the frame shift and the other containing the downstream part. Exon boundaries were determined according to the domains in Reusch et al. (2004). Primer sequences and melting temperatures are listed in table 1.1 and their locations on the gene are given in figure 1.1.

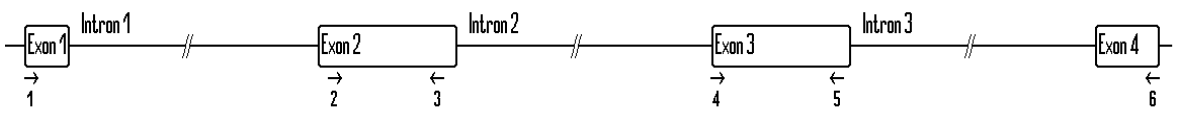


Figure 1.1 Schematic representation of MHC IIA. Arrows indicate the position and orientation of the PCR primers, numbering is according to table 1.1.

Table 1.1 PCR primers used to amplify MHC IIA.

	Primer name	Sequence 5´-3´	T_M (°C)	Location
1.	AaMHCIIAE1F1	ATGAATCACTCCATGTTCACAGC	61.0	Exon 1
2.	AaMHCIIAE2F	AGCCTGCCRGACCAATGACACTGCTC	72.0	Exon 2
3.	AaMHCIIAE2R	TGAACARSCAYGTYCAAGTTATYAAT	62.0	Exon 2
4.	AaMHCIIAE3F1	GATCCTCCTCAGAGYACAATCTWTTC	65.0	Exon 3
5.	AaMHCIIAE3R	GCCCCTTGTGCTCCACGCTGCAGGAA	74.0	Exon 3
6.	AaMHCIIAE4R	GTTGCAGTTATTTCCTTTGATGAGGA	63.0	Exon 4

T_M, melting temperature

For designing primers in exon 1 and exon 4 (table 1.1, figure 1.1), the same alignment was used. However, exon boundaries were re-examined using annotations in NCBI which lead to slightly different starting and ending positions of all exons.

Table 1.2 PCR primers used to amplify MHC IIB from genomic and cDNA.

	Primer name	Sequence 5´-3´	T_M (°C)	Location
1.	AaMHCIIBE1F4	CTTTATTTACTTTACTGCTTGGA	56.0	Exon 1
2.	AaMHCIIBE2F3	AGTGYCGTTTCAGYTCCAGMGAYCTG	69.0	Exon 2
3.	AaMHCIIBE2F	GACCTGCAGGACCTCGAGTWCATTG	69.0	Exon 2
4.	MHCIIaaF1	TATTGGCTACACTGCACTTGGAGTG	62.0	Exon 2
5.	AaMHCIIBE2R	ATTCAGCATTAGGCCTGCAGTAA	61.0	Exon 2
6.	AaMHCIIBE2R2	CTCACYTGRMTWATCCAGTATGG	62.0	Exon 2
7.	AaMHCIIBI2R	CAAGTTTAAGCTTCWGAACATATTTG	60.0	Intron 2
8.	AaMHCIIBE3F3	TTCTACCCCAGAGGAATCAAAATGAC	65.0	Exon 3
9.	AaMHCIIBE3R3	TGGATCTGATAGTACCAGTTTCCAT	63.0	Exon 3
10.	AaMHCIIBE3R2	TGCTCCACCWKGCAGGAGATTTKCTC	69.0	Exon 3
11.	AaMHCIIBE4R	CGGTGGATTTCTTCTTGTAATAGA	60.0	Exon 4

T_M, melting temperature

MHC class IIB

For designing primers in conserved regions of exons 2 and 3 of MHC IIB, three sequences from EeelBase (Coppe *et al.* 2010) (contig IDs: eeel_rep_c36960, eeel2_c1172, eeel2_c3234) were aligned with 13 published teleost sequences (*Oncorhynchus mykiss* [gi|306966138], *Salmo salar* [gi|57338492], *Salvelinus alpines* [gi|205933548], *Dicentrarchus labrax* [gi|115738470], *Cyprinus carpio jian* [gb|HQ380378.1], *Epinephelus akaara* [gb|EU399184.1], *Morone saxatilis* [gb|L33967.1], *Leiocassis longirostris* [gb|GQ478337.1], *Oreochromis niloticus* [gb|JN181166.1], *Gasterosteus aculeatus* [gi|51449916], *Stizostedion vitreum* [gb|AY158837.1 and gb|AY158838.1]). Two of the eel sequences were split into two sequences, because they had a shift in their reading frame when comparing them to NCBI. Domain boundaries were identified according to Reusch et al. (2004). Sequences and melting temperatures are given in table 1.2 and positions on the gene are indicated in figure 1.2.

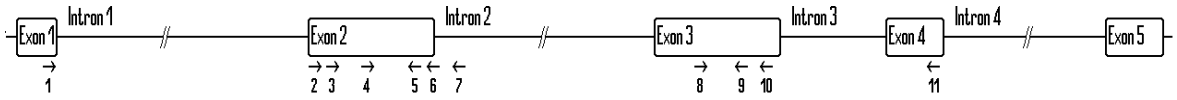


Figure 1.2 Scheme of MHC IIB. Primer positions and orientation are indicated by arrows according to the numbering in table 1.2.

For the design of primers in exon 1 and exon 4 (table 1.2, figure 1.2) a similar alignment was used which included four additional teleost sequences: *Coregonus clupeaformis* (gi|317119976), *Salvelinus fontinalis* (gi|186898224), *Lutjanus argentimaculatus* (gb|JQ655274.1), and *Barbus intermedius* (gi|45433844). Additionally, the exon annotation has been double-checked with annotations in NCBI and the exon boundaries do not exactly correspond to the domains outlined in Reusch et al. (2004). The primer for exon 4 was designed in a conserved region of the teleost alignment, whereas the primer in exon 1 was designed at the end of the exon due to the lack sequence information for eels.
In addition, primers that would amplify exon 2 and part of intron 2 were designed based on the sequences obtained with the previous primers (table 1.2, figure 1.2).

2.3. Cloning and sequencing

PCRs were carried out in a 20 µl reaction containing 1× DreamTaq buffer (Thermo Scientific, St. Leon-Rot, Germany) including 20 mM $MgCl_2$, 0.5 mM of each dNTP, 0.5 µM of each primer, 0.5 U DreamTaq polymerase, and 1 µl of template DNA. Cycling conditions were as follows: an initiation step of 95 °C for 1 min, 35 cycles of 95 °C for 15 s, annealing temperature for 15 s, and 72 °C for 30 s - 1 min, depending on the expected fragment length, followed by a final extension of 72 °C for 15 min. The annealing temperatures and extension times are given in tables 1.3 and 1.4. In order to reduce PCR artefacts two independent PCR reactions were pooled (Lenz & Becker 2008)and run on a 1.5 % agarose gel at 45 V for 4 h, excised, and purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany) with slight modifications. 300 µl of NTI buffer were added per 100 mg of agarose gel and

Table 1.3 PCR conditions for the amplification of MHC IIA.

Primer combination		T_A (°C)	t_E (s)
AaMHCIIAE1F1	AaMHCIIAE3R	68.0	60
AaMHCIIAE2F	AaMHCIIAE2R	61.0	30
AaMHCIIAE2F	AaMHCIIAE3R	74.0	60
AaMHCIIAE3F1	AaMHCIIAE4R	63.0	60

T_A, annealing temperature, t_E, extension time

Table 1.4 PCR conditions for amplifying MHC IIB.

Primer combination		T_A (°C)	t_E (s)
AaMHCIIBE1F4	AaMHCIIBE3R3	61.0	60
AaMHCIIBE1F4	AaMHCIIBE4R	61.0	60
AaMHCIIBE2F3	AaMHCIIBE2R2	59.0	30
AaMHCIIBE2F3	AaMHCIIBI2R	61.0	30
AaMHCIIBE2F	AaMHCIIBE2R	65.0	30
AaMHCIIBE2F	AaMHCIIBE3R3	66.0	60
AaMHCIIBE2F	AaMHCIIBE3R2	69.0	60
MHCIIaaF1	AaMHCIIBE2R	64.0	30
MHCIIaaF1	AaMHCIIBE3R2	64.0	60
AaMHCIIBE3F3	AaMHCIIBE4R	64.0	60

T_A, annealing temperature, t_E, extension time

incubated at 50 °C for 10 min. The washing step consisted of twice 500 µl and once 150 µl of NT3 buffer, followed by a drying step of 2 min at 11000 rpm. The samples were then incubated for 5 min at 60 °C. DNA was eluted in 30 µl of pre-warmed NE buffer.

Purified PCR products were cloned into chemically competent cells with the Qiagen® PCR Cloning^{plus} kit (Hilden, Germany) or the TOPO® TA Cloning® kit for subcloning, with TOP10 competent cells (Invitrogen, Darmstadt, Germany). Ligation was performed according to the manufacturer´s instructions but reducing the recommended amount by half for both kits. Incubation for the Qiagen PCR cloning kit was at 4 °C for 1 h or 2.5 h and for the TOPO TA cloning kit at room for 30 minutes. Transformation followed the manufacturer´s recommendations with slight modifications. Five and 3 µl ligation reaction were added to 25 µl of competent cells for the Qiagen PCR cloning kit and the TOPO TA cloning kit, respectively. Cells were incubated on ice, heat-shocked at 42 °C, and then 150 µl of S.O.C. medium were added. Cells were incubated at 37 °C for 30 min and plated on selective plates containing kanamycin or ampicillin antibiotics. Eight to 48 positive colonies were picked per sample and transferred in 25 µl of HPLC water. Clones were denatured at 95 °C for 10 min and the insertion size was checked on a gel after PCR amplification. PCR products of 8 - 32 clones per sample were cleaned with ExoSAP (Thermo Scientific, St. Leon-Rot, Germany) and sequenced on an ABI 3130xl Genetic Analyser (Applied Biosystems, Darmstadt, Germany) with the ABI BigDye Terminator v3.1 Cycle Sequencing kit using 0.25 - 0.5 µl of BigDye mix and 1 µl of PCR product. Cycling conditions were 1 min at 96 °C followed by 25 cycles of 10 s at 96 °C, 5 s at 50 °C and 6 min at 60 °C.

For MHC IIA, the forward primer AaMHCIIAE2F was used in combination with the reverse primers AaMHCIIAE2R and AaMHCIIAE3R. Furthermore, the primer combinations AaMHCIIAE1F / AaMHCIIAE3R and AaMHCIIAE3F1 / AaMHCIIAE4R were used to amplify exon 1 and exon 4, respectively. Since cloning of gDNA from glass eels using forward and reverse primers in exons 2 and 3, respectively, was not successful due to the large size of intron 2, amplification and cloning for all primer combinations were performed on cDNA of the adult eels. Coding sequences containing the putative start codon and ending right before the putative stop codon were assembled by overlapping the fragments obtained with the different primer combinations.

For MHC IIB, the primer combination AaMHCIIBE2F / AaMHCIIBE3R2 was used on gDNA from glass eels and each of the forward primers AaMHCIIBE2F and MHCIIaaF1 were used with each of the reverse primers AaMHCIIBE2R and AaMHCIIBE3R2 on adult gDNA. However, cloning of fragments including intron 2 was only marginally successful for adult eels. The primer combinations AaMHCIIBE2F / AaMHCIIBE2R, AaMHCIIBE2F / AaMHCIIBE3R2, and AaMHCIIBE2F / AaMHCIIBE3R3 were further used on cDNA of the adult eels to confirm transcription of the identified alleles. For the amplification of downstream regions of exon 3, the primers AaMHCIIBE3F3 and AaMHCIIBE4R were used on gDNA and the primers AaMHCIIBE1F4 / AaMHCIIBE3R3 were used on cDNA to amplify the 5´-end of exon 2.

2.4. Data analysis

The sequences were blasted against the NCBI database for verification and then aligned to other teleost sequences in BioEdit (Hall 1999). Calculations of the best-fit substitution model and mean nucleotide distance were performed in MEGA v5.20 (Tamura *et al.* 2011). This program was also used to test for overall positive selection using a Z-test with the modified Nei-Gojobori method with Jukes-Cantor correction and a transition/transversion bias as determined by the substitution model. Significance was estimated with 10´000 bootstrap replicates. For MHC IIA, tests were performed separately for the entire allele, the α_1-domain, and the allele without α_1. For MHC IIB, positive selection was calculated for the entire allele and separately for the β_1-domain.

Site-specific positive selection was inferred by CodeML of the package PAML v4.7 (Yang 2007). Three models that do not allow for positive selection (M1a, M7, M8a) were compared to two models that allow certain sites to evolve under positive selection (M2a, M8). The analysis was run using the default parameters specified in the CodeML manual. The most likely evolutionary scenario was determined by likelihood ratio tests and positively selected sites were calculated by the Bayes empirical Bayes approach (Yang *et al.* 2005).

The programs DnaSP v5.10 (Librado & Rozas 2009) and Geneconv (Sawyer 1999) as well as the MaxChi method implemented in the stepwise package in R (Graham *et al.* 2005; R Core Team 2013) were used to detect signs of recombination and gene conversion. DnaSP and Geneconv were run with the default settings. MaxChi was run with a window half-width of 30 polymorphic sites and 1000 Monte Carlo replicates. Significant breakpoint at adjacent sites were counted as one recombination event (Graham *et al.* 2005).

For visualizing phylogenetic relationships maximum likelihood and neighbour-joining trees were constructed in MEGA v5.20 using the best-fit substitution model and the respective parameters. The branch support was tested by 500 and 1000 bootstrap replicates for ML and NJ trees, respectively. Since MHC alleles often show genealogic inconsistencies which are not well represented by trees, Neighbour-net networks have been constructed in SplitsTree4 (Huson & Bryant 2006) for the entire alleles and the α_1-domain and β_1-domain separately. Partial α_1 and β_1 sequences of the Japanese eel (Ac. no. gi|471296315, gi|471296500, gi|471296745, gi|471296952, gi|471304152, gi|471303063, gi|471303479, gi|471304253, gi|471304172, gi|471296560) have been included in the networks. Bootstrap values were obtained by 1000 replicates.

3. Results

3.1. MHC class IIA

Sequence diversity

Fourteen putative MHC IIA alleles were isolated from cDNA of four *A. anguilla* individuals, seven of which included the complete coding sequence. The alleles translate into 14 distinct amino acid sequences and none of them comprised a premature stop codon (figure 1.3). The alleles were named *MhcAnan-DXA*1-14* according to the nomenclature proposed by Klein et

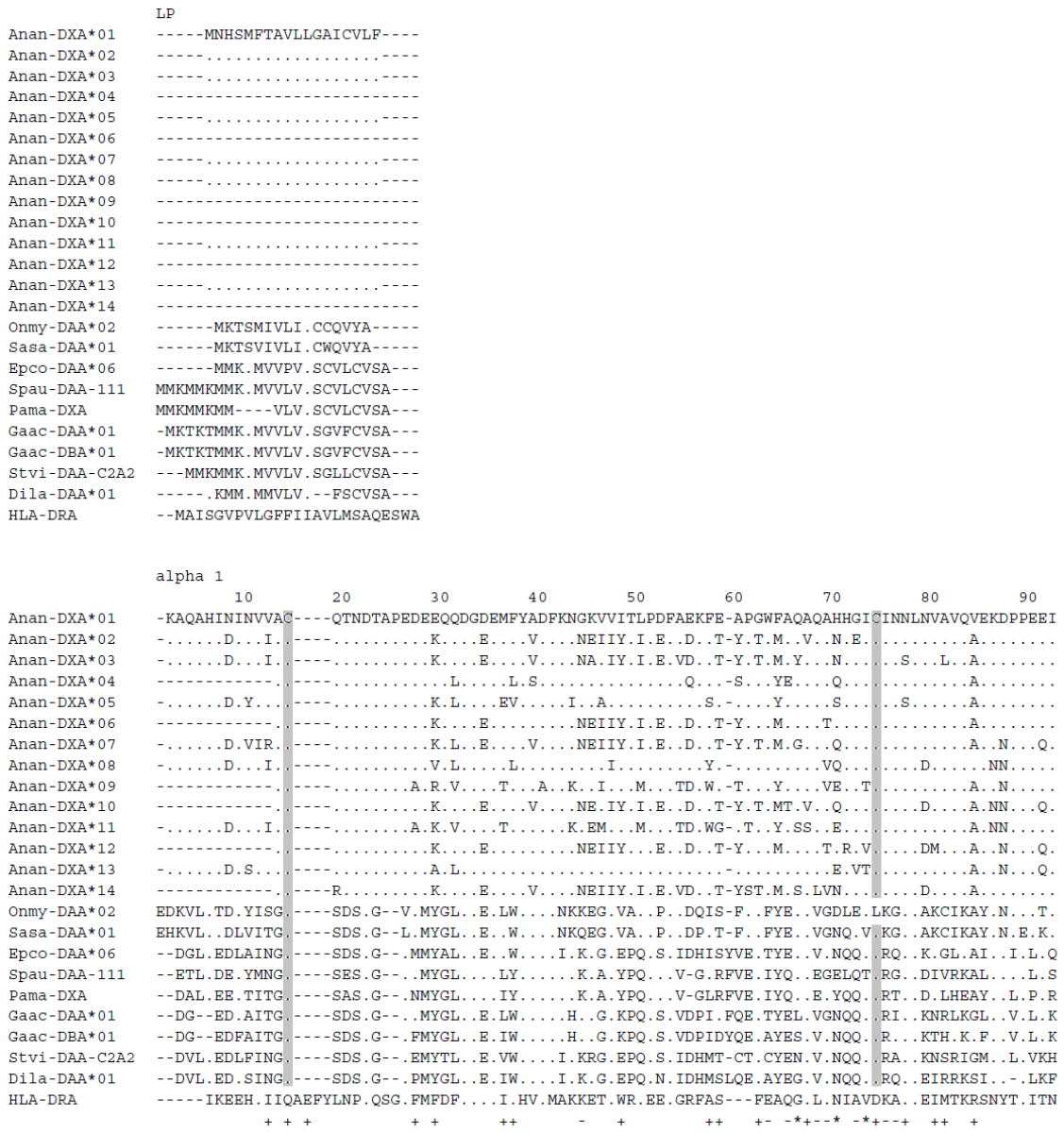


Figure 1.3 Amino acid alignment of *Anguilla anguilla* MHC IIA sequences with sequences from other teleosts and human DQA (*Onchorhynchus mykiss*, *Salmo salar*, *Epinephelus coioides*, *Spratus aurata*, *Pagrus major*, *Gasterosteus aculeatus*, *Stizostedion vitreum*, *Dicentrarchus labrax* (Ac. nos. as indicated in the material and method section, Ac. no. Dila-DAA*01: gb|DQ821106.1)). Conserved cysteine residues are shaded in light grey and conserved glycine residues of the TM domain are shaded in dark grey. ABS in humans are indicated by +, TCR binding sites by -, and residues involved in either antigen or TCR binding by *. o indicate the main residues interacting with the CD4 co-receptor in humans.

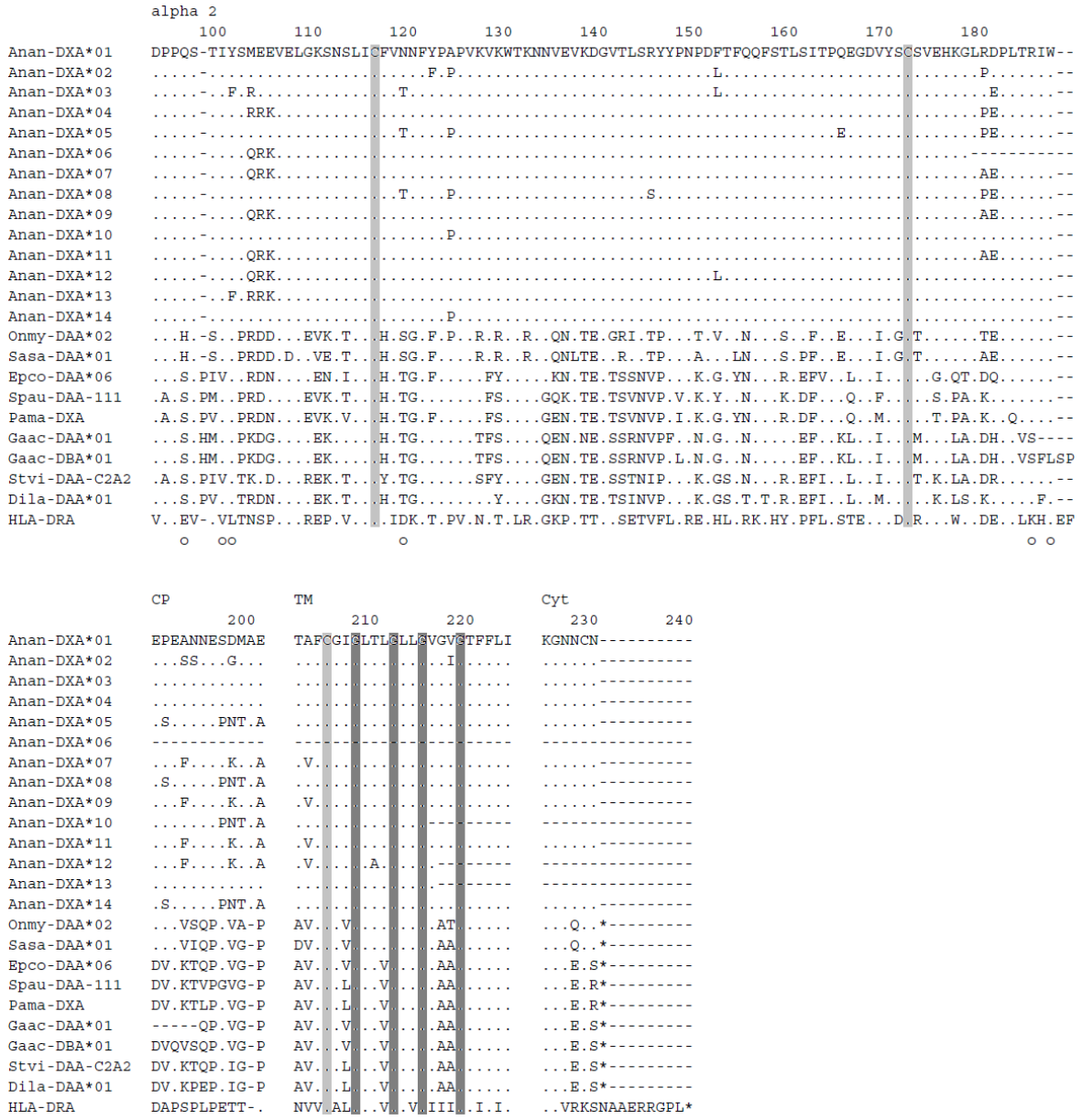


Figure 1.3 continued.

al (1990). Three to four alleles were detected per individual suggesting the presence of at least two loci.

The sequences of 717 base pairs contain the putative start codon and end right before the putative stop codon. The sequences show all the typical features of MHC IIA as inferred from other teleost sequences, a leader peptide (57 bp), an α_1-domain (255 bp), an α_2-domain (282 bp), a connecting peptide (36 bp), a transmembrane domain (69 bp), and a cytoplasmic tail (18 bp). The mature protein is 220 amino acids long and contains all four conserved cysteines (positions 13, 68, 109, 165) that are potentially forming disulphide bridges, including the teleost-specific cysteine residues in the α_1-domain (figure 1.3). Furthermore, the $C^{196}xxG^{199}xxxG^{203}xxG^{206}xxxG^{210}$ motive of the transmembrane domain is also conserved, as is the polarity of the main residues presumably involved in the interaction with the CD4

Table 1.5 Non-synonymous and synonymous substitution rates for MHC IIA and MHC IIB.

Gene	Domains	dN	dS	dN/dS	p-value
MHC IIA	LP, α1, α2, CP, TM, Cyt	0.083 (0.011)	0.053 (0.011)	1.56	0.017
	α1	0.166 (0.026)	0.060 (0.020)	2.76	< 0.001
	LP, α2, CP, TM, Cyt	0.035 (0.008)	0.056 (0.016)	0.61	1.000
MHC IIB	LP, β1, α2, CP, TM, Cyt	0.090 (0.014)	0.101 (0.018)	0.89	1.000
	β1	0.131 (0.024)	0.124 (0.027)	1.06	0.408

dN, number of nonsynonymous substitution per nonsynonymous site, dS, number of synonymous substitutions per synonymous site, standard errors are given in brackets, p-values indicate the probability of positive selection calculated using a Z-test, LP, leader peptide, CP, connecting peptide, TM, transmembrane domain, Cyt, cytoplasmic tail

co-receptor. In total, 68 variable sites were identified in the amino acid alignment, 46 (67.6 %) and 12 (17.6 %) of which are located in the α_1-domain and the α_2-domain, respectively. The leader peptide and the cytoplasmic tail are conserved among all the alleles. At the nucleotide level, 130 sites are segregating with 87 (66.9 %) and 27 (20.8 %) of the sites located in the α_1-domain and α_2-domain, respectively. The mean nucleotide distance (±SE) among alleles is 0.136 (±0.022).

Selection

An excess of non-synonymous substitutions over synonymous substitutions is indicative of positive selection. A Z-test revealed overall positive selection for the entire alleles and separately for the α_1-domain (table 1.5). However, the signal is completely absent when excluding the α_1-domain from the analysis. Models allowing for positively selected sites (PSS) fit the data considerably better than models not allowing for positive selection (table 1.6).

Table 1.6 Summary of site-specific likelihood ratio tests for positive selection in MHC IIA.

Gene	LRT	2Δl	p-value	Estimate for ω >1	Proportion of PSS	PSS
MHC IIA	M1a vs M2a	81.766	<0.0001	6.71254	0. 23807	***42***, **44**, *49*, *50*, **59**, 63, ***70***, **72**, **76**, ***78***, *79*, **81**, ***82***, 91, **99**, 103, **114**, **115**, 163
	M7 vs M8	81.818	<0.0001	7.65097	0.20136	***42***, **44**, *49*, *50*, 58, **59**, **63**, 67, ***70***, *71*, **72**, 73, **76**, ***78***, ***79***, **81**, ***82***, *85*, 91, **99**, **103**, **114**, **115**, 163
	M8a vs M8	81.761	<0.0001			
MHC IIB	M1a vs M2a	68.868	1.11E-15	8.21475	0.07553	***33***, ***43***, **58**, ***66***, 68, **71**, 72, *90*
	M7 vs M8	70.176	5.55E-16	8.15487	0.07569	***33***, ***43***, **58**, ***66***, **68**, **71**, **72**, 77, ***90***
	M8a vs M8	66.925	3.33E-16			

LRT, models compared by the likelihood ratio test, $2\Delta\ell = 2(\ell b\text{-}\ell a)$, ω = dN/dS, PSS, positively selected sites inferred by Bayes empirical Bayes posterior probabilities, PSS in **bold** are inferred at 99 % level, PSS in *italics* correspond to human ABS.

Seven out of 16 PSS and 9 out of 21 PSS in the α_1-domain inferred by models M2a and M8, respectively, correspond to the ABS in human HLA as indicated in Reche & Reinherz (2003).

Recombination

Geneconv detected eight events of recombination or gene conversion that involve large fragments with proposed breakpoints at the beginning or the end of the domains. Since the analysis is performed on cDNA, the true breakpoints are most likely located in the introns. DnaSP and MaxChi estimated a minimum number of 21 and 16 recombination events, respectively. In contrast to the exchange of large fragments detected by Geneconv, these events represent mainly micro-recombinations that involve only a few base pairs.

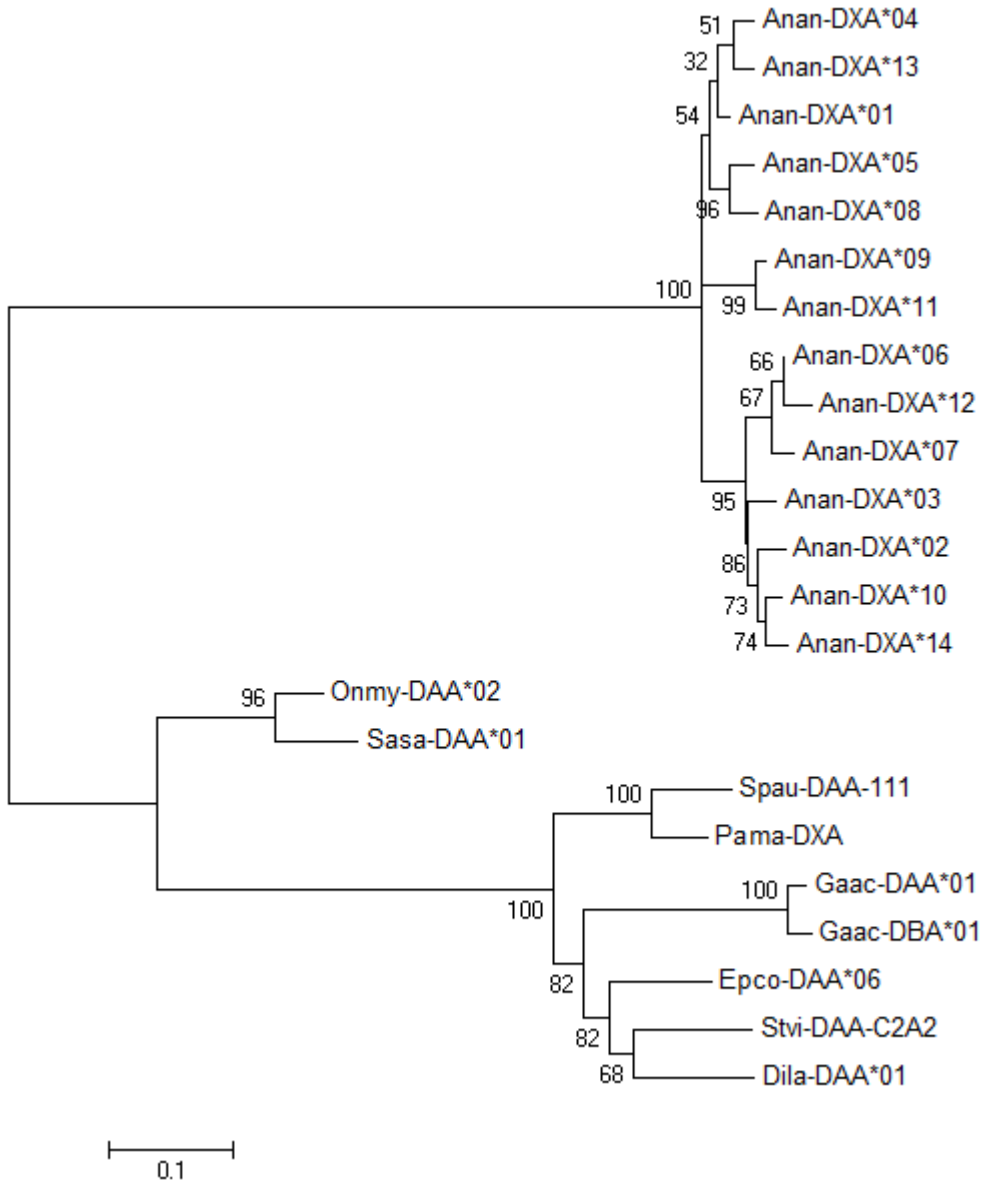
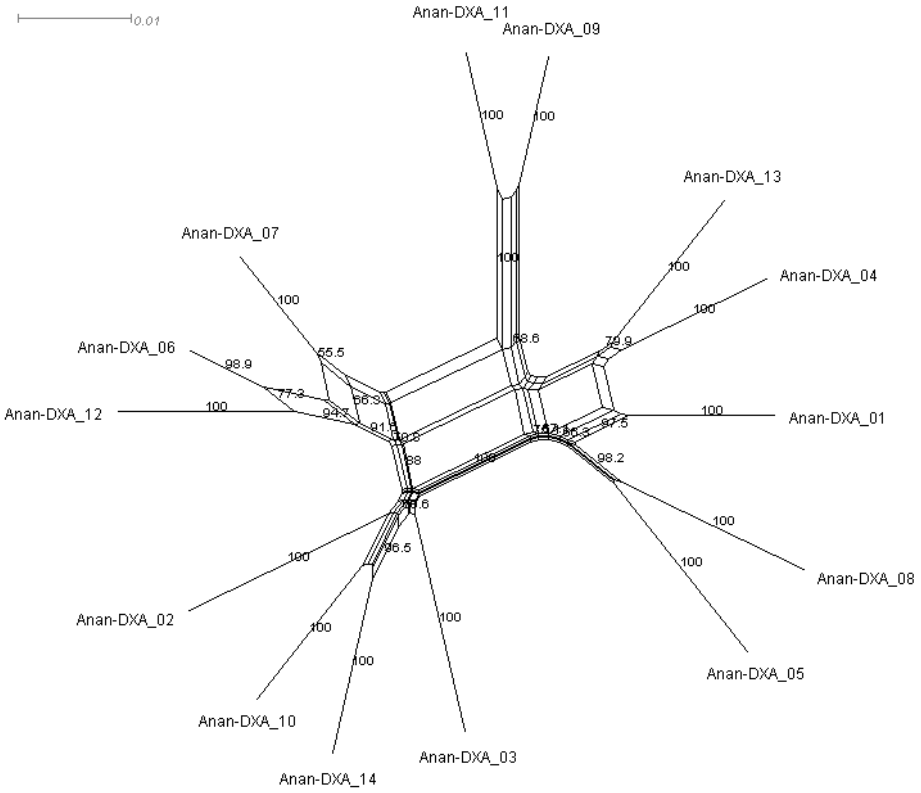
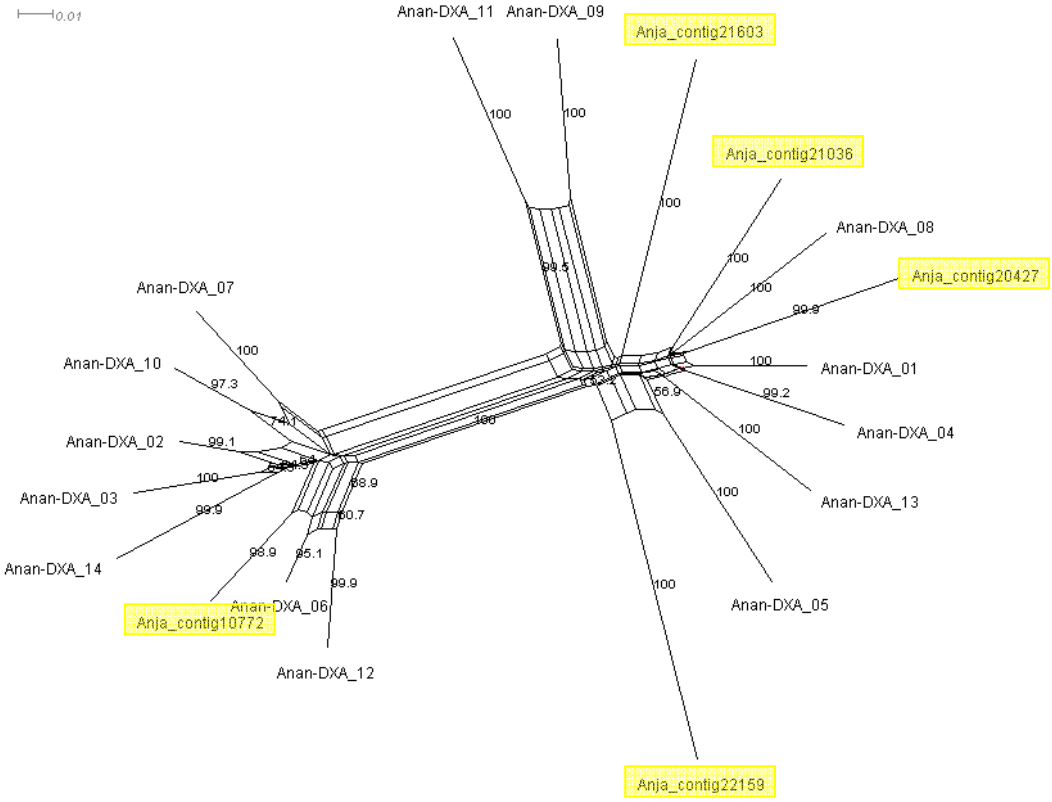


Figure 1.4 Maximum likelihood tree of complete MHC IIA alleles showing the relationship between *A. anguilla* alleles and other teleosts. Bootstrap values were obtained with 500 replicates. *A. anguilla* alleles fall into two, maybe three major clusters. A neighbour-joining tree (not shown) gave similar results.



a



b

Figure 1.5 Neighbour-net networks for MHC IIA of (a) complete *A. anguilla* alleles and (b) α_1-domains of A*. anguilla* and *A. japonica*. Only bootstrap values above 50 are shown. *A. japonica* alleles are highlighted.

Phylogenetic analysis

The phylogenetic tree supports the clear separation between the *A. anguilla* and the euteleostean sequences that is indicated in the alignment (figure 1.4). Within the eel sequences, there are two, maybe three major clusters of alleles, indicating at least one

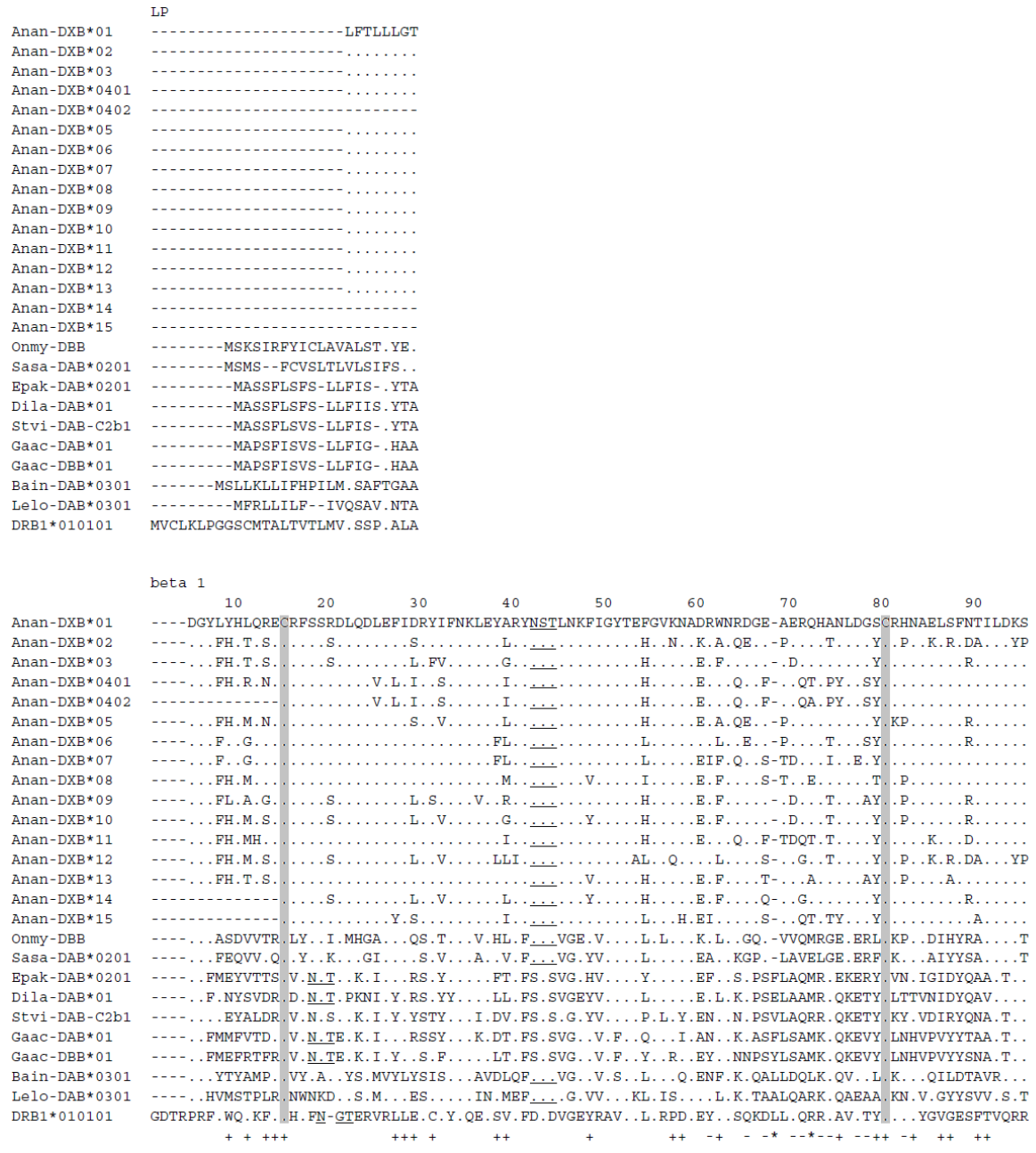


Figure 1.6 Amino acid alignment of MHC IIB sequences of *A. anguilla*, additional teleost species, and HLA-DRB1 (Barbus intermedius and Leiocassis longirostris *Onchorhynchus mykiss*, *Salmo salar*, *Epinephelus akaara, Dicentrarchus labrax*, *Stizostedion vitreum*, *Gasterosteus aculeatus, Barbus intermedius, Leiocassis longirostris* (Ac. nos. as indicated in the material and method section, Ac. no. Dila-DAB*01: gb|DQ821110.1)). Conserved cysteines in the β_1- and β_2-domain are shaded in light grey and conserved glycines in the TM are shaded in dark grey. N-linked glycosylation sites are underlined. ABS in humans are indicated by +, TCR contact sites by -, and sites that can interact with either the antigen or the TCR are denoted with *. Main sites contacting the CD4 co-receptor are indicated by o.

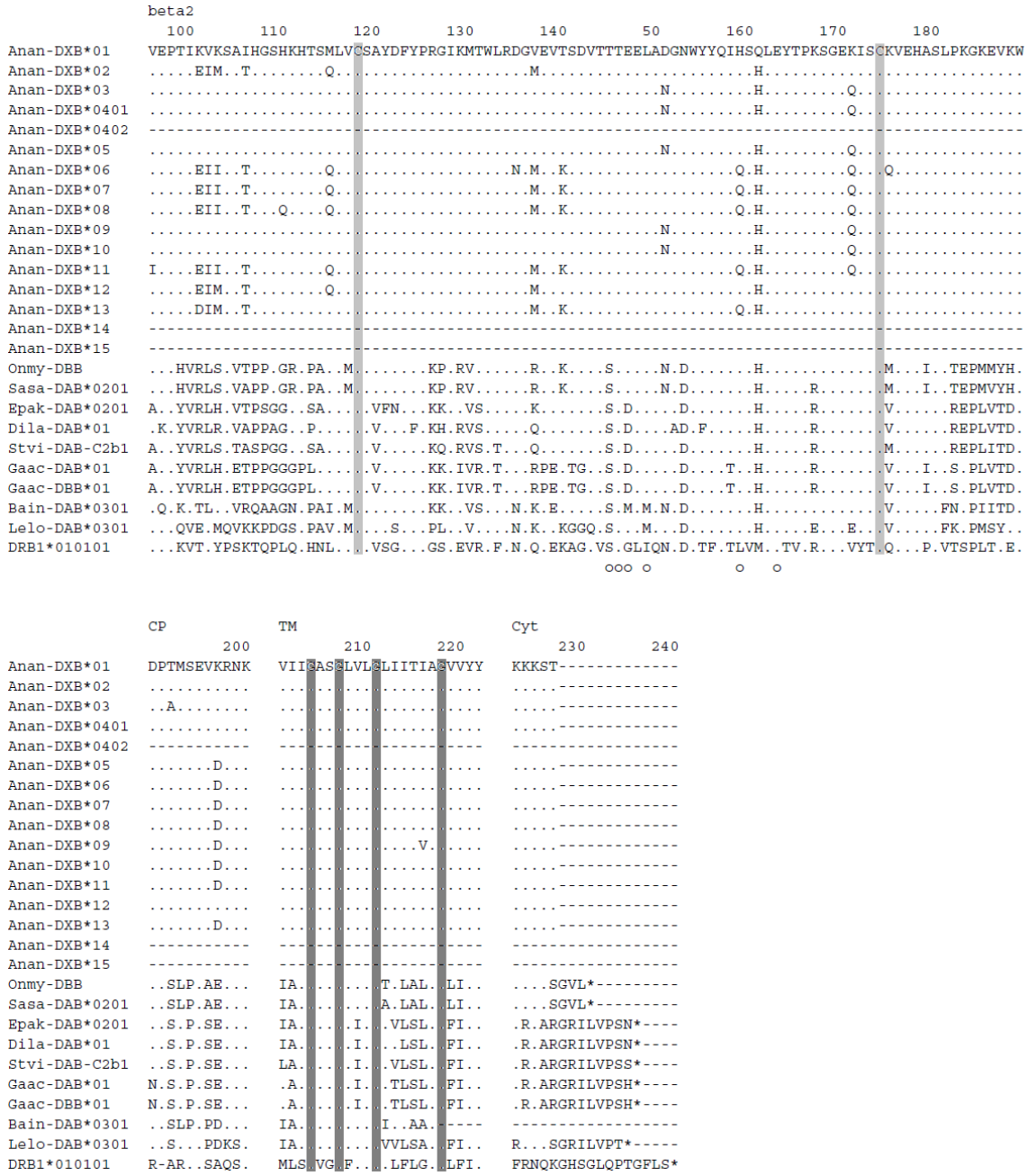


Figure 1.6 continued.

duplication event. Within these clusters bootstrap support is relatively low. However, the Neighbour-net network indicates an additional subcluster (figure 1.5a). For the α_1-domain the separation of the two major clusters is much more pronounced, whereas the subcluster including alleles *Anan-DXA*09* and *Anan-DXA*11* is less resolved and the potential subcluster including alleles *Anan-DXA*06*, *Anan-DXA*07*, and *Anan-DXA*12* disappears (figure 1.5b). Additionally, the sequences of *A. japonica* are intermingled with the *A. anguilla* sequences, indicating trans-species polymorphism.

3.2. MHC class IIB

Sequence diversity

Putative MHC IIB alleles were isolated from genomic DNA and cDNA of four A. anguilla individuals. The fragments contained parts of exon 1, the complete exon 2 and exon 3, and parts of exon 4. For a subset, intron 2 and intron 3 sequences were isolated. Each individual possesses three to six alleles in the genome, however, a maximum of four alleles were found in cDNA. In total, 16 different nucleotide sequences were found that correspond to 16 amino

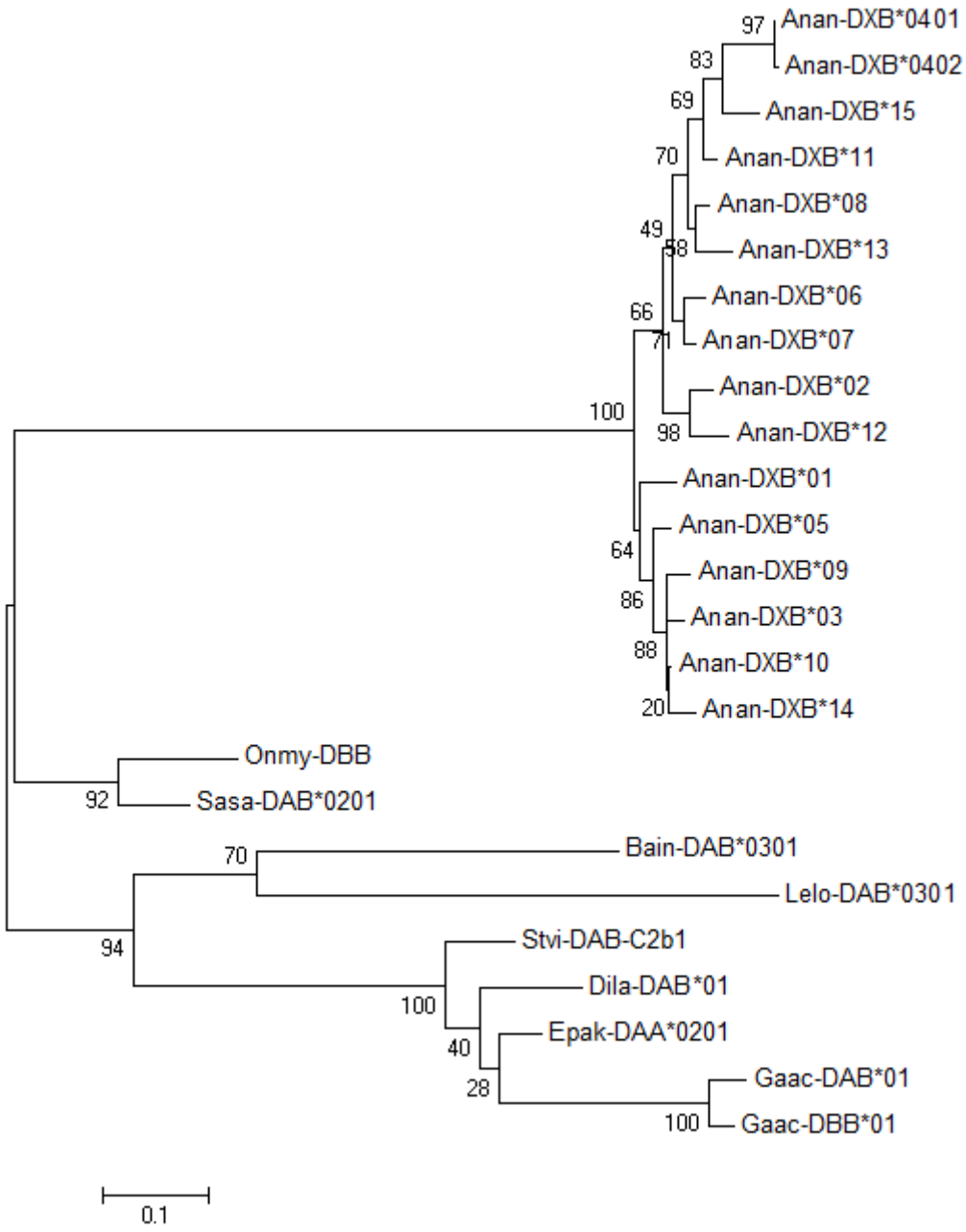
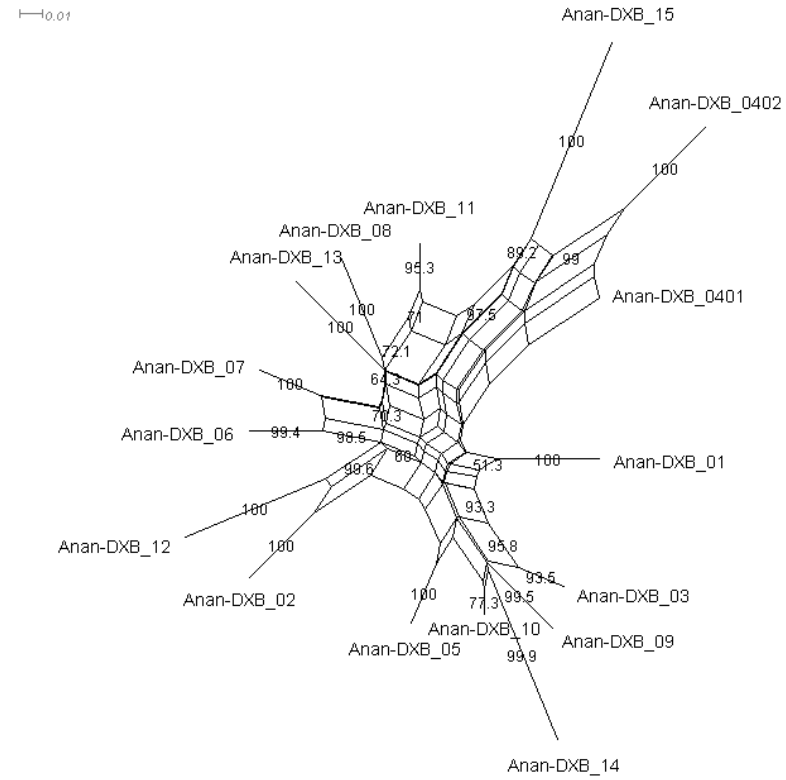
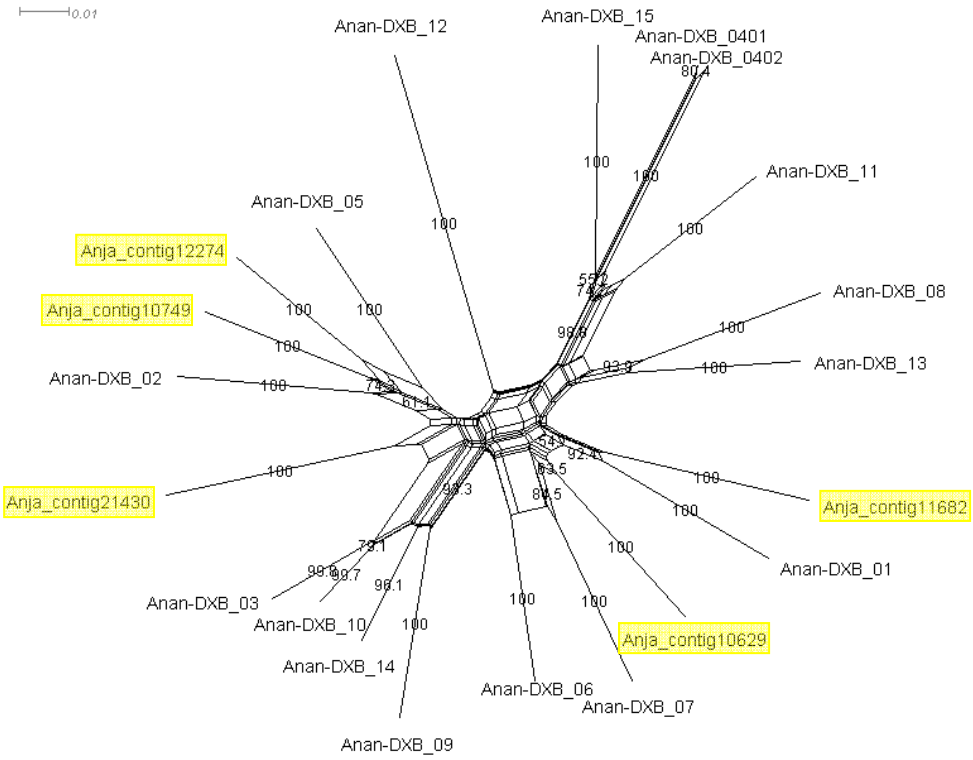


Figure 1.7 Maximum likelihood tree showing the relationship among *A. anguilla* MHC IIB alleles and those of euteleosts. Bootstrap values were obtained by replicating 500 times. Anan-DXB alleles do not form clear clusters. Similar results were obtained for a neighbour-joining tree (data not shown).



a



b

Figure 1.8 Neighbour-net networks for a) entire alleles of *A. anguilla* MHC IIB and b) β1-domains of *A. anguilla* and *A. japonica*. Bootstrap values above 50 are shown. *A. japonica* alleles are highlighted.

acid sequences (figure 1.6). Thirteen of them were present in both genomic and cDNA. No premature stop codon was found in any of the sequences. The alleles were named *MhcAnan-DXB*01-15* following the proposed nomenclature (Klein *et al.* 1990).
The coding region is 693 nucleotides long and translates into 230 amino acids, 222 of which contribute to the mature protein. The length of exons 2 and 3 is 270 bp and 282 bp, respectively. For intron 2, two lineages could be isolated which differ considerably in their nucleotide composition. However, 87 nucleotides downstream of exon 2 and 18 nucleotides upstream of exon 3 are relatively conserved. For one lineage several indels could be identified resulting in length differences (894-916 nucleotides). The second lineage is 669 nucleotides long. Since only one sequence could be isolated, no information about variability is available. In contrast to intron 2, intron 3 is short (130 bp) and more conserved belonging to only one lineage. Both, intron 2 and intron 3 contain the consensus GT-AG splice sites.
In the coding region, the typical features of MHC IIB proteins could be identified: a partial leader peptide (26 bp), a complete β_1-domain and β_2-domain (270 bp and 282 bp), a connecting peptide (33 bp), a transmembrane domain (66 bp), and a partial cytoplasmic tail (16 bp). Furthermore, the conserved cysteine residues at positions 13, 75, 113, and 169 are present (figure 1.6) as is the characteristic motive of the transmembrane domain (G199xxG202xxxG206xxxxxxG213). Residues corresponding to the main interaction sites with the CD4 co-receptor in human HLA show identical or conservative substitutions. Overall, 61 out of 230 amino acid sites are segregating with 43 (70.5 %) and 15 (24.6 %) located in the β_1-domain and β_2-domain, respectively. In the nucleotide alignment, 128 bp are variable out of which 83 (69.5 %) and 32 (25.0 %) occur in the β_1-domain and β_2-domain, respectively. The mean nucleotide distance (±SE) among alleles is 0.102 (±0.011)

Selection

No signature of overall positive selection was detected neither for the entire alleles nor the β_1-domain (table 1.5). Additionally, the rate of synonymous substitutions does not differ among domains, however, the rate of non-synonymous substitutions is considerably lower outside the β_1-domain. Despite the lack of overall positive selection, models that allow certain sites to be positively selected (PSS) explain the observed pattern significantly better than models that do not include positive selection (table 1.6). For models M2a and M8, 4 of the 8 and 9 PSS, respectively, located in the β_1-domain correspond to ABS in human HLA (Reche & Reinherz 2003).

Recombination

A minimum of 13 and 7 recombination events were calculated by DnaSP and MaxChi, respectively. Geneconv detected 10 recombination or gene conversion breakpoints. The exchanged fragments have an average length of 68 bp and are mainly located in the β_1-domain.

Phylogenetic analysis

The phylogenetic separation of eel and euteleostean MHC IIB sequences is supported by large bootstrap values (figure 1.7). Within the *A. anguilla* sequences three clusters are

supported by moderate bootstrap values, however, a network does not indicate any clear clustering (figure 1.8a). The absence of such clusters is even more pronounced in the β_1-domain (figure 1.8b). Furthermore, the sequences of *A. anguilla* and *A. japonica* form one cluster, indicating trans-species polymorphism among anguillids.

4. Discussion

Sequence diversity

In the European eel, three to four MHC class IIA alleles per individual were isolated from cDNA indicating the presence of at least two loci. For MHC IIB, up to six alleles were found per individual, however, not more than four were detected in cDNA. The alleles not isolated from cDNA might be pseudogenes or have acquired different functions. MHC genes are assumed to evolve under a birth-and-death scenario and gene duplications and loss of functionality are commonly observed (Klein *et al.* 1998b; Nei & Rooney 2005). However, it cannot be excluded these alleles are expressed in very low numbers or in tissues other than the liver. Similar numbers of alleles and potential loci for MHC IIA and MHC IIB have been found for cyprinids and certain bass species (van Erp *et al.* 1996; Kruiswijk *et al.* 2004).

The MHC class IIA and class IIB alleles contained all the typical features of functional proteins (figures 1.3 and 1.6). The length of the different domains is in the range of that described for other teleost species (Stet *et al.* 2002; Kruiswijk *et al.* 2004; Reusch *et al.* 2004; Silva *et al.* 2007). However, the connecting peptide of the *Anan-DXA* alleles is one residue longer than in most described teleostean MHC IIA alleles. The cysteine residues needed for the correct structure of the molecule are located at the same positions as in all other fish sequences. Similar to catfish, none of the alleles of the European eel possesses an N-glycosylation site in the α-chain which is contrary to many other fish species for which such a site is present in some or all of the alleles (Godwin *et al.* 2000; Grimholt *et al.* 2000; Cuesta *et al.* 2006). N-glycosylation sites are important for the correct folding and assembly of MHC heterodimers. However, Godwin et al (2000) speculated that the presence of an N-glycosylation site in the β-chain might be sufficient for assembling functional MHC molecules. The typical motives of the MHC IIA and MHC IIB transmembrane domains are conserved among A. anguilla sequences and alleles of other teleosts. It is important for the formation of α/β-heterodimers and is conserved among all studied vertebrate taxa (King & Dixon 2010).

Polymorphism in *Anan-DXA* alleles is high in comparison with other teleosts. While 14 different alleles have been isolated from four individuals in the European eel, only eight have been found in six zebrafish and four in three Atlantic salmon individuals (Sultmann *et al.* 1993; Grimholt *et al.* 2000). Furthermore, over 50 % of the residues in the α_1-domain of the *Anan-DXA* alleles are variable. For MHC IIB, allelic variation in the European eel is in the upper range of what has been described for teleosts. For silvery minnow, 25 alleles were isolated from 9 individuals and in seahorses nine alleles were detected in ten individuals (Osborne & Turner 2011; Bahr & Wilson 2012), whereas 16 alleles were present in four individuals of the European eel. The sequence diversity at MHC IIB is in a similar range as in other teleosts (Stet *et al.* 2002; Reusch *et al.* 2004).

In contrast to most vertebrates which usually show lower sequence polymorphism at MHC IIA than MHC IIB genes (Stet *et al.* 2002; Reche & Reinherz 2003; Reusch *et al.* 2004; Wegner 2008; Sin *et al.* 2012), MHC IIA in the European eel is as polymorphic as MHC IIB. A similar pattern has only been found for the DQA genes in mammals which show similar levels of variability as MHC IIB genes.

Mode of evolution

Evolution of the MHC class II genes of the European eel seems to be driven by a combination of positive selection and recombination. The MHC IIA of the European eel experienced strong positive selection, particularly in the α_1-domain which is involved in antigen binding. In this domain, d_N values are increased almost threefold over d_S values. Furthermore, roughly 43 % of the positively selected sites (PSS) inferred by site-specific models correspond to antigen-binding sites (ABS) in humans. In contrast to most vertebrate species, no overall signal of positive selection was detected for MHC IIB, neither for the entire alleles nor the β_1-domain (Shum *et al.* 2001; Wegner 2008; Sin *et al.* 2012). However, some salmonid species also lack a clear signal of positive selection at MHC IIB (Aguilar & Garza 2007). Despite the absence of overall positive selection, several sites inferred to be involved in antigen binding seem to be subjected to strong positive selection. Strong purifying selection or neutral processes at sites outside the ABS might be the reason that no overall signal for positive selection could be detected.

The synonymous substitution rates in MHC IIA and MHC IIB of the European eel are highly elevated in comparison to euteleosten genes. The rates in MHC IIA exceed those in the stickleback fourfold, whereas for MHC IIB, the rates are up to ten times higher than in other fish species (Reusch *et al.* 2004; Aguilar & Garza 2007; Bahr & Wilson 2012).

Phylogenetic analysis

The phylogeny of the MHC II genes mirrors the relationship inferred from mitochondrial DNA (Inoue *et al.* 2001; Near *et al.* 2012), with a clear separation of *A. anguilla* from euteleosts. Alleles from *A. anguilla* and *A. japonica*, however, cannot be assigned to species-specific clusters, indicating that old allelic lineages persisted through the first radiation event in anguillids proposed to have happened about 20 Mya when the two species diverged from the common ancestor (Minegishi *et al.* 2005; Teng *et al.* 2009). The occurrence of trans-species polymorphism is a common phenomenon in MHC genes and has been observed in all vertebrate taxa (Glaberman & Caccone 2008; Kikkawa *et al.* 2009; Kiemnec-Tyburczy *et al.* 2010; Lenz *et al.* 2013a).

The characterization provided here will hopefully facilitate further studies of MHC of the European eel. Eels might be particularly interesting systems, because they allow studying the evolution of MHC solely influenced by parasites without strong neutral or demographic processes. Additionally, this characterization might enable studying the role of MHC genes on the species´ survival in the presence of an invasive parasite.

Chapter 2: Population structure of MHC IIB genes

1. Introduction

The major histocompatibility complex (MHC) triggers the adaptive immune response in vertebrates by presenting antigens to helper T cells. Its genes show an exceptionally high polymorphism within and among populations which is generally explained by balancing selection (Apanius *et al.* 1997; Bernatchez & Landry 2003; Eizaguirre & Lenz 2010). Three hypotheses of parasite-mediated selection gained most attention.

First, frequency-dependent selection or rare allele advantage assumes an advantage for individuals carrying rare resistance alleles, because the most common parasite will be adapted to efficiently exploit the most common host genotype (van Valen 1973). This co-evolution between host and parasites will lead to cyclic fluctuations of MHC alleles. Parasites indeed seem to adapt to familiar host genotypes and specific genotype-genotype interactions were found in wild and experimental populations (Dionne *et al.* 2009; Eizaguirre *et al.* 2012b; Kubinak *et al.* 2012)

Second, heterozygote advantage assumes that heterozygosity confers resistance to a wider set of parasites than homozygosity (Doherty & Zinkernagel 1975). The advantage can be a result of dominance or overdominance (Penn *et al.* 2002; Oliver *et al.* 2009). Additionally, divergent alleles are proposed to bind a wider and more dissimilar set of antigens and, therefore, increase resistance against diverse parasites (Wakeland *et al.* 1990; Lenz *et al.* 2009; Lenz 2011; Eizaguirre *et al.* 2012a).

Third, spatial and temporal variation in parasite communities might exert different selective pressures over space and time. This would maintain a large repertoire of alleles within a metapopulation and ultimately lead to differentiated and locally adapted populations (Hedrick 2002). Locally differentiated MHC pools could be shown to be correlated with differences in parasite communities by several studies (Dionne *et al.* 2007; Loiseau *et al.* 2009; Eizaguirre *et al.* 2012a). On the other hand, Fraser and Neff (2010) found similar MHC pools between distant guppy populations that might stem from similar selection pressures by the same parasite. However, susceptibility and resistance alleles to the same parasite might also be population-specific (Loiseau *et al.* 2011). Differences in MHC pools generally seem to be unlinked to differentiation at neutral markers and geographic distances in many species with some showing signs of divergent selection and others signs of homogenizing selection (Landry & Bernatchez 2001; Bernatchez & Landry 2003; Ekblom *et al.* 2007; Evans *et al.* 2010; Cammen *et al.* 2011). Differences in habitat type with putatively different parasite communities are commonly used to explain the observed pattern. However, differentiation at MHC is usually inferred by comparing it to differentiation at microsatellites or mitochondrial DNA and strong neutral processes such as genetic drift might limit the extent to which selection can be detected (Wegner 2008). At small geographic scales where gene flow among populations is relatively frequent, selection is regularly observed at MHC. In contrast, at a larger geographic scale restricted gene flow and genetic drift are expected to be much stronger and mask a potential signal of selection leaving a signature that is

consistent with neutral evolution (Landry & Bernatchez 2001; Aguilar & Garza 2006; Evans *et al.* 2010; Spurgin & Richardson 2010).

The European eel (*Anguilla anguilla*) is an excellent model to study the effects of natural selection due to its low genetic differentiation at neutral markers across its distribution range (Wirth & Bernatchez 2001; Dannewitz *et al.* 2005; Als *et al.* 2011).

The European eel inhabits freshwater systems and estuaries of the European continent, including Iceland, the Baltic and the Black Sea, as well as North Africa, and the Azores (Tesch 2003). Its catadromous life cycle involves two very distinct migrations across the North Atlantic (figure 2.1). Adult silver eels migrate to the Sargasso Sea where they presumably spawn once and then die. After hatching, the leptocephali larvae drift back to the European and African continent with the oceanic currents, where they metamorphose into glass eels and enter coastal waters to migrate upstream. Thereafter, they spend several years as yellow eels feeding in continental waters before silvering and migrating back to the Sargasso Sea (Schmidt 1923; Tesch 2003; Aarestrup *et al.* 2009).

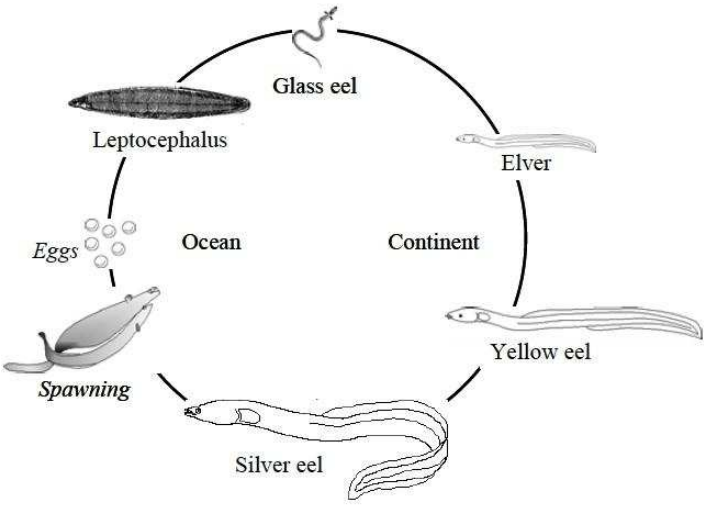


Figure 2.1 Life cycle of the European eel modified from ICES 2012.

Populations of the European eel were considered to be part of one undifferentiated panmictic population until several studies found weak but significant differentiation which is consistent with an isolation-by-distance (IBD) scenario (Daemen *et al.* 2001; Wirth & Bernatchez 2001; Maes & Volckaert 2002). Three geographically separated clusters were evident, a Mediterranean, an Atlantic and a North Sea/Baltic Sea clade. Studies controlling for age differences detected temporal variation between cohorts and arrival waves indicating an isolation-by-time scenario (Dannewitz *et al.* 2005; Maes *et al.* 2006). However, more recent studies on glass eels and adult eels could not support earlier IBD or IBT pattern, but rather found no genetic differences on a large scale and genetic patchiness on a smaller scale (Palm *et al.* 2009; Pujolar *et al.* 2009; Pujolar *et al.* 2011a). Furthermore, larvae caught in the Sargasso Sea showed no geographic or temporal genetic differentiation at nuclear

markers (Als *et al.* 2011). These later findings indicate a panmictic population with a relatively small number of successfully reproducing adults during each spawning event followed by random larval dispersal.

In the early 1980s, the swim bladder nematode *Anguillicola* crassus was introduced into Europe from East Asia and spread rapidly across the entire distribution range of the European eel (Kirk 2003). The introduction coincides with the decline of eel recruitment to less than 1 % of the pre-1980 levels (Dekker 2003; ICES 2012). In contrast to the native host of *A. crassus*, the Japanese eel *A. japonica*, the European eel is highly susceptible to infections (Lefebvre & Crivelli 2004). The prevalence is about 60 – 70 % and the infectivity in *A. anguilla* is greatly enhanced compared to *A. japonica* (Weclawski *et al.* 2013). Therefore, *A. crassus* infections are considered to be among the major causes for the collapse (Kirk 2003).

Infected individuals show a greatly reduced lumen and thickened wall of the swim bladder which can lead to the complete damage of the organ. Furthermore, they show a reduced swimming ability and lower resistance to hypoxia (Gollock *et al.* 2005; Palstra *et al.* 2007) which might prevent the completion of the spawning migration. Additionally, mass mortalities are associated with *A. crassus* infections (Kirk 2003). However, there seems to be little or no impacts on the continental stages of eels in the absence of additional stressors (Lefebvre *et al.* 2013).

Despite the supposedly strong impacts of *A. crassus* on its life history, the European eel is incapable of mounting an effective immune response (Knopf *et al.* 2000). Antibodies were only detected at a late stage of infection against adult nematodes. Furthermore, the response was highly variable among individuals.

The variability in the immune response and the stabilized prevalence might be induced by differences in MHC alleles among individuals and eventually lead to purging of susceptibility alleles from the population. As a consequence, the allele pool might be reduced and homogenized among subpopulations.

2. Material & Methods

2.1. Samples

Glass eels were collected in December and January in Adour, France for three consecutive cohorts in 2010, 2011, and 2012. Twenty-five to thirty individuals were sampled for each cohort. An additional 14 and 15 glass eel samples were collected in Burrishoole, Ireland and Viskan, Sweden, respectively (table 2.1). Adult eels were collected from 13 populations in Europe. Sampling was designed to screen both large and small geographical scales (table 2.1). Additionally, 15 American eels (*A. rostrata)* from the river Schwentine, Germany and 5 from the East coast of North America were included.

2.2. Nucleic acid extraction and cDNA synthesis

DNA was extracted from with the DNeasy® 96 Blood and Tissue kit (Qiagen, Hilden, Germany) following the manufacturer´s recommendations but eluting twice in 50 µl of AE buffer. Concentrations were measured on a NanoDrop 1000 and samples were standardized to a concentration of 10 ng/µl.

Table 2.1 Populations and sample characteristics.

Population		N	Life stage	Year	Species	Salinity	Infection
Adour, France	AD	25	Glass eels	2010	*A. anguilla*	Freshwater	
		24	Glass eels	2011	*A. anguilla*	Freshwater	
		30	Glass eels	2012	*A. anguilla*	Freshwater	
America	AM	5	Adult eels		*A. rostrata*		
Bann Lower, Ireland	BL	11	Adult eels		*A. anguilla*	Marine	
Bann Toome, Ireland	BT	16	Adult eels		*A. anguilla*	Freshwater	
Boretree, Ireland	SLB	15	Adult eels		*A. anguilla*	Marine	
Burrishoole, Ireland	BU	14	Glass eels		*A. anguilla*		
		17	Adult eels		*A. anguilla*		
Denmark	DK	17	Adult eels		*A. anguilla*	Brackish	
Finland	Fin	18	Adult eels		*A. anguilla*	Freshwater	
Glynn Lagoon, Ireland	GL	19	Adult eels		*A. anguilla*	Brackish	
Larne Lagoon, Ireland	LC	16	Adult eels		*A. anguilla*	Brackish	
Larne Lough, Ireland	LL	12	Adult eels		*A. anguilla*	Marine	
Lough Comber, Ireland	SLC	11	Adult eels		*A. anguilla*	Marine	
Portugal	Pt	15	Adult eels		*A. anguilla*		
Quoile, Ireland	Q	11	Adult eels		*A. anguilla*	Brackish	
Schwentine, Germany	Ger	12	Adult eels		*A. anguilla*	Freshwater	
		24	Adult eels		*A. anguilla*	Freshwater	infected
		23	Adult eels		*A. anguilla*	Freshwater	uninfected
		13	Adult eels		*A. rostrata*	Freshwater	
Viskan, Sweden	Swe	15	Glass eels		*A. anguilla*		

N, sample size, Infection, infection with the swim bladder nematode *Anguillicola crassus*

2.3. MHC genotyping

First, a PCR was run with the forward primer AaMHCIIBE2F3 (5´-AGTGYCGTTTCAGYTCCAGMGAYCTG-3´) and the reverse primer AaMHCIIBE2R2 (5´-CTCACYTGRMTWATCCAGTATGG-3´, both primers have been isolated within chapter I) in a final volume of 10 µl containing 1× DreamTaq buffer (Thermo Scientific, St. Leon-Rot, Germany) including 20 mM MgCl2, 0.5 mM of each dNTP, 0.5 µM of each primer, 0.5 U DreamTaq polymerase, and 10 ng of template DNA. The thermal profile contained and initial step of 95 °C for 1 min, 10 cycles of 95 °C for 15 s, 59 °C for 15 s, and 72 °C for 30 s, and a final extention step of 72 °C for 7 min. Two separate reactions were run per individual. The PCR products were then diluted 1:5 in HPLC and used as template in the second PCR with the same concentrations as the initial PCR but in a volume of 25 µl. This reconditioning step reduces the accumulation of PCR artefacts in the final product (Lenz & Becker 2008). Additionally, modified forward and reverse primers were used which consisted of an adaptor for sequencing on a 454 platform (F: CCATCTCATCCCTGCGTGTCTCGACTCAG, R: CCTATCCCCTGTGTGCCTTGGCAGTCTCAG), followed by a 10 bp tag (MID) and the sequence of the MHC specific primers AaMHCIIBE2F3 and MHCIIBE2R2 for the forward and reverse primers, respectively. Two unique combinations of tagged primers were used for each individual. The thermal profile was the same as for the initial PCR but with an annealing temperature of 57 °C. Two microlitres of each PCR were run on a 1.5 % agarose gel for 50 min at 80 V. Then, the products were cleaned using the MinElute® 96 UF PCR purification kit (Qiagen, Hilden, Germany) following the manufacturer´s instructions. PCR products were eluted in 20 µl HPLC water by pipeting up and down 25 times. Concentrations were measured on a NanoDrop 1000 and 30 ng per sample were pooled, but not more than 6 µl. Six times 40 µl of this pool were run on a 1 % agarose gel for 3 h at 45 V and stained in ethidium bromide. The bands were excised and purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany) following the modifications outlined in part 1. The six bands were pooled on one column for purification. The concentration was measured on a NanoDrop 1000 and the sample was sequenced on a Roche 454 sequencing platform (LGC Genomics, Berlin, Germany).

Custom Perl scripts were then used to call the alleles and assign them to individuals. Only reads containing both forward and reverse primers were kept and assigned to individuals based on the MID tags. For both, primers and MID tags the occurrence of one nucleotide mismatch or insertion/deletion due to sequencing errors was allowed. Then, MID tags and primer sequences were removed. Putative alleles of each individual were kept in the final dataset if they (1) were present in both PCR replicates and (2) occurred at a frequency of at least 10 % of the most common allele within a single PCR reaction.

2.4. Data analysis

The sequences were aligned in BioEdit (Hall 1999). The best-fit nucleotide model was calculated using MEGA v5.20 (Tamura *et al.* 2011). The Z-test implemented in MEG was used estimate overall positive selection using the modified Nei-Gojobori method with Jukes-

Cantor correction and the transition/transversion ratio indicated by the best-fit nucleotide model. Significance was determined by 10´000 bootstrap replicates. Positively selected sites were assessed in CodeML implemented in the program PAML (Yang 2007). The models M1a, M7, and M8a, not allowing for positive selection were compared to the models M2a and M8 that allow for positively selected sites. A likelihood ratio test was used to determine which evolutionary scenario explains the data. Sites under positive selection were calculated with the Bayes empirical Bayes approach (Yang *et al.* 2005). Positive selection was assessed for the entire dataset and separately for adult eels and glass eels, as well as for the two life stages of the Burrishoole population. Running models 7, 8, and 8a was not successful for the largest dataset including all individuals.

In order to estimate differences in MHC composition among populations, life stages, salinities (freshwater, brackish, and marine), and between the two species, *A. anguilla* and *A. rostrata* an analysis of similarity implemented in the software Primer 6 was performed (Clarke 1993). Additionally, genetic distances of MHC pools between infected and uninfected individuals from the river Schwentine were calculated. Correlations between distances of MHC and microsatellite data (M. Baltazar-Soares, unpublished) among populations were assessed with a Mantel test. Furthermore, differences in the mean numbers of alleles per individual among populations, life stages, and the two species, as well as their interaction were calculated with a linear model for the entire dataset and for the Burrishoole population separately. The Mantel test and the linear model were run in the software R (R Core Team 2013).

3. Results

A total of 202 alleles were found in 363 individuals including glass eels and adult eels of *A. anguilla* as well as adult eels of *A. rostrata* all of which translated into unique amino acid sequences. The sequences comprise a fragment of MHC IIB exon 2 and vary in length from 198 to 202 base pairs. There is no species-specific clustering for the alleles isolated from the two species (figure 2.2). Furthermore, *A. rostrata* shares 21 out of the 39 alleles with *A. anguilla*. A Z-test revealed no overall positive selection neither for the entire dataset nor the four different subsets, adult eels, glass eels, and adult eels and glass eels of the Burrishoole population (Z = 0.236-1.194, p = 0.117-0.407). However, models allowing for positive selection fit the data considerably better than models assuming neutral evolution and infer several sites to be strongly selected (table 2.2). Similar sites were inferred across the different datasets.

Table 2.2 Summary of likelihood ratio tests for site-specific models of positive selection

Group	LRT	2Δl	p-value	Estimate for ω >1	Proportion of PSS	PSS
All individuals	M1a vs M2a	573.529	<0.0001	7.15341	0.27888	**7**, **16**, **17**, **26**, **32**, **40**, **42**, **45**, **46**, 47, **50**, **51**, **53**, **58**, **59**, **62**, **67**, **69**
Adult eels	M1a vs M2a	511.767	<0.0001	6.91363	0.27789	**7**, **16**, **17**, **26**, **32**, **40**, **42**, **45**, **46**, 47, **50**, **51**, **53**, **58**, **62**, **67**, **69**
	M7 vs M8	448.116	<0.0001	6.77153	0.27809	**7**, 16, **17**, **26**, **32**, **40**, **42**, **45**, **46**, 47, **50**, **51**, **53**, **58**, 59, **62**, **67**, **69**
	M8a vs M8	429.097	<0.0001			
Glass eels	M1a vs M2a	370.504	<0.0001	6.60652	0.25332	**7**, **16**, **17**, **26**, **32**, **40**, **42**, **45**, **46**, **50**, **51**, **53**, **58**, **62**, **67**, **69**
	M7 vs M8	324.971	<0.0001	6.54816	0.25326	**7**, **16**, **17**, **26**, **32**, **40**, **42**, **45**, **46**, **50**, **51**, **53**, **58**, **62**, **67**, **69**
	M8a vs M8	319.375	<0.0001			
Burrishoole adult eels	M1a vs M2a	93.902	<0.0001	7.53847	0.23651	**7**, 16, **17**, **40**, 42, **45**, **46**, **53**, **58**, **62**, **67**, **69**
	M7 vs M8	95.844	<0.0001	7.66316	0.23793	**7**, 16, **17**, 32, **40**, 42, **45**, **46**, 50, **53**, **58**, **62**, **67, 69**
	M8a vs M8	89.012	<0.0001			
Burrishoole glass eels	M1a vs M2a	109.834	<0.0001	7.03859	0.23028	**17**, **26**, **32**, **40**, **45**, **46**, **53**, **58**, **67**, **69**
	M7 vs M8	108.244	<0.0001	7.00305	0.23147	**17**, **26**, **32**, **40**, 42, **45**, **46**, **53**, **58**, 62, **67**, **69**
	M8a vs M8	107.424	<0.0001			

LRT, models compared by the likelihood ratio test, $2\Delta\ell = 2(\ell b - \ell a)$, ω = dN/dS, PSS, positively selected sites inferred by Bayes empirical Bayes posterior probabilities, PSS in **bold** are inferred at 99 % level.

Since the composition of the MHC allele pools does not differ between the two species (R = -0,021, p = 0.986, figure 3), *A. anguilla* and *A. rostrata* individuals from the river Schwentine were treated as one population. There seem to be no differences in MHC allele pools neither among populations, life stages, and salinity ranges (R = -0.021, p = 0.986, R = -0.008, p = 0. 790, and R = -0.017, p = 0.927, respectively, figure 2.3), nor between *A. crassus* infected and uninfected individuals of the Schwentine population (R = -0.009, p = 0. 681). Additionally, no pairwise comparison of MHC allele pools among neither populations nor the glass eel cohorts from Adour is significant. However, differentiation at microsatellites and MHC among populations does not correlate (Mantel test: Z = -0.023, p = 0.761). When comparing the mean number of alleles per individual again there are no differences among populations, but adult eels and glass eels differ significantly across all individuals and separately in the Burrishoole population (t = 2.843, p = 0.005 and t = 3.072, p = 0.005, respectively, figure 2.4).

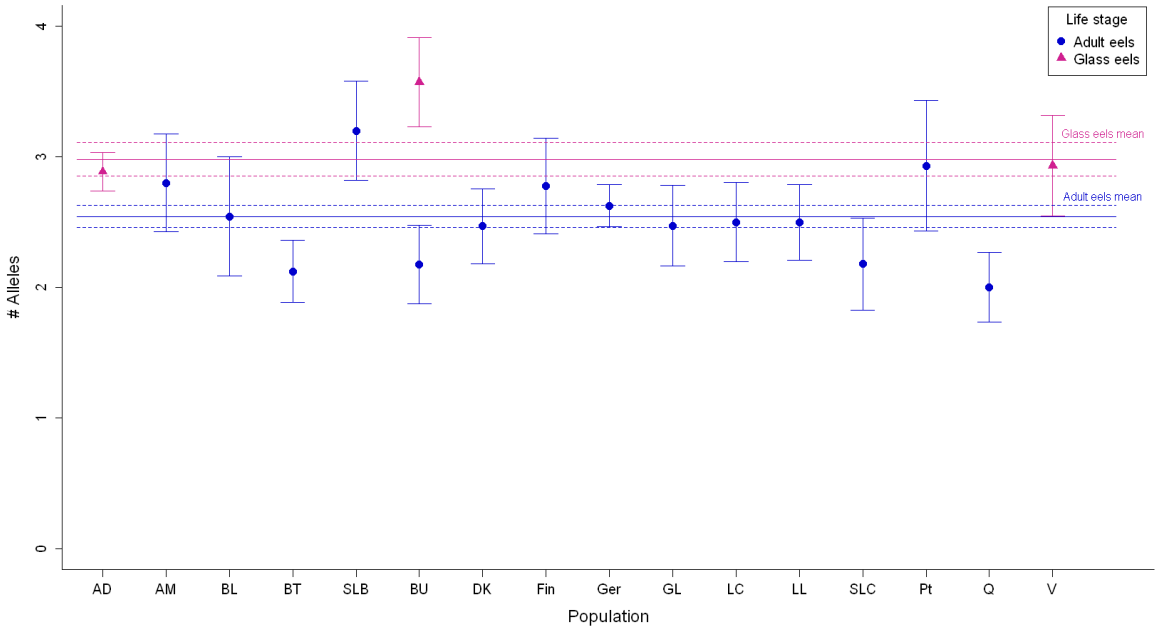


Figure 2.4 Mean number of alleles for each population. Lines indicate the mean number of alleles for the two life stages ±SE. Population labels correspond to labels in table 1.

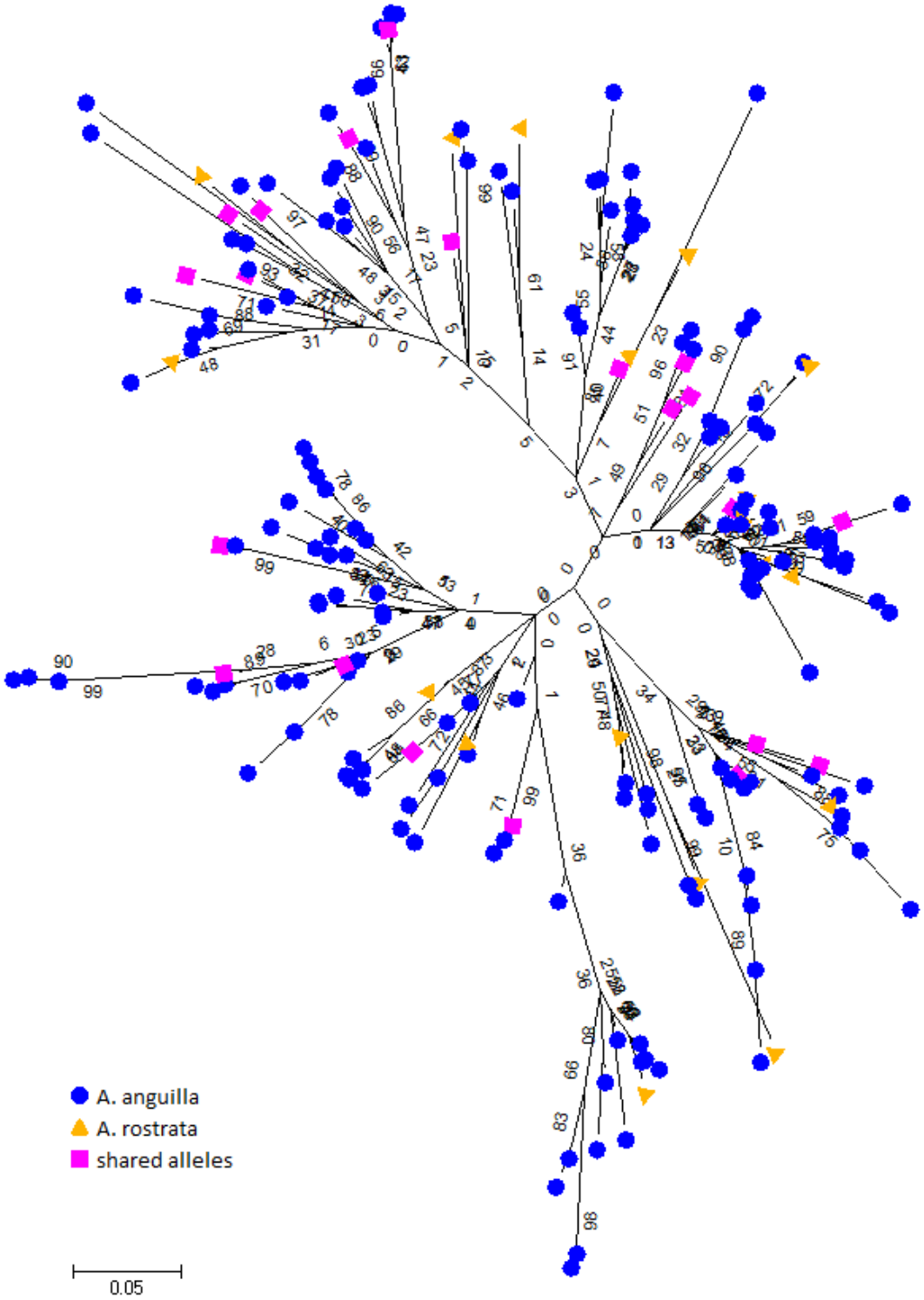


Figure 2.2 Maximum likelihood tree indicating the relationship among *A. anguilla* and *A. rostrata* alleles. Bootstrap values were obtained by 500 replicates.

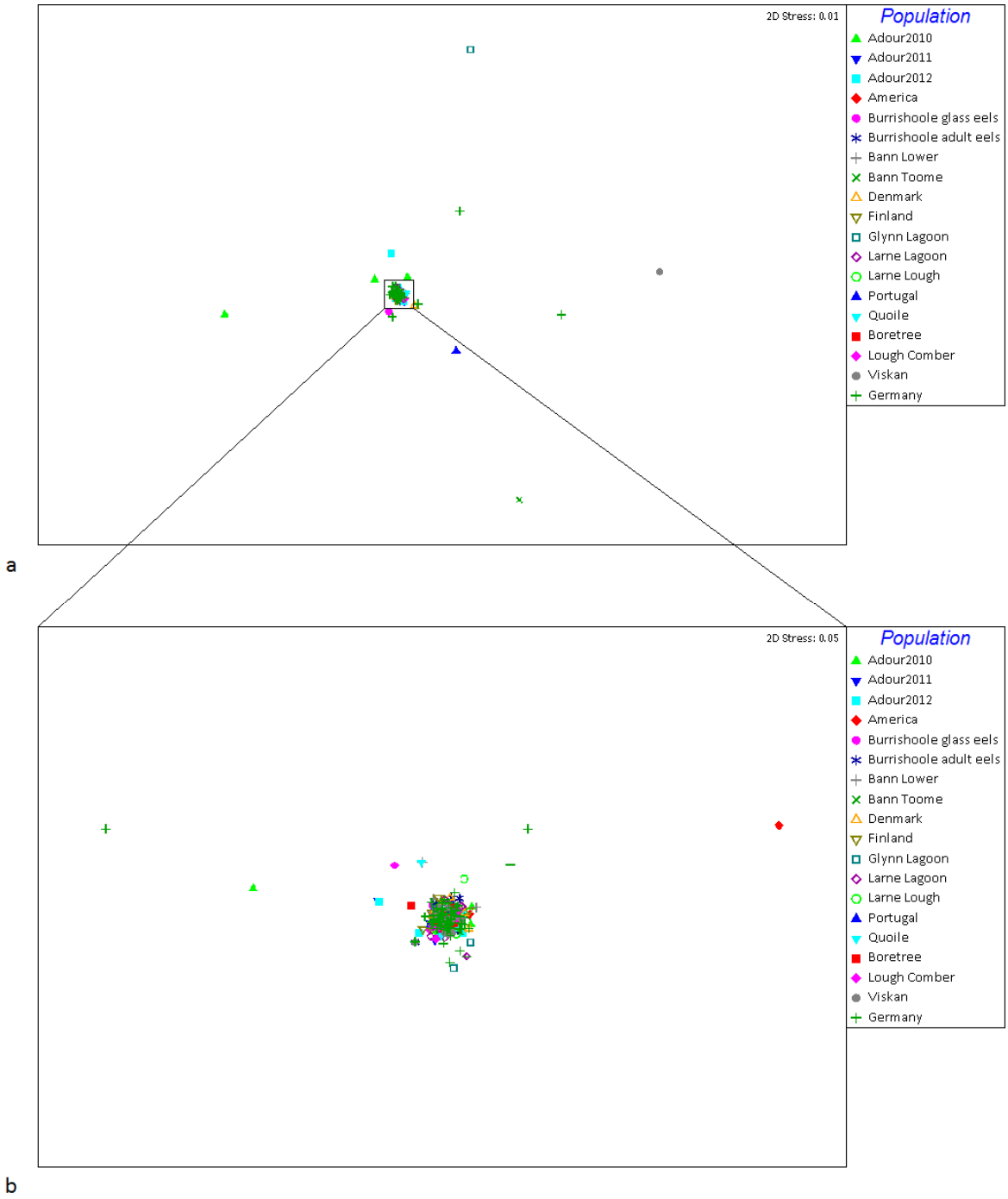


Figure 2.3 MDS plot showing the genetic distance at MHC among individuals of the different population for (a) the entire dataset and (b) a more detailed view of the central cluster.

4. Discussion

MHC alleles of the European eel and the American eel do not fall into separate phylogenetic clusters. Furthermore, a considerable number of alleles is shared between the species indicating the presence of trans-species polymorphism (TSP). TSP arises from the maintenance of allelic lineages over long evolutionary periods that span speciation events and it is a phenomenon commonly observed in nature (Klein *et al.* 1998a; Ottova *et al.* 2005; Kikkawa *et al.* 2009; Lenz *et al.* 2013a).

Unlike in other species, no signal of overall positive selection could be detected at exon 2 of MHC IIB of the European eel (Landry & Bernatchez 2001; Bahr & Wilson 2011; Kuduk *et al.* 2012). However, evolution models infer several sites to be positively selected. Processes acting on the remaining sites might be strong enough to eliminate an overall signal of positive selection.

Selection might also act on the number of MHC alleles per individual which is assumed to be the result of parasite diversity among habitats (Eizaguirre *et al.* 2011). In the European eel, selection seems to act in favour of a higher copy number in glass eels compared to adult eels, but not among different populations. It might be a response not to varying parasite pressure, but to the population decline in order to increase allelic diversity after a possible bottleneck event. However, according to microsatellite data from EST the European eel did not undergo a genetic bottleneck (Pujolar *et al.* 2011b). Alternatively, individuals with a high number of alleles might be deselected. Since MHC diversity and T cell diversity within an individual are assumed to be directly linked, a large number of MHC alleles might be disadvantageous (Woelfing *et al.* 2009).

When comparing the MHC pool diversity no differences were detected for comparisons among populations, life stages, salinity ranges, or infection status. This was not expected, since MHC pools show local adaptation among habitats or habitat types for many taxa (Ekblom *et al.* 2007; Cammen *et al.* 2011; Eizaguirre *et al.* 2011). The panmictic or near-panmictic life history of the eel, however, might prevent local adaptation if there is only weak or moderate differential selective pressure among habitats. On the other hand, differentiation at neutrally evolving microsatellite loci and the MHC does not correlate suggesting the presence of some form selection on MHC. These contradictory findings might be due to incorrect assumptions of the units experiencing similar selection. Parasite communities in the foraging habitats might not be the driving force in shaping MHC allele composition, but rather female philopatry within the Sargasso Sea (Baltazar-Soares *et al.* submitted) might determine the structure of MHC allele pools as seen in turtles (Stiebens *et al.* 2013). To verify this assumption, however, further analysis is needed.

Final discussion

The European eel possesses at least two potentially expressed and functional MHC class IIA and MHC class IIB loci. Sequence variation is high in both genes with genetic distances among alleles are in the upper range of what has been described for teleosts (Reusch *et al.* 2004; Silva *et al.* 2007; Bahr & Wilson 2012). Similarly, extremely high allelic variation at exon 2 of MHC IIB was found in an extensive population survey, that exceeds the number of alleles in many species (Atlantic salmon: 18 in 666 individuals (Landry & Bernatchez 2001), guppy: 39 in 412 individuals (Fraser & Neff 2010), great snipe: 50 in 175 individuals (Ekblom *et al.* 2007), brown bear: 16 in 234 individuals (Kuduk *et al.* 2012)). Furthermore, several sites of both genes are inferred to be subjected to positive selection of which about 50 % correspond to human antigen-binding sites (ABS). This is in line with the situation in other teleosts for which polymorphic sites outside the putative ABS have also been found, indicating that the exact shape of the antigen-binding groove in fish might be different from human HLA molecules (Stet *et al.* 2002; Aguilar & Garza 2007; Wegner 2008). In contrast to the current view, MHC IIA alleles of the European eel seem to be as variable as MHC IIB alleles and experience similar selective pressures (Reche & Reinherz 2003; Reusch *et al.* 2004; Arbanasic *et al.* 2013).

Selection on MHC IIB at the population level could not be detected which might be attributable to the near-panmictic life history of the European eel (Palm *et al.* 2009; Als *et al.* 2011; Pujolar *et al.* 2011a). The proposed random spawning in large aggregations resulting in extreme levels of gene flow followed by random dispersal of the larvae might prevent local adaptation and selection within generations might not be strong enough to overcome this signal. However, genetic differentiation among populations at MHC did not follow genetic differentiation at neutrally evolving microsatellites, suggesting some form of balancing selection. Non-random spawning in the Sargasso Sea due to female philopatry could be a possible explanation for this observation (Baltazar-Soares *et al.* submitted). Under this scenario, structure would arise according to mitochondrial haplotypes and then be mixed during the drift across the Atlantic Ocean and randomly distributed among geographic locations.

Balancing selection might also result in the retention of allelic lineages over long evolutionary times leading to the trans-species polymorphism (TSP) (Klein *et al.* 1998a). For MHC IIA and MHC IIB genes, TSP was present among *A. anguilla* and both, *A. rostrata* and *A. japonica* suggesting that allelic lineages have persisted through the first radiation event in anguillids during which the Atlantic eels have been phylogenetically separated from *A. japonica* (Minegishi *et al.* 2005; Teng *et al.* 2009). Trans-species polymorphism is common in nature and has been observed for most studied taxa (Ottova *et al.* 2005; Glaberman & Caccone 2008; Kikkawa *et al.* 2009).

To conclude, the European eel possesses at least two loci for each MHC II genes that show the typical characteristics of functional genes and selection seems to play an important role in shaping and maintaining the genetic structure. Natural selection might also be important

in maintaining variation in the MHC pool of the entire eel stock, however, the mechanisms and selective agents remain to be identified.

Acknowledgements

First of all, I would like to thank Dr. Christophe Eizaguirre for giving me the opportunity to do my master thesis in his group and for being a committed and patient supervisor. Then, I would like to thank Miguel Soares and Melinda Hofmann for helpful and other assistance, comments, and discussions in every part of the work. Katrin Beining has given technical assistance and without Dr. Freddy Chain and Dr. Till Bayer data analysis would have failed. Last, many thanks to Dr. Enno Prigge, Dr. Chris Harrod, and everyone else who contributed samples.

References

Aarestrup, K., Okland, F., Hansen, M.M., Righton, D., Gargan, P., Castonguay, M. *et al.* (2009). Oceanic Spawning Migration of the European Eel (Anguilla anguilla). *Science*, 325, 1660-1660.

Aguilar, A. & Garza, J.C. (2006). A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (Oncorhynchus mykiss Walbaum). *Molecular Ecology*, 15, 923-937.

Aguilar, A. & Garza, J.C. (2007). Patterns of historical balancing selection on the salmonid major histocompatibility complex class II beta gene. *Journal of Molecular Evolution*, 65, 34-43.

Alcaide, M. (2010). On the relative roles of selection and genetic drift in shaping MHC variation. *Molecular Ecology*, 19, 3842-3844.

Als, T.D., Hansen, M.M., Maes, G.E., Castonguay, M., Riemann, L., Aarestrup, K. *et al.* (2011). All roads lead to home: panmixia of European eel in the Sargasso Sea. *Molecular Ecology*, 20, 1333-1346.

Apanius, V., Penn, D., Slev, P.R., Ruff, L.R. & Potts, W.K. (1997). The nature of selection on the major histocompatibility complex. *Critical Reviews in Immunology*, 17, 179-224.

Arbanasic, H., Huber, D., Kusak, J., Gomercic, T., Hrenovic, J. & Galov, A. (2013). Extensive polymorphism and evidence of selection pressure on major histocompatibility complex DLA-DRB1, DQA1 and DQB1 class II genes in Croatian grey wolves. *Tissue Antigens*, 81, 19-27.

Bahr, A. & Wilson, A.B. (2011). The impact of sex-role reversal on the diversity of the major histocompatibility complex: Insights from the seahorse (Hippocampus abdominalis). *Bmc Evolutionary Biology*, 11.

Bahr, A. & Wilson, A.B. (2012). The evolution of MHC diversity: Evidence of intralocus gene conversion and recombination in a single-locus system. *Gene*, 497, 52-57.

Baltazar-Soares, M., Biastoch, A., Harrod, C., Hanel, R., Marohn, L., Prigge, E. *et al.* (submitted). Recruitment collapse and population structure of the European eel shaped by local ocean current dynamics.

Bernatchez, L. & Landry, C. (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology*, 16, 363-377.

Brown, J.H., Jardetzky, T.S., Gorga, J.C., Stern, L.J., Urban, R.G., Strominger, J.L. *et al.* (1993). 3-DIMENSIONAL STRUCTURE OF THE HUMAN CLASS-II HISTOCOMPATIBILITY ANTIGEN HLA-DR1. *Nature*, 364, 33-39.

Cammen, K., Hoffman, J.I., Knapp, L.A., Harwood, J. & Amos, W. (2011). Geographic variation of the major histocompatibility complex in Eastern Atlantic grey seals (Halichoerus grypus). *Molecular Ecology*, 20, 740-752.

Clarke, K.R. (1993). NONPARAMETRIC MULTIVARIATE ANALYSES OF CHANGES IN COMMUNITY STRUCTURE. *Australian Journal of Ecology*, 18, 117-143.

Coppe, A., Pujolar, J.M., Maes, G.E., Larsen, P.F., Hansen, M.M., Bernatchez, L. *et al.* (2010). Sequencing, de novo annotation and analysis of the first Anguilla anguilla transcriptome: EeelBase opens new perspectives for the study of the critically endangered european eel. *Bmc Genomics*, 11.

Cuesta, A., Esteban, M.A. & Meseguer, J. (2006). Cloning, distribution and up-regulation of the teleost fish MHC class II alpha suggests a role for granulocytes as antigen-presenting cells. *Molecular Immunology*, 43, 1275-1285.

Daemen, E., Cross, T., Ollevier, F. & Volckaert, F.A.M. (2001). Analysis of the genetic structure of European eel (Anguilla anguilla) using microsatellite DNA and mtDNA markers. *Marine Biology*, 139, 755-764.

Dannewitz, J., Maes, G.E., Johansson, L., Wickstrom, H., Volckaert, F.A.M. & Jarvi, T. (2005). Panmixia in the European eel: a matter of time. *Proceedings of the Royal Society B-Biological Sciences*, 272, 1129-1137.

De Charleroy, D., Grisez, L., Thomas, K., Belpaire, C. & Ollevier, F. (1990). THE LIFE-CYCLE OF ANGUILLICOLA-CRASSUS. *Diseases of Aquatic Organisms*, 8, 77-84.

Decaestecker, E., Gaba, S., Raeymaekers, J.A.M., Stoks, R., Van Kerckhoven, L., Ebert, D. *et al.* (2007). Host-parasite 'Red Queen' dynamics archived in pond sediment. *Nature*, 450, 870-U816.

Dekker, W. (2003). Did lack of spawners cause the collapse of the European eel, Anguilla anguilla? *Fisheries Management and Ecology*, 10, 365-376.

Dionne, M., Miller, K.M., Dodson, J.J. & Bernatchez, L. (2009). MHC standing genetic variation and pathogen resistance in wild Atlantic salmon. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 364, 1555-1565.

Dionne, M., Miller, K.M., Dodson, J.J., Caron, F. & Bernatchez, L. (2007). Clinal variation in mhc diversity with temperature: Evidence for the role of host-pathogen interaction on local adaptation in Atlantic salmon. *Evolution*, 61, 2154-2164.

Doherty, P.C. & Zinkernagel, R.M. (1975). BIOLOGICAL ROLE FOR MAJOR HISTOCOMPATIBILITY ANTIGENS. *Lancet*, 1, 1406-1409.

Eizaguirre, C. & Lenz, T.L. (2010). Major histocompatibility complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. *Journal of Fish Biology*, 77, 2023-2047.

Eizaguirre, C., Lenz, T.L., Kalbe, M. & Milinski, M. (2012a). Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters*, 15, 723-731.

Eizaguirre, C., Lenz, T.L., Kalbe, M. & Milinski, M. (2012b). Rapid and adaptive evolution of MHC genes under parasite selection in experimental vertebrate populations. *Nature Communications*, 3.

Eizaguirre, C., Lenz, T.L., Sommerfeld, R.D., Harrod, C., Kalbe, M. & Milinski, M. (2011). Parasite diversity, patterns of MHC II variation and olfactory based mate choice in diverging three-spined stickleback ecotypes. *Evolutionary Ecology*, 25, 605-622.

Ekblom, R., Saether, S.A., Jacobsson, P., Fiske, P., Sahlman, T., Grahn, M. *et al.* (2007). Spatial pattern of MHC class II variation in the great snipe (Gallinago media). *Molecular Ecology*, 16, 1439-1451.

Evans, M.L. & Neff, B.D. (2009). Major histocompatibility complex heterozygote advantage and widespread bacterial infections in populations of Chinook salmon (Oncorhynchus tshawytscha). *Molecular Ecology*, 18, 4716-4729.

Evans, M.L., Neff, B.D. & Heath, D.D. (2010). MHC genetic structure and divergence across populations of Chinook salmon (Oncorhynchus tshawytscha). *Heredity*, 104, 449-459.

Fraser, B.A. & Neff, B.D. (2010). Parasite mediated homogenizing selection at the MHC in guppies. *Genetica*, 138, 273-278.

Friedland, K.D., Miller, M.J. & Knights, B. (2007). Oceanic changes in the Sargasso Sea and declines in recruitment of the European eel. *Ices Journal of Marine Science*, 64, 519-530.

Glaberman, S. & Caccone, A. (2008). Species-specific evolution of class I MHC genes in iguanas (Order : Squamata; subfamily : Iguaninae). *Immunogenetics*, 60, 371-382.

Godwin, U.B., Flores, M., Quiniou, S., Wilson, M.R., Miller, N.W., Clem, L.W. *et al.* (2000). MHC class II A genes in the channel catfish (Ictalurus punctatus). *Developmental and Comparative Immunology*, 24, 609-622.

Gollock, M.J., Kennedy, C.R. & Brown, J.A. (2005). European eels, Anguilla anguilla (L.), infected with Anguillicola crassus exhibit a more pronounced stress response to severe hypoxia than uninfected eels. *Journal of Fish Diseases*, 28, 429-436.

Graham, J., McNeney, B. & Seillier-Moiseiwitsch, F. (2005). Stepwise detection of recombination breakpoints in sequence alignments. *Bioinformatics*, 21, 589-595.

Grimholt, U., Getahun, A., Hermsen, T. & Stet, R.J.M. (2000). The major histocompatibility class II alpha chain in salmonid fishes. *Developmental and Comparative Immunology*, 24, 751-763.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98.

Hedrick, P.W. (2002). Pathogen resistance and genetic variation at MHC loci. *Evolution*, 56, 1902-1908.

Huson, D.H. & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23, 254-267.

ICES (2012). Report of the Joint EIFAAC/ICES Working Group on Eels (WGEEL). ICES CM 2012/ACOM:18 Copenhagen, Denmark, p. 824.

Inoue, J.G., Miya, M., Tsukamoto, K. & Nishida, M. (2001). A mitogenomic perspective on the basal teleostean phylogeny: Resolving higher-level relationships with longer DNA sequences. *Molecular Phylogenetics and Evolution*, 20, 275-285.

Janeway, C.A., Travers, P., Walport, M. & Shlomchik, M.J. (2001). *Immunobiology : the immune system in health and disease*. 5th edn. Garland Publishing, New York, USA.

Kekalainen, J., Vallunen, J.A., Primmer, C.R., Rattya, J. & Taskinen, J. (2009). Signals of major histocompatibility complex overdominance in a wild salmonid population. *Proceedings of the Royal Society B-Biological Sciences*, 276, 3133-3140.

Kiemnec-Tyburczy, K.M., Richmond, J.Q., Savage, A.E. & Zamudio, K.R. (2010). Selection, trans-species polymorphism, and locus identification of major histocompatibility complex class II beta alleles of New World ranid frogs. *Immunogenetics*, 62, 741-751.

Kikkawa, E.F., Tsuda, T.T., Sumiyama, D., Naruse, T.K., Fukuda, M., Kurita, M. *et al.* (2009). Trans-species polymorphism of the Mhc class II DRB-like gene in banded penguins (genus Spheniscus). *Immunogenetics*, 61, 341-352.

King, G. & Dixon, A.M. (2010). Evidence for role of transmembrane helix-helix interactions in the assembly of the Class II major histocompatibility complex. *Molecular Biosystems*, 6, 1650-1661.

Kirk, R.S. (2003). The impact of Anguillicola crassus on European eels. *Fisheries Management and Ecology*, 10, 385-394.

Klein, J., Bontrop, R.E., Dawkins, R.L., Erlich, H.A., Gyllensten, U.B., Heise, E.R. *et al.* (1990). NOMENCLATURE FOR THE MAJOR HISTOCOMPATIBILITY COMPLEXES OF DIFFERENT SPECIES - A PROPOSAL. *Immunogenetics*, 31, 217-219.

Klein, J., Sato, A., Nagl, S. & O'HUigin, C. (1998a). Molecular trans-species polymorphism. *Annual Review of Ecology and Systematics*, 29, 1-+.

Klein, J., Sato, A. & O'HUigin, C. (1998b). Evolution by gene duplication in the major histocompatibility complex. *Cytogenetics and Cell Genetics*, 80, 123-127.

Knights, B. (2003). A review of the possible impacts of long-term oceanic and climate changes and fishing mortality on recruitment of anguillid eels of the Northern Hemisphere. *Science of the Total Environment*, 310, 237-244.

Knopf, K., Naser, K., van der Heijden, M.H.T. & Taraschewski, H. (2000). Humoral immune response of European eel Anguilla anguilla experimentally infected with Anguillicola crassus. *Diseases of Aquatic Organisms*, 42, 61-69.

Kruiswijk, C.P., Hermsen, T., Fujiki, K., Dixon, B., Savelkoul, H.F.J. & Stet, R.J.M. (2004). Analysis of genomic and expressed major histocompatibility class Ia and class II genes in a hexaploid Lake Tana African 'large' barb individual (Barbus intermedius). *Immunogenetics*, 55, 770-781.

Kubinak, J.L., Ruff, J.S., Hyzer, C.W., Slev, P.R. & Potts, W.K. (2012). Experimental viral evolution to specific host MHC genotypes reveals fitness and virulence trade-offs in alternative MHC types. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 3422-3427.

Kuduk, K., Babik, W., Bojarska, K., Sliwinska, E.B., Kindberg, J., Taberlet, P. *et al.* (2012). Evolution of major histocompatibility complex class I and class II genes in the brown bear. *Bmc Evolutionary Biology*, 12.

Kuroda, N., Figueroa, F., O'HUigin, C. & Klein, J. (2002). Evidence that the separation of Mhc class II from class I loci in the zebrafish, Danio rerio, occurred by translocation. *Immunogenetics*, 54, 418-430.

Landry, C. & Bernatchez, L. (2001). Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (Salmo salar). *Molecular Ecology*, 10, 2525-2539.

Lefebvre, F., Fazio, G., Mounaix, B. & Crivelli, A.J. (2013). Is the continental life of the European eel Anguilla anguilla affected by the parasitic invader Anguillicoloides crassus? *Proceedings of the Royal Society B-Biological Sciences*, 280.

Lefebvre, F.S. & Crivelli, A.J. (2004). Anguillicolosis: dynamics of the infection over two decades. *Diseases of Aquatic Organisms*, 62, 227-232.

Lenz, T.L. (2011). COMPUTATIONAL PREDICTION OF MHC II-ANTIGEN BINDING SUPPORTS DIVERGENT ALLELE ADVANTAGE AND EXPLAINS TRANS-SPECIES POLYMORPHISM. *Evolution*, 65, 2380-2390.

Lenz, T.L. & Becker, S. (2008). Simple approach to reduce PCR artefact formation leads to reliable genotyping of MHC and other highly polymorphic loci - Implications for evolutionary analysis. *Gene*, 427, 117-123.

Lenz, T.L., Eizaguirre, C., Kalbe, M. & Milinski, M. (2013a). EVALUATING PATTERNS OF CONVERGENT EVOLUTION AND TRANS-SPECIES POLYMORPHISM AT MHC IMMUNOGENES IN TWO SYMPATRIC STICKLEBACK SPECIES. *Evolution*, n/a-n/a.

Lenz, T.L., Mueller, B., Trillmich, F. & Wolf, J.B.W. (2013b). Divergent allele advantage at MHC-DRB through direct and maternal genotypic effects and its consequences for allele pool composition and mating. *Proceedings of the Royal Society B-Biological Sciences*, 280.

Lenz, T.L., Wells, K., Pfeiffer, M. & Sommer, S. (2009). Diverse MHC IIB allele repertoire increases parasite resistance and body condition in the Long-tailed giant rat (Leopoldamys sabanus). *Bmc Evolutionary Biology*, 9.

Librado, P. & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451-1452.

Loiseau, C., Richard, M., Garnier, S., Chastel, O., Julliard, R., Zoorob, R. *et al.* (2009). Diversifying selection on MHC class I in the house sparrow (Passer domesticus). *Molecular Ecology*, 18, 1331-1340.

Loiseau, C., Zoorob, R., Robert, A., Chastel, O., Julliard, R. & Sorci, G. (2011). Plasmodium relictum infection and MHC diversity in the house sparrow (Passer domesticus). *Proceedings of the Royal Society B-Biological Sciences*, 278, 1264-1272.

Lu, D.Q., Yi, S.B., Yao, M., Li, Y.W., Liu, X.C., Zhang, Y. *et al.* (2012). Identification and expression analysis of major histocompatibility complex IIB gene in orange-spotted grouper Epinephelus coioides. *Journal of Fish Biology*, 81, 165-180.

Luijckx, P., Fienberg, H., Duneau, D. & Ebert, D. (2013). A Matching-Allele Model Explains Host Resistance to Parasites. *Current Biology*, 23, 1085-1088.

Maes, G.E., Pujolar, J.M., Hellemans, B. & Volckaert, F.A.M. (2006). Evidence for isolation by time in the European eel (Anguilla anguilla L.). *Molecular Ecology*, 15, 2095-2107.

Maes, G.E. & Volckaert, F.A.M. (2002). Clinal genetic variation and isolation by distance in the European eel Anguilla anguilla (L.). *Biological Journal of the Linnean Society*, 77, 509-521.

Malaga-Trillo, E., Zaleska-Rutczynska, Z., McAndrew, B., Vincek, V., Figueroa, F., Sultmann, H. *et al.* (1998). Linkage relationships and haplotype polymorphism among cichlid Mhc class II B loci. *Genetics*, 149, 1527-1537.

Milinski, M. (2006). The major histocompatibility complex, sexual selection, and mate choice. *Annual Review of Ecology Evolution and Systematics*, 37, 159-186.

Minegishi, Y., Aoyama, J., Inoue, J.G., Miya, M., Nishida, M. & Tsukamoto, K. (2005). Molecular phylogeny and evolution of the freshwater eels genus Anguilla based on the whole mitochondrial genome sequences. *Molecular Phylogenetics and Evolution*, 34, 134-146.

Murray, B.W., Shintani, S., Sultmann, H. & Klein, J. (2000). Major histocompatibility complex class II A genes in cichlid fishes: identification, expression, linkage relationships, and haplotype variation. *Immunogenetics*, 51, 576-586.

Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P. *et al.* (2012). Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 13698-13703.

Nei, M., Gu, X. & Sitnikova, T. (1997). Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 7799-7806.

Nei, M. & Rooney, A.P. (2005). Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics*, 39, 121-152.

Oliver, M.K., Telfer, S. & Piertney, S.B. (2009). Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water vole (Arvicola terrestris). *Proceedings of the Royal Society B-Biological Sciences*, 276, 1119-1128.

Ono, H., Ohuigin, C., Vincek, V. & Klein, J. (1993). EXON-INTRON ORGANIZATION OF FISH MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-II B-GENES. *Immunogenetics*, 38, 223-234.

Osborne, M.J. & Turner, T.F. (2011). Isolation and characterization of major histocompatibility class II beta genes in an endangered North American cyprinid fish, the Rio Grande silvery minnow (Hybognathus amarus). *Fish & Shellfish Immunology*, 30, 1275-1282.

Ottova, E., Simkova, A., Martin, J.F., de Bellocq, J.G., Gelnar, M., Allienne, J.F. *et al.* (2005). Evolution and trans-species polymorphism of MHC class II beta genes in cyprinid fish. *Fish & Shellfish Immunology*, 18, 199-222.

Palm, S., Dannewitz, J., Prestegaard, T. & Wickstrom, H. (2009). Panmixia in European eel revisited: no genetic difference between maturing adults from southern and northern Europe. *Heredity*, 103, 82-89.

Palstra, A.P., Heppener, D.F.M., van Ginneken, V.J.T., Szekely, C. & van den Thillart, G. (2007). Swimming performance of silver eels is severely impaired by the swim-bladder parasite Anguillicola crassus. *Journal of Experimental Marine Biology and Ecology*, 352, 244-256.

Penn, D.J., Damjanovich, K. & Potts, W.K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 11260-11264.

Pujolar, J.M., Bevacqua, D., Andrello, M., Capoccioni, F., Ciccotti, E., De Leo, G.A. *et al.* (2011a). Genetic patchiness in European eel adults evidenced by molecular genetics and population dynamics modelling. *Molecular Phylogenetics and Evolution*, 58, 198-206.

Pujolar, J.M., Bevacqua, D., Capoccioni, F., Ciccotti, E., De Leo, G.A. & Zane, L. (2009). Genetic variability is unrelated to growth and parasite infestation in natural populations of the European eel (Anguilla anguilla). *Molecular Ecology*, 18, 4604-4616.

Pujolar, J.M., Bevacqua, D., Capoccioni, F., Ciccotti, E., De Leo, G.A. & Zane, L. (2011b). No apparent genetic bottleneck in the demographically declining European eel using molecular genetics and forward-time simulations. *Conservation Genetics*, 12, 813-825.

R Core Team (2013). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, Austria, http://www.R-project.org/.

Raeymaekers, J.A.M., Hablutzel, P.I., Gregoir, A.F., Bamps, J., Roose, A.K., Vanhove, M.P.M. *et al.* (2013). Contrasting parasite communities among allopatric colour morphs of the Lake Tanganyika cichlid Tropheus. *Bmc Evolutionary Biology*, 13.

Reche, P.A. & Reinherz, E.L. (2003). Sequence variability analysis of human class I and class II MHC molecules: Functional and structural correlates of amino acid polymorphisms. *Journal of Molecular Biology*, 331, 623-641.

Reusch, T.B.H. & Langefors, A. (2005). Inter- and intralocus recombination drive MHC class IIB gene diversification in a teleost, the three-spined stickleback Gasterosteus aculeatus. *Journal of Molecular Evolution*, 61, 531-U545.

Reusch, T.B.H., Schaschl, H. & Wegner, K.M. (2004). Recent duplication and inter-locus gene conversion in major histocompatibility class II genes in a teleost, the three-spined stickleback. *Immunogenetics*, 56, 427-437.

Sato, A., Dongak, R., Hao, L., Shintani, S. & Sato, T. (2012). Organization of Mhc class II A and B genes in the tilapiine fish Oreochromis. *Immunogenetics*, 64, 679-690.

Sato, A., Figueroa, F., Murray, B.W., Malaga-Trillo, E., Zaleska-Rutczynska, Z., Sultmann, H. *et al.* (2000). Nonlinkage of major histocompatibility complex class I and class II loci in bony fishes. *Immunogenetics*, 51, 108-116.

Sawyer, S.A. (1999). GENECONV: A computer package for the statistical detection of gene conversion. Distributed by the author Department of Mathematics, Washington University in St. Louis, http://www.math.wustl.edu/~sawyer.

Schmidt, J. (1923). The Breeding Places of the Eel. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 211, 179-208.

Shum, B.P., Guethlein, L., Flodin, L.R., Adkison, M.A., Hedrick, R.P., Nehring, R.B. *et al.* (2001). Modes of salmonid MHC class I and II evolution differ from the primate paradigm. *Journal of Immunology*, 166, 3297-3308.

Silva, D.S.P., Reis, M.I.R., Nascimento, D.S., do Vale, A., Pereira, P.J.B. & dos Santos, N.M.S. (2007). Sea bass (Dicentrarchus labrax) invariant chain and class II major histocompatibility complex: Sequencing and structural analysis using 3D homology modelling. *Molecular Immunology*, 44, 3758-3776.

Sin, Y.W., Dugdale, H.L., Newman, C., Macdonald, D.W. & Burke, T. (2012). MHC class II genes in the European badger (Meles meles): characterization, patterns of variation, and transcription analysis. *Immunogenetics*, 64.

Spurgin, L.G. & Richardson, D.S. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B-Biological Sciences*, 277, 979-988.

Spurgin, L.G., van Oosterhout, C., Illera, J.C., Bridgett, S., Gharbi, K., Emerson, B.C. *et al.* (2011). Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations. *Molecular Ecology*, 20, 5213-5225.

Star, B., Nederbragt, A.J., Jentoft, S., Grimholt, U., Malmstrom, M., Gregers, T.F. *et al.* (2011). The genome sequence of Atlantic cod reveals a unique immune system. *Nature*, 477, 207-210.

Stet, R.J.M., de Vries, B., Mudde, K., Hermsen, T., van Heerwaarden, J., Shum, B.P. *et al.* (2002). Unique haplotypes of co-segregating major histocompatibility class II A and class II B alleles in Atlantic salmon (Salmo salar) give rise to diverse class II genotypes. *Immunogenetics*, 54, 320-331.

Stiebens, V.A., Merino, S.E., Roder, C., Chain, F.J.J., Lee, P.L.M. & Eizaguirre, C. (2013). Living on the edge: how philopatry maintains adaptive potential. *Proceedings of the Royal Society B: Biological Sciences*, 280.

Sultmann, H., Mayer, W.E., Figueroa, F., Ohuigin, C. & Klein, J. (1993). ZEBRAFISH MHC CLASS-II ALPHA-CHAIN-ENCODING GENES - POLYMORPHISM, EXPRESSION, AND FUNCTION. *Immunogenetics*, 38, 408-420.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, 28, 2731-2739.

Teng, H.Y., Lin, Y.S. & Tzeng, C.S. (2009). A New Anguilla Species and a Reanalysis of the Phylogeny of Freshwater Eels. *Zoological Studies*, 48, 808-822.

Tesch, F.-W. (2003). *The Eel*. 3rd edn. Blackwell Science Ltd, Oxford, UK.

van Erp, S.H.M., Egberts, E. & Stet, R.M. (1996). Characterization of class IIA and B genes in a gynogenetic carp clone. *Immunogenetics*, 44, 192-202.

van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, 1, 1-30.

Wakeland, E.K., Boehme, S., She, J.X., Lu, C.C., McIndoe, R.A., Cheng, I. *et al.* (1990). ANCESTRAL POLYMORPHISMS OF MHC CLASS-II GENES - DIVERGENT ALLELE ADVANTAGE. *Immunologic Research*, 9, 115-122.

Weclawski, U., Heitlinger, E.G., Baust, T., Klar, B., Petney, T., Han, Y.S. *et al.* (2013). Evolutionary divergence of the swim bladder nematode Anguillicola crassus after colonization of a novel host, Anguilla anguilla. *Bmc Evolutionary Biology*, 13.

Wegner, K.M. (2008). Historical and contemporary selection of teleost MHC genes: did we leave the past behind? *Journal of Fish Biology*, 73, 2110-2132.

Wegner, K.M., Reusch, T.B.H. & Kalbe, M. (2003). Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *Journal of Evolutionary Biology*, 16, 224-232.

Wielgoss, S., Taraschewski, H., Meyer, A. & Wirth, T. (2008). Population structure of the parasitic nematode Anguillicola crassus, an invader of declining North Atlantic eel stocks. *Molecular Ecology*, 17, 3478-3495.

Wirth, T. & Bernatchez, L. (2001). Genetic evidence against panmixia in the European eel. *Nature*, 409, 1037-1040.

Wirth, T. & Bernatchez, L. (2003). Decline of North Atlantic eels: a fatal synergy? *Proceedings of the Royal Society of London Series B-Biological Sciences*, 270, 681-688.

Woelfing, B., Traulsen, A., Milinski, M. & Boehm, T. (2009). Does intra-individual major histocompatibility complex diversity keep a golden mean? *Philosophical Transactions of the Royal Society B-Biological Sciences*, 364, 117-128.

Xu, T.J., Chen, S.L., Ji, X.S. & Sha, Z.X. (2009). Molecular cloning, genomic structure, polymorphism and expression analysis of major histocompatibility complex class IIA and IIB genes of half-smooth tongue sole (Cynoglossus semilaevis). *Fish & Shellfish Immunology*, 27, 192-201.

Xu, T.J., Sun, Y.N., Shi, G., Cheng, Y.Z. & Wang, R.X. (2011). Characterization of the Major Histocompatibility Complex Class II Genes in Miiuy Croaker. *Plos One*, 6.

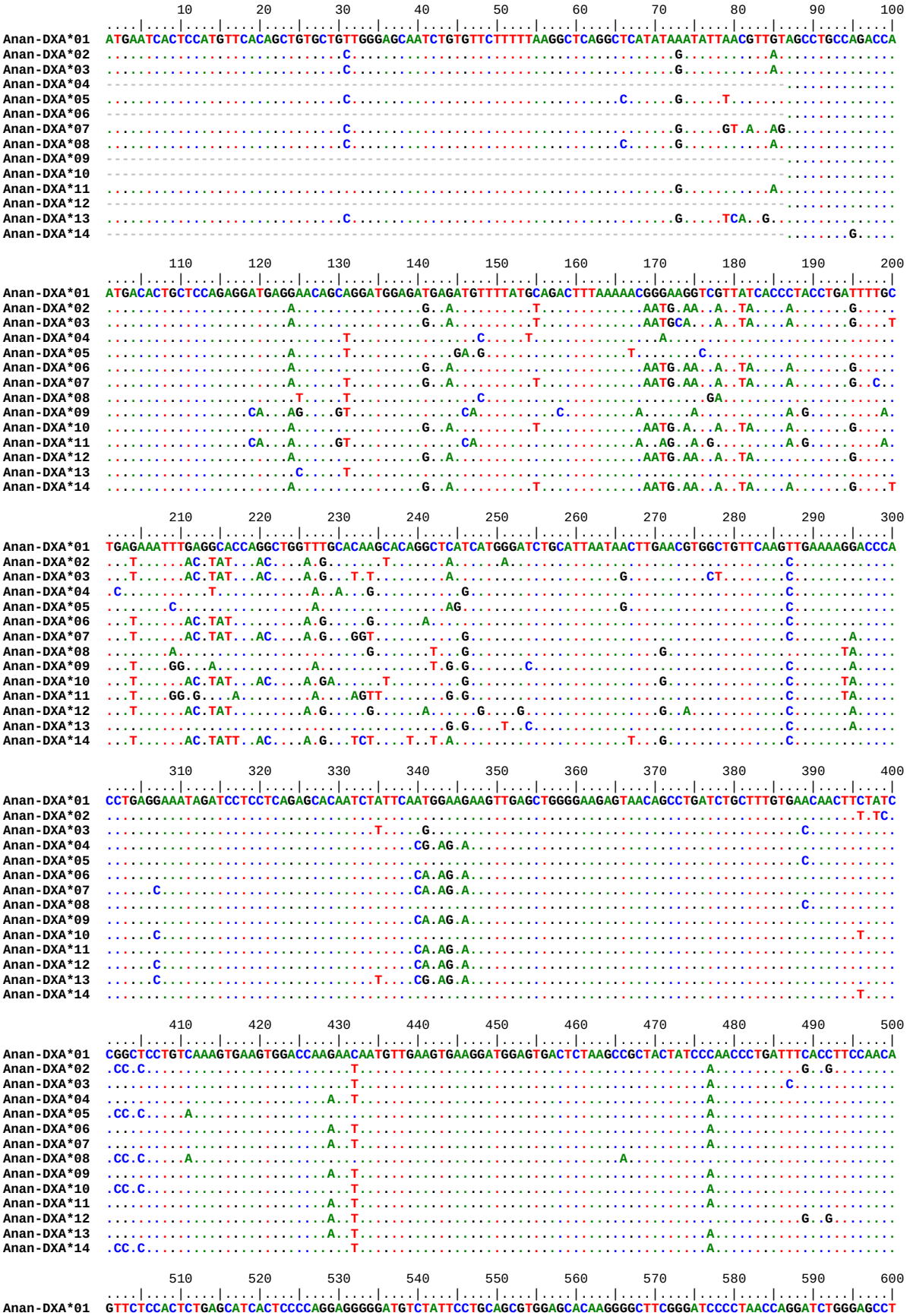
Yang, Z.H. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24, 1586-1591.

Yang, Z.H., Wong, W.S.W. & Nielsen, R. (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution*, 22, 1107-1118.

Picture title page: http://mobile.wwf.ch//fisch

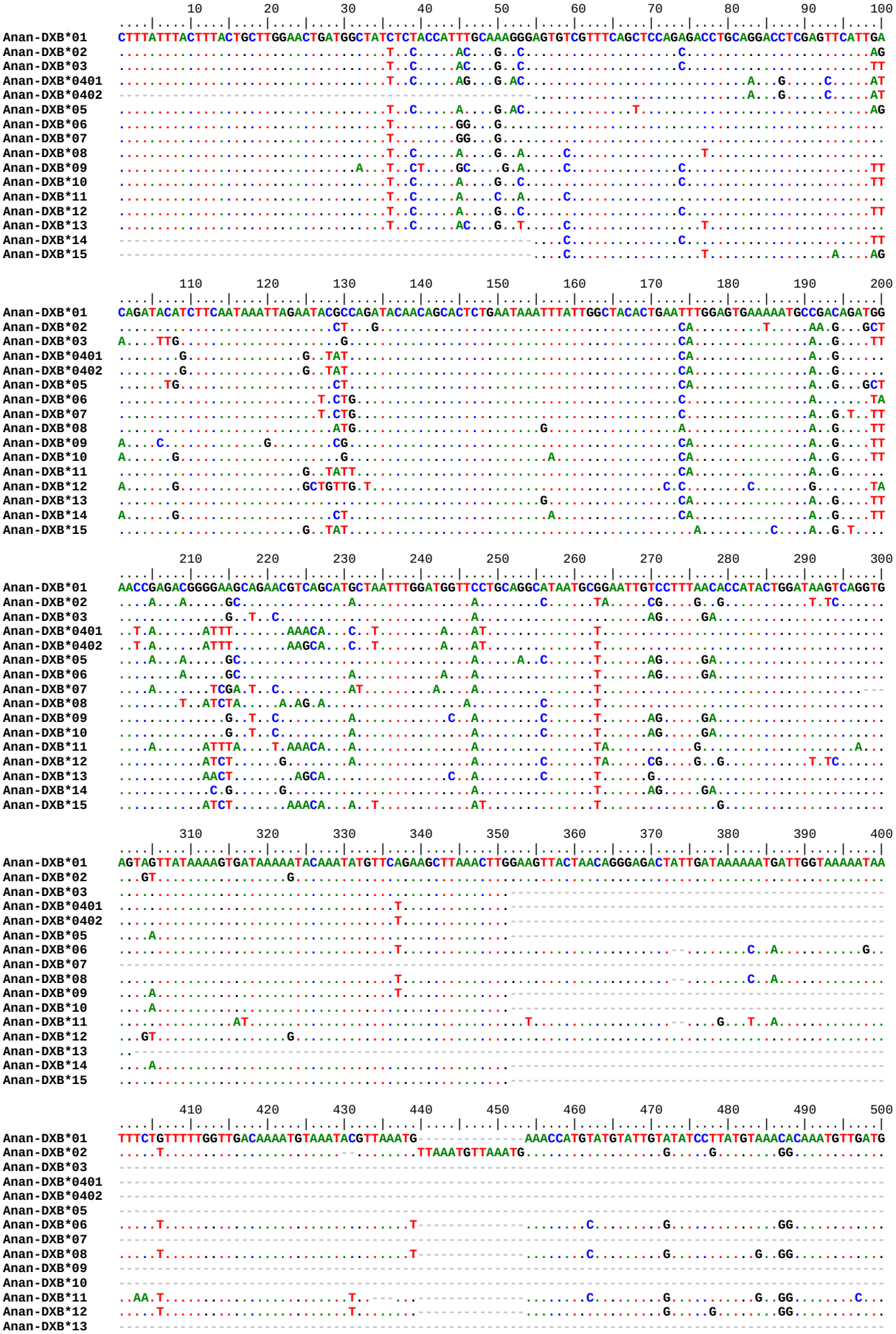
Appendix 1

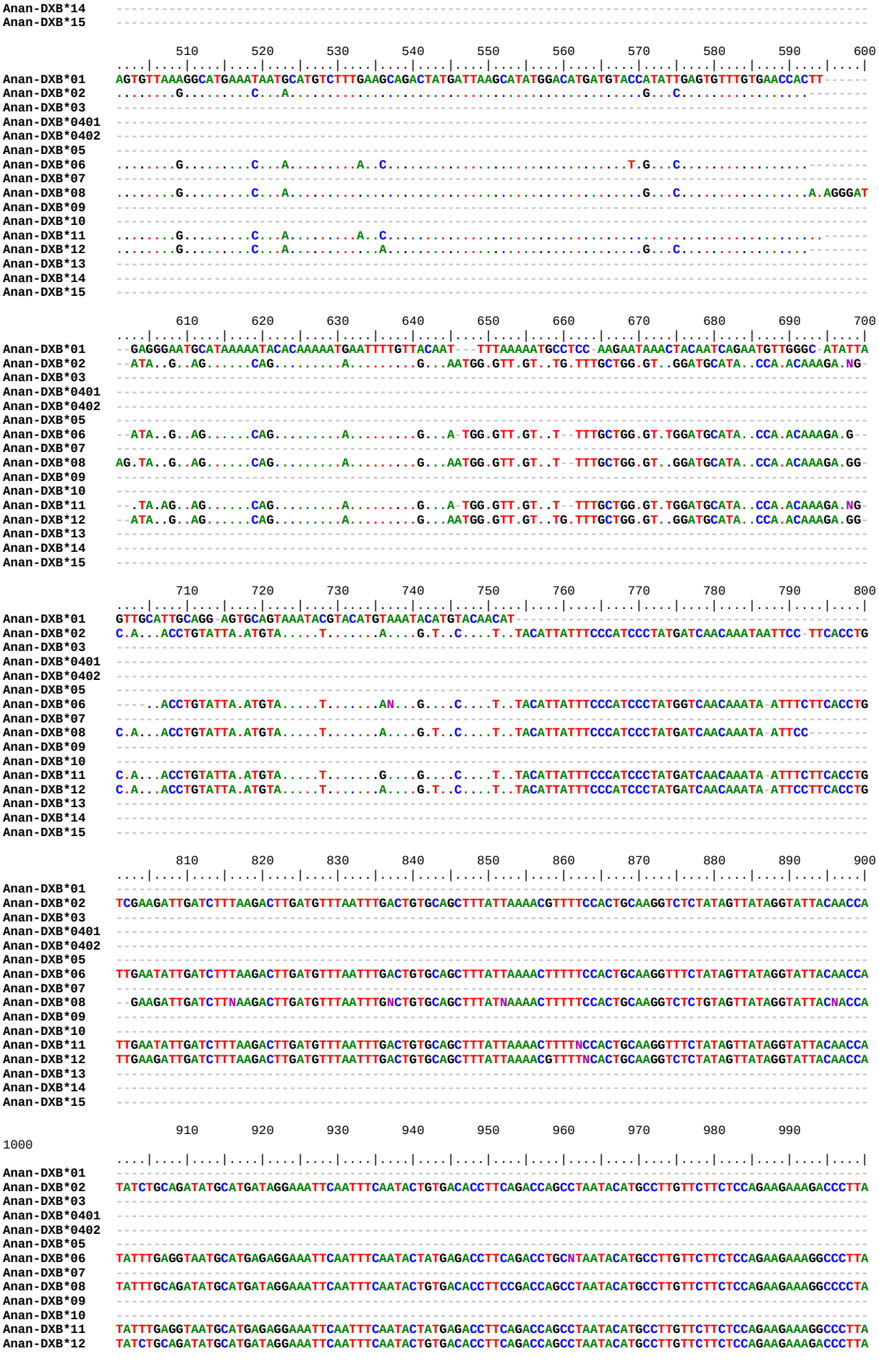
Nucleotide sequences of MHC IIA alleles

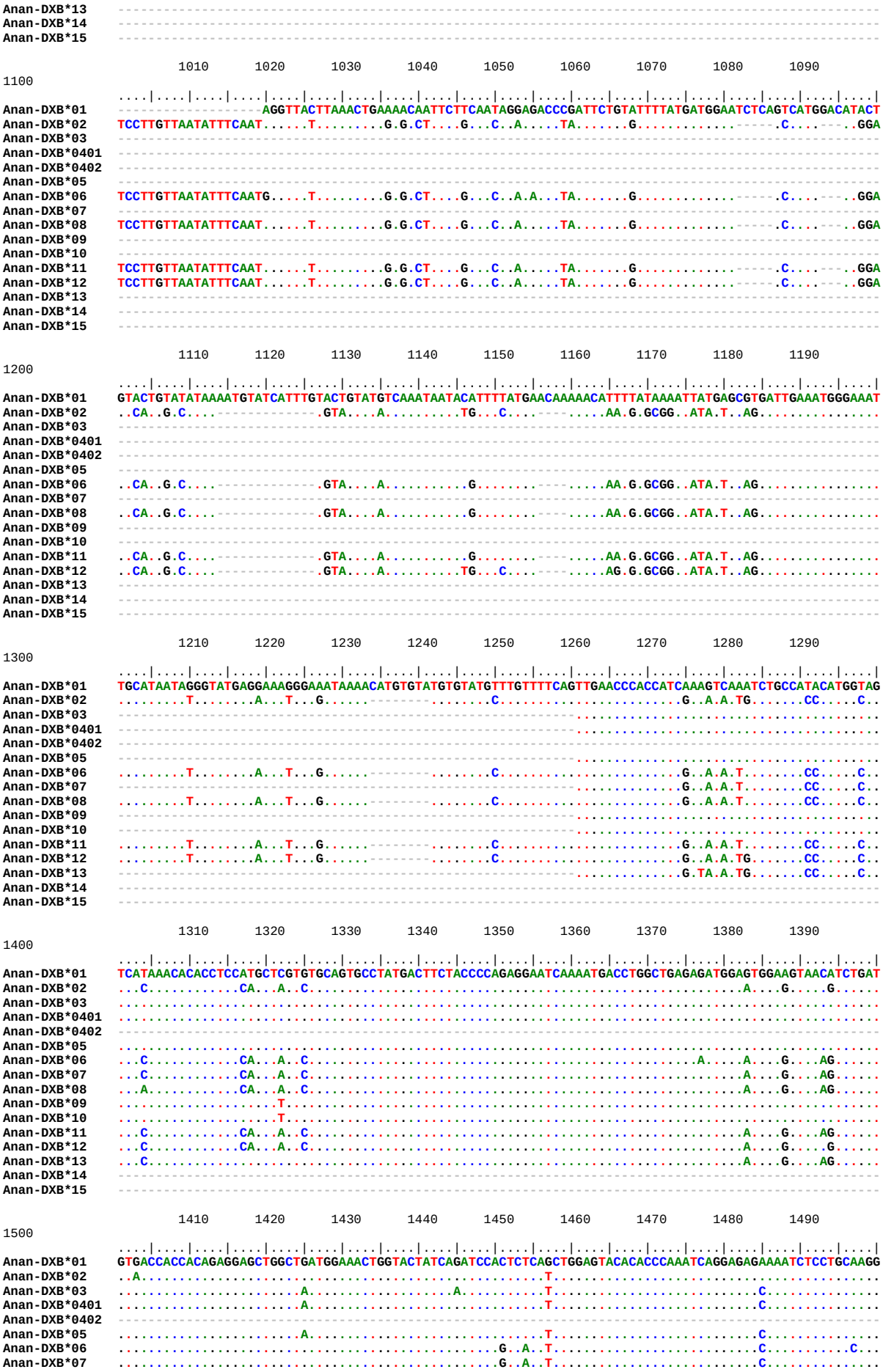


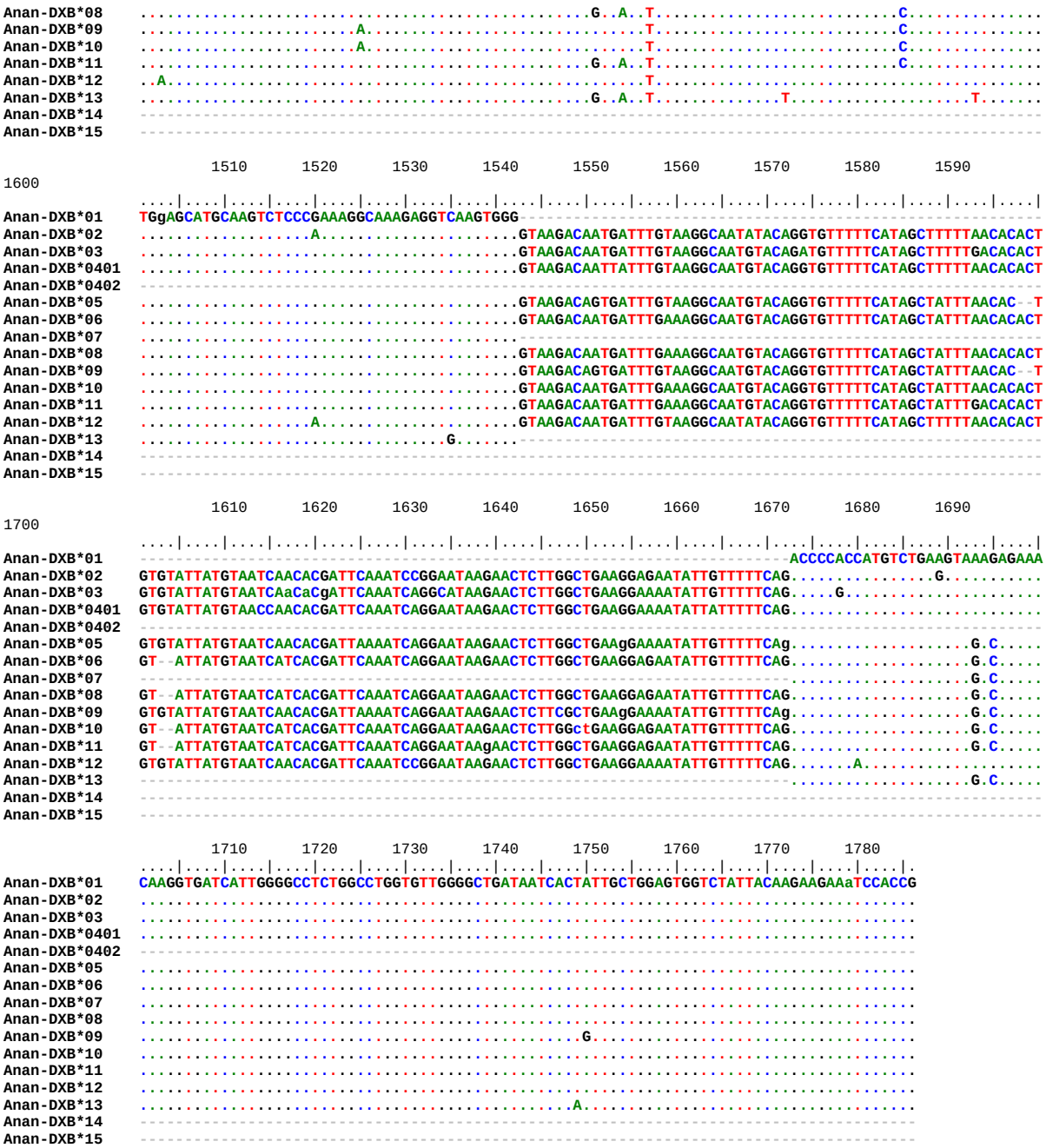

Anan-DXA*02
Anan-DXA*03
Anan-DXA*04
Anan-DXA*05
Anan-DXA*06
Anan-DXA*07
Anan-DXA*08
Anan-DXA*09
Anan-DXA*10
Anan-DXA*11
Anan-DXA*12
Anan-DXA*13
Anan-DXA*14
610 620 630 640 650 660 670 680 690 700
Anan-DXA*01 GAAGCTAACAATGAATCAGACATGGCAGAGACAGCATTCTGTGGAATAGGACTGACCCTGGGTCTGCTGGGAGTGGGAGTTGGAACTTTCTTCCTCATCA
Anan-DXA*02
Anan-DXA*03
Anan-DXA*04
Anan-DXA*05
Anan-DXA*06
Anan-DXA*07
Anan-DXA*08
Anan-DXA*09
Anan-DXA*10
Anan-DXA*11
Anan-DXA*12
Anan-DXA*13
Anan-DXA*14
710
Anan-DXA*01 AAGGAAATAACTGCAAC
Anan-DXA*02
Anan-DXA*03
Anan-DXA*04
Anan-DXA*05
Anan-DXA*06
Anan-DXA*07
Anan-DXA*08
Anan-DXA*09
Anan-DXA*10
Anan-DXA*11
Anan-DXA*12
Anan-DXA*13
Anan-DXA*14

Nucleotide sequences of MHC IIB alleles



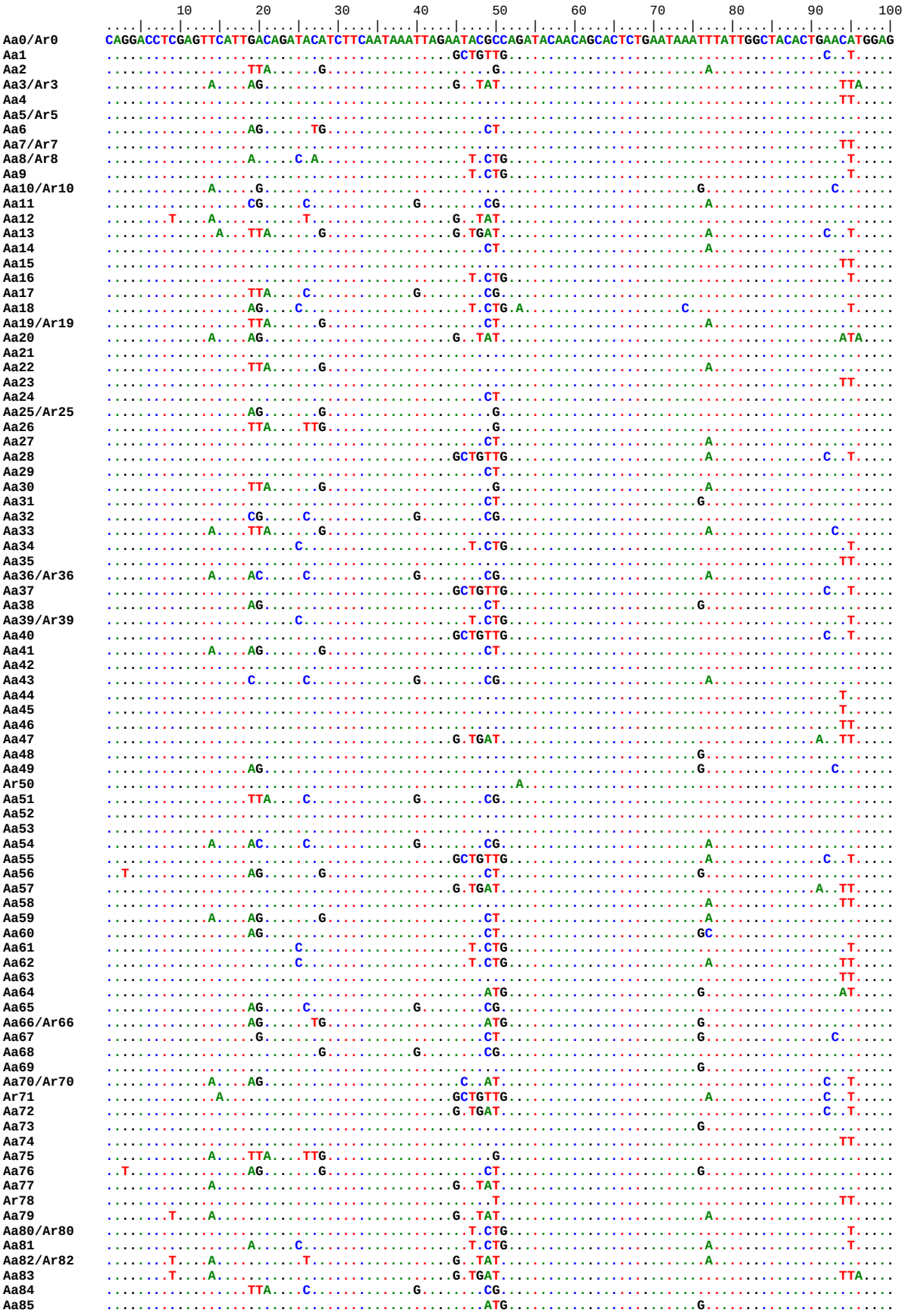

Anan-DXB*14
Anan-DXB*15
510 520 530 540 550 560 570 580 590 600
Anan-DXB*01 AGTGTTAAAGGCATGAAATAATGCATGTCTTTGAAGCAGACTATGATTAAGCATATGGACATGATGTACCATATTGAGTGTTTGTGAACCACTT
Anan-DXB*02 .....G.....C...A.....G...C.....
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06 .....G.....C...A.....A..C.....T.G...C.....
Anan-DXB*07
Anan-DXB*08 .....G.....C...A.....G...C.....A.AGGGAT
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11 .....G.....C...A.....A..C.....
Anan-DXB*12 .....G.....C...A.....A.....G...C.....
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
610 620 630 640 650 660 670 680 690 700
Anan-DXB*01 GAGGGAATGCATAAAAATACACAAAAATGAATTTTGTTACAAT TTTAAAAATGCCTCC AAGAATAAACTACAATCAGAATGTTGGGC ATATTA
Anan-DXB*02 ATA..G..AG.....CAG.....A.....G...AATGG.GTT.GT..TG.TTTGCTGG.GT..GGATGCATA..CCA.ACAAAGA.NG
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06 ATA..G..AG.....CAG.....A.....G...A TGG.GTT.GT..T TTTGCTGG.GT.TGGATGCATA..CCA.ACAAAGA.G
Anan-DXB*07
Anan-DXB*08 AG.TA..G..AG.....CAG.....A.....G...AATGG.GTT.GT..T TTTGCTGG.GT..GGATGCATA..CCA.ACAAAGA.GG
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11 .TA.AG..AG.....CAG.....A.....G...A TGG.GTT.GT..T TTTGCTGG.GT.TGGATGCATA..CCA.ACAAAGA.NG
Anan-DXB*12 ATA..G..AG.....CAG.....A.....G...AATGG.GTT.GT..TG.TTTGCTGG.GT..GGATGCATA..CCA.ACAAAGA.GG
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
710 720 730 740 750 760 770 780 790 800
Anan-DXB*01 GTTGCATTGCAGG AGTGCAGTAAATACGTACATGTAAATACATGTACAACAT
Anan-DXB*02 C.A...ACCTGTATTA.ATGTA.....T.....A....G.T..C....T..TACATTATTTCCCATCCCTATGATCAACAAATAATTCC TTCACCTG
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06 ..ACCTGTATTA.ATGTA.....T.....AN...G....C....T..TACATTATTTCCCATCCCTATGGTCAACAAATA ATTTCTTCACCTG
Anan-DXB*07
Anan-DXB*08 C.A...ACCTGTATTA.ATGTA.....T.....A....G.T..C....T..TACATTATTTCCCATCCCTATGATCAACAAATA ATTCC
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11 C.A...ACCTGTATTA.ATGTA.....T.....G....G....C....T..TACATTATTTCCCATCCCTATGATCAACAAATA ATTTCTTCACCTG
Anan-DXB*12 C.A...ACCTGTATTA.ATGTA.....T.....A....G.T..C....T..TACATTATTTCCCATCCCTATGATCAACAAATA ATTCCTTCACCTG
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
810 820 830 840 850 860 870 880 890 900
Anan-DXB*01
Anan-DXB*02 TCGAAGATTGATCTTTAAGACTTGATGTTTAATTTGACTGTGCAGCTTTATTAAAACGTTTTCCACTGCAAGGTCTCTATAGTTATAGGTATTACAACCA
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06 TTGAATATTGATCTTTAAGACTTGATGTTTAATTTGACTGTGCAGCTTTATTAAAACTTTTTCCACTGCAAGGTTTCTATAGTTATAGGTATTACAACCA
Anan-DXB*07
Anan-DXB*08 GAAGATTGATCTTNAAGACTTGATGTTTAATTTGNCTGTGCAGCTTTATNAAAACTTTTTCCACTGCAAGGTCTCTGTAGTTATAGGTATTACNACCA
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11 TTGAATATTGATCTTTAAGACTTGATGTTTAATTTGACTGTGCAGCTTTATTAAAACTTTTNCCACTGCAAGGTTTCTATAGTTATAGGTATTACAACCA
Anan-DXB*12 TTGAAGATTGATCTTTAAGACTTGATGTTTAATTTGACTGTGCAGCTTTATTAAAACGTTTTNCACTGCAAGGTCTCTATAGTTATAGGTATTACAACCA
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
910 920 930 940 950 960 970 980 990
1000
Anan-DXB*01
Anan-DXB*02 TATCTGCAGATATGCATGATAGGAAATTCAATTTCAATACTGTGACACCTTCAGACCAGCCTAATACATGCCTTGTTCTTCTCCAGAAGAAAGACCCTTA
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06 TATTTGAGGTAATGCATGAGAGGAAATTCAATTTCAATACTATGAGACCTTCAGACCTGCNTAATACATGCCTTGTTCTTCTCCAGAAGAAAGGCCCTTA
Anan-DXB*07
Anan-DXB*08 TATTTGCAGATATGCATGATAGGAAATTCAATTTCAATACTGTGACACCTTCCGACCAGCCTAATACATGCCTTGTTCTTCTCCAGAAGAAAGGCCCCTA
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11 TATTTGAGGTAATGCATGAGAGGAAATTCAATTTCAATACTATGAGACCTTCAGACCAGCCTAATACATGCCTTGTTCTTCTCCAGAAGAAAGGCCCTTA
Anan-DXB*12 TATCTGCAGATATGCATGATAGGAAATTCAATTTCAATACTGTGACACCTTCAGACCAGCCTAATACATGCCTTGTTCTTCTCCAGAAGAAAGACCCTTA

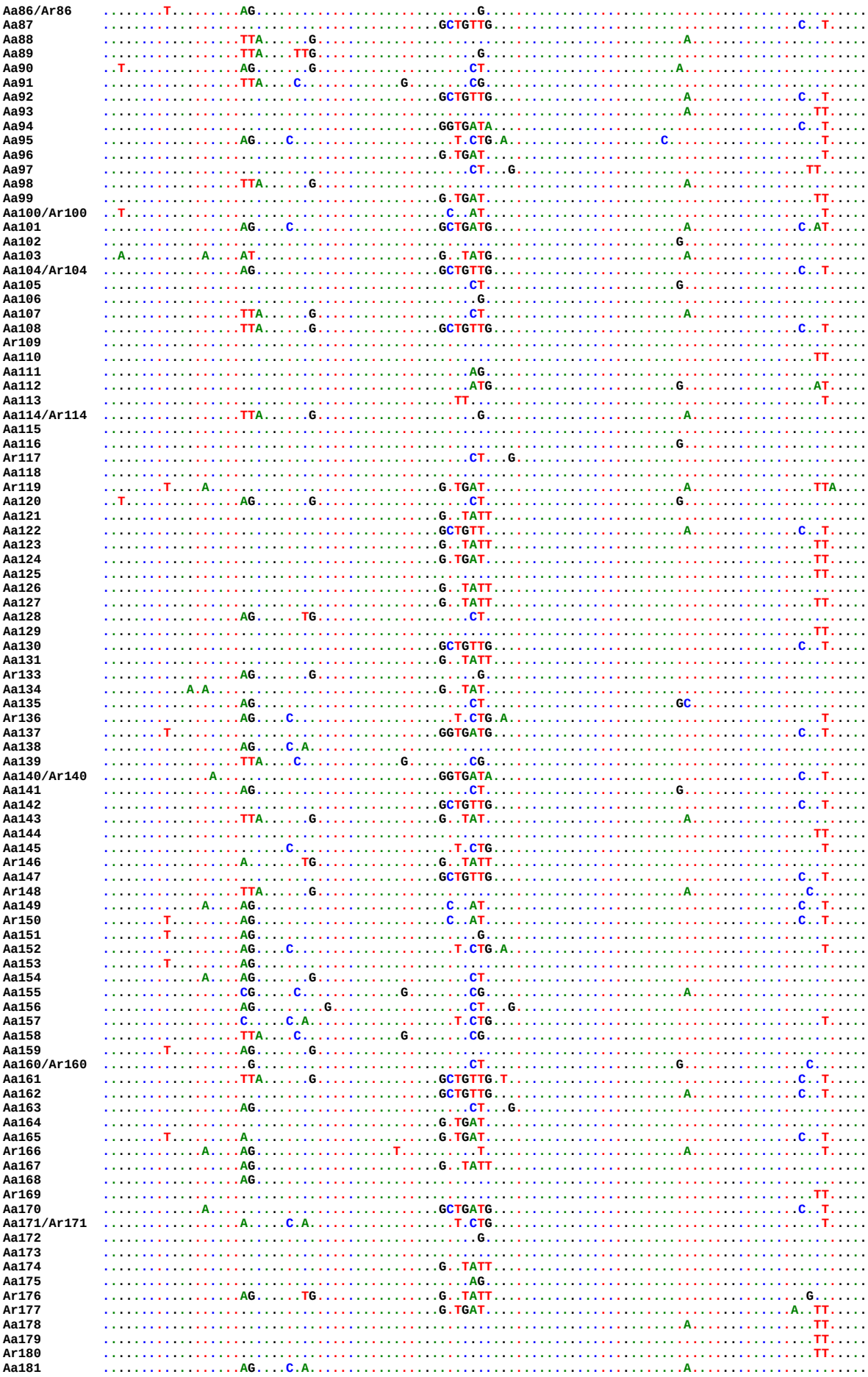

Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
Anan-DXB*01 AGGTTACTTAAACTGAAAACAATTCTTCAATAGGAGACCCGATTCTGTATTTTATGATGGAATCTCAGTCATGGACATACT
Anan-DXB*02 TCCTTGTTAATATTTCAAT.....T.....G.G.CT....G...C..A.....TA.....G.....C.....GGA
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06 TCCTTGTTAATATTTCAATG....T.....G.G.CT....G...C..A.A...TA.....G.....C.....GGA
Anan-DXB*07
Anan-DXB*08 TCCTTGTTAATATTTCAAT.....T.....G.G.CT....G...C..A.....TA.....G.....C.....GGA
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11 TCCTTGTTAATATTTCAAT.....T.....G.G.CT....G...C..A.....TA.....G.....C.....GGA
Anan-DXB*12 TCCTTGTTAATATTTCAAT.....T.....G.G.CT....G...C..A.....TA.....G.....C.....GGA
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
Anan-DXB*01 GTACTGTATATAAAATGTATCATTTGTACTGTATGTCAAATAATACATTTTATGAACAAAAACATTTTATAAAATTATGAGCGTGATTGAAATGGGAAAT
Anan-DXB*02 ..CA..G.C....GTA....A.....TG...C.....AA.G.GCGG..ATA.T..AG.....
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06 ..CA..G.C....GTA....A.....G.....AA.G.GCGG..ATA.T..AG.....
Anan-DXB*07
Anan-DXB*08 ..CA..G.C....GTA....A.....G.....AA.G.GCGG..ATA.T..AG.....
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11 ..CA..G.C....GTA....A.....G.....AA.G.GCGG..ATA.T..AG.....
Anan-DXB*12 ..CA..G.C....GTA....A.....TG...C.....AG.G.GCGG..ATA.T..AG.....
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
Anan-DXB*01 TGCATAATAGGGTATGAGGAAAGGGAAATAAAACATGTGTATGTGTATGTTTGTTTTCAGTTGAACCCACCATCAAAGTCAAATCTGCCATACATGGTAG
Anan-DXB*02 .....T.....A...T...G.....C.....G..A.A.TG.....CC.....C..
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06 .....T.....A...T...G.....C.....G..A.A.T.....CC.....C..
Anan-DXB*07
Anan-DXB*08 .....T.....A...T...G.....C.....G..A.A.T.....CC.....C..
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11 .....T.....A...T...G.....C.....G..A.A.T.....CC.....C..
Anan-DXB*12 .....T.....A...T...G.....C.....G..A.A.TG.....CC.....C..
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
Anan-DXB*01 TCATAAACACACCTCCATGCTCGTGTGCAGTGCCTATGACTTCTACCCCAGAGGAATCAAAATGACCTGGCTGAGAGATGGAGTGGAAGTAACATCTGAT
Anan-DXB*02 ...C.....CA...A..C.....A....G.....G.....
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06 ...C.....CA...A..C.....A....A....G....AG.....
Anan-DXB*07 ...C.....CA...A..C.....A....G....AG.....
Anan-DXB*08 ...A.....CA...A..C.....A....G....AG.....
Anan-DXB*09 .....T.....
Anan-DXB*10 .....T.....
Anan-DXB*11 ...C.....CA...A..C.....A....G....AG.....
Anan-DXB*12 ...C.....CA...A..C.....A....G.....G.....
Anan-DXB*13 ...C.....A....G....AG.....
Anan-DXB*14
Anan-DXB*15
1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
Anan-DXB*01 GTGACCACCACAGAGGAGCTGGCTGATGGAAACTGGTACTATCAGATCCACTCTCAGCTGGAGTACACACCCAAATCAGGAGAGAAAATCTCCTGCAAGG
Anan-DXB*02 ..A.....T.....
Anan-DXB*03 .....A.....A.....T.....C.....
Anan-DXB*0401 .....A.....T.....C.....
Anan-DXB*0402
Anan-DXB*05 .....A.....T.....C.....
Anan-DXB*06 .....G..A..T.....C.....C...
Anan-DXB*07 .....G..A..T.....C.....

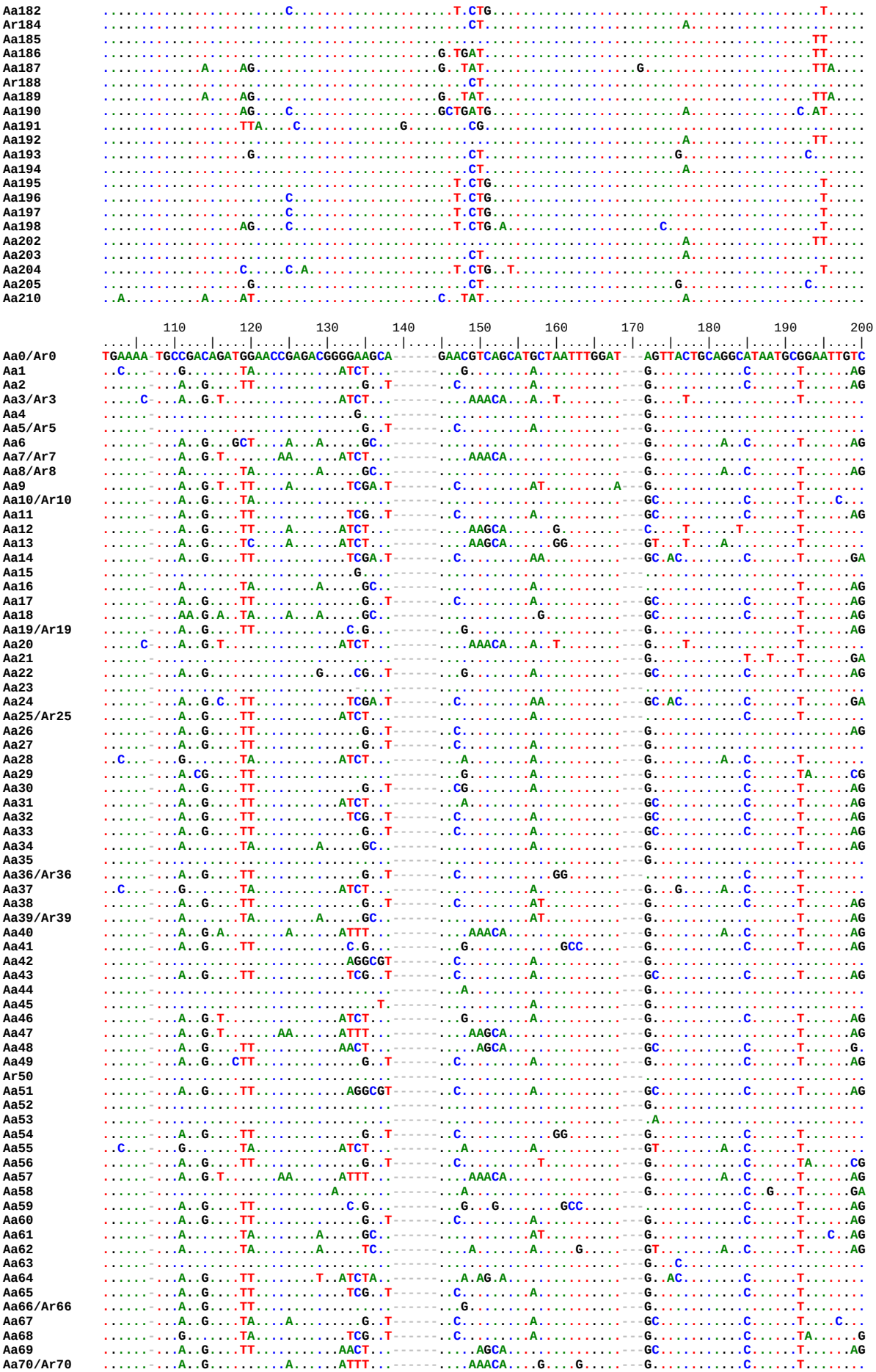

Anan-DXB*08
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11
Anan-DXB*12
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
Anan-DXB*01
Anan-DXB*02
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06
Anan-DXB*07
Anan-DXB*08
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11
Anan-DXB*12
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
1610 1620 1630 1640 1650 1660 1670 1680 1690 1700
Anan-DXB*01
Anan-DXB*02
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06
Anan-DXB*07
Anan-DXB*08
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11
Anan-DXB*12
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15
1710 1720 1730 1740 1750 1760 1770 1780
Anan-DXB*01
Anan-DXB*02
Anan-DXB*03
Anan-DXB*0401
Anan-DXB*0402
Anan-DXB*05
Anan-DXB*06
Anan-DXB*07
Anan-DXB*08
Anan-DXB*09
Anan-DXB*10
Anan-DXB*11
Anan-DXB*12
Anan-DXB*13
Anan-DXB*14
Anan-DXB*15

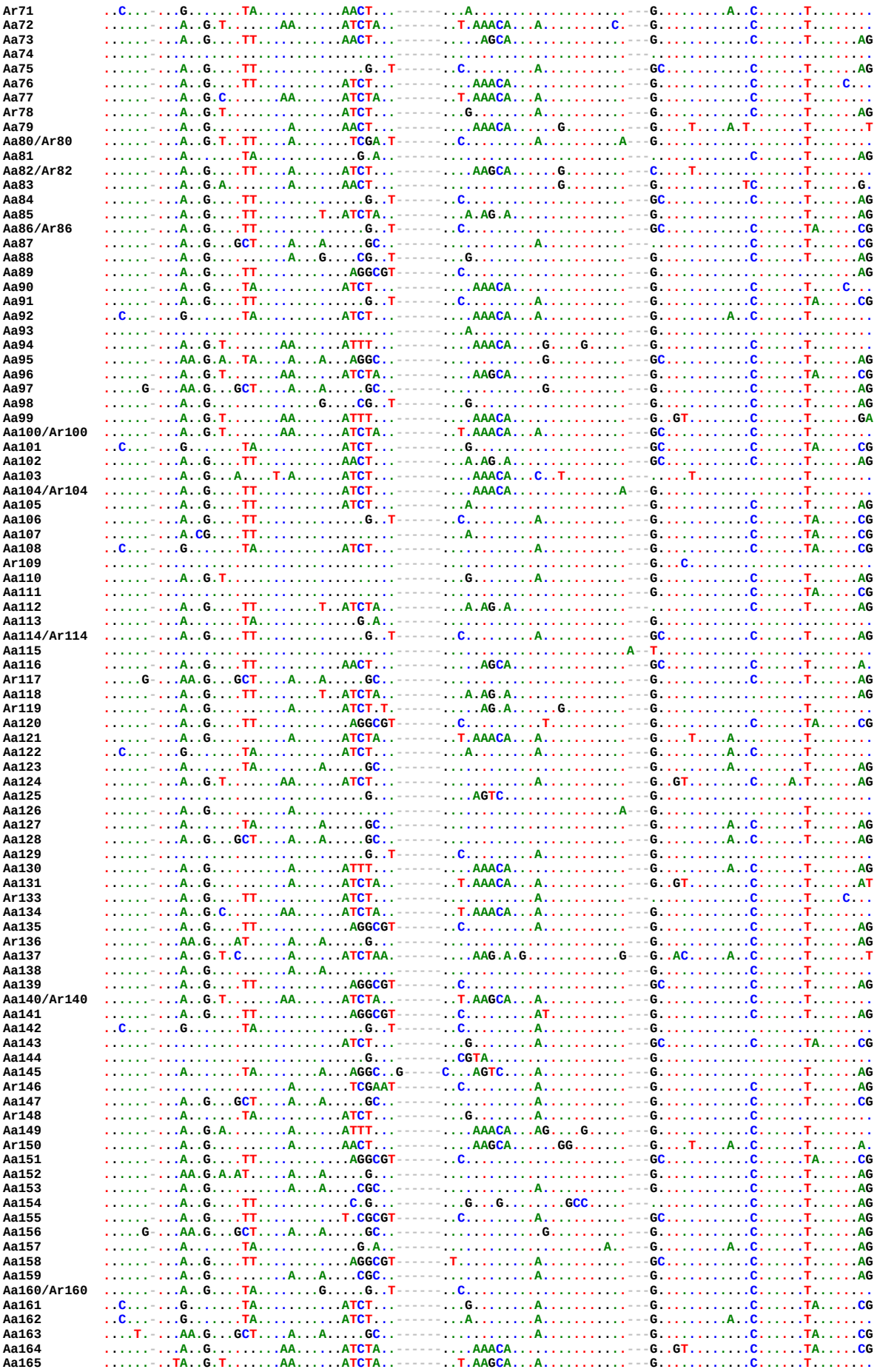
Appendix 2

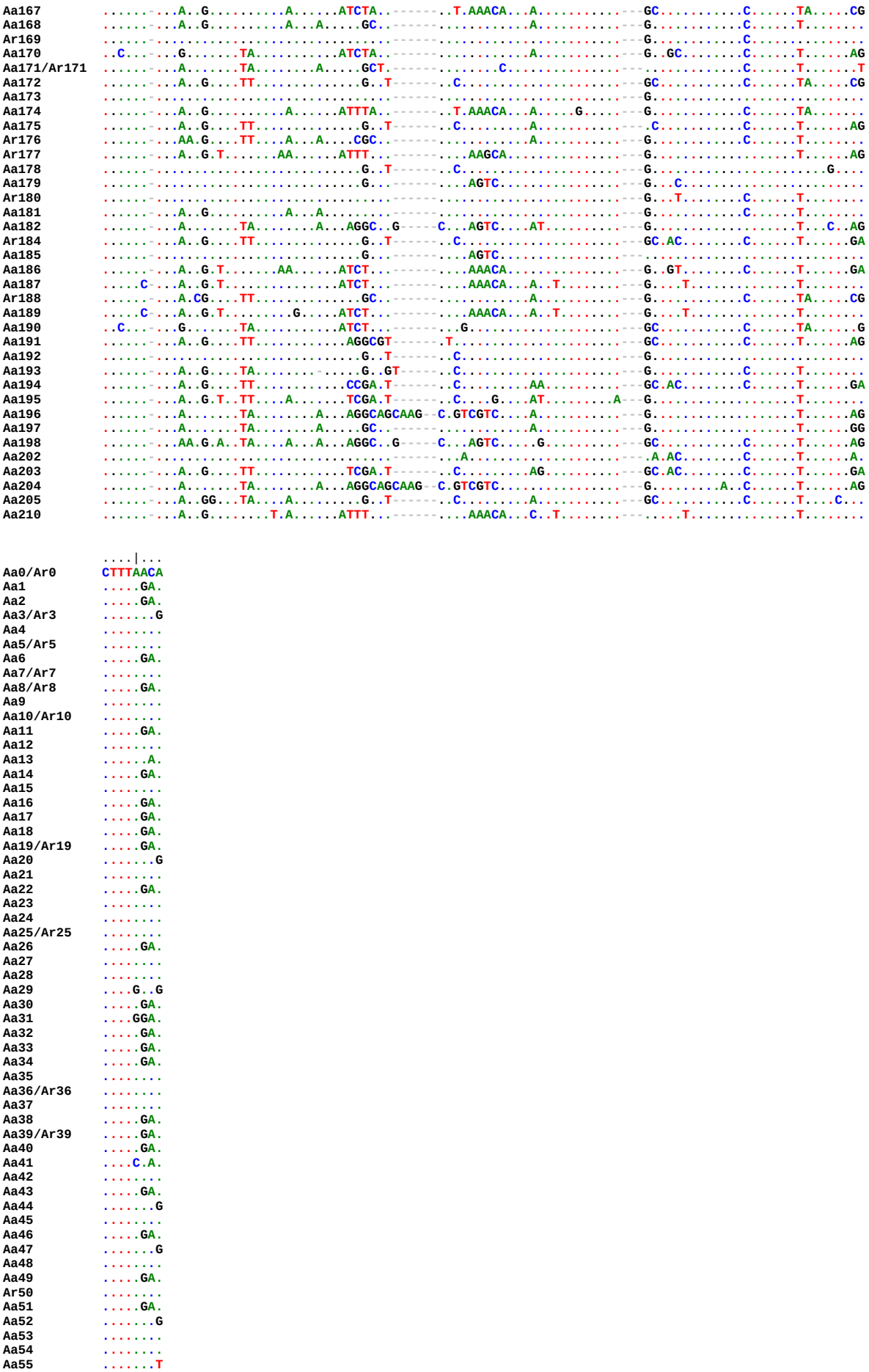
Nucleotide sequences MHC IIB exon 2




Aa86/Ar86
Aa87
Aa88
Aa89
Aa90
Aa91
Aa92
Aa93
Aa94
Aa95
Aa96
Aa97
Aa98
Aa99
Aa100/Ar100
Aa101
Aa102
Aa103
Aa104/Ar104
Aa105
Aa106
Aa107
Aa108
Ar109
Aa110
Aa111
Aa112
Aa113
Aa114/Ar114
Aa115
Aa116
Ar117
Aa118
Ar119
Aa120
Aa121
Aa122
Aa123
Aa124
Aa125
Aa126
Aa127
Aa128
Aa129
Aa130
Aa131
Ar133
Aa134
Aa135
Ar136
Aa137
Aa138
Aa139
Aa140/Ar140
Aa141
Aa142
Aa143
Aa144
Aa145
Ar146
Aa147
Ar148
Aa149
Ar150
Aa151
Aa152
Aa153
Aa154
Aa155
Aa156
Aa157
Aa158
Aa159
Aa160/Ar160
Aa161
Aa162
Aa163
Aa164
Aa165
Ar166
Aa167
Aa168
Ar169
Aa170
Aa171/Ar171
Aa172
Aa173
Aa174
Aa175
Ar176
Ar177
Aa178
Aa179
Ar180
Aa181


Aa182
Ar184
Aa185
Aa186
Aa187
Ar188
Aa189
Aa190
Aa191
Aa192
Aa193
Aa194
Aa195
Aa196
Aa197
Aa198
Aa202
Aa203
Aa204
Aa205
Aa210
110
120
130
140
150
160
170
180
190
200
Aa0/Ar0
Aa1
Aa2
Aa3/Ar3
Aa4
Aa5/Ar5
Aa6
Aa7/Ar7
Aa8/Ar8
Aa9
Aa10/Ar10
Aa11
Aa12
Aa13
Aa14
Aa15
Aa16
Aa17
Aa18
Aa19/Ar19
Aa20
Aa21
Aa22
Aa23
Aa24
Aa25/Ar25
Aa26
Aa27
Aa28
Aa29
Aa30
Aa31
Aa32
Aa33
Aa34
Aa35
Aa36/Ar36
Aa37
Aa38
Aa39/Ar39
Aa40
Aa41
Aa42
Aa43
Aa44
Aa45
Aa46
Aa47
Aa48
Aa49
Ar50
Aa51
Aa52
Aa53
Aa54
Aa55
Aa56
Aa57
Aa58
Aa59
Aa60
Aa61
Aa62
Aa63
Aa64
Aa65
Aa66/Ar66
Aa67
Aa68
Aa69
Aa70/Ar70


Ar71
Aa72
Aa73
Aa74
Aa75
Aa76
Aa77
Ar78
Aa79
Aa80/Ar80
Aa81
Aa82/Ar82
Aa83
Aa84
Aa85
Aa86/Ar86
Aa87
Aa88
Aa89
Aa90
Aa91
Aa92
Aa93
Aa94
Aa95
Aa96
Aa97
Aa98
Aa99
Aa100/Ar100
Aa101
Aa102
Aa103
Aa104/Ar104
Aa105
Aa106
Aa107
Aa108
Ar109
Aa110
Aa111
Aa112
Aa113
Aa114/Ar114
Aa115
Aa116
Ar117
Aa118
Ar119
Aa120
Aa121
Aa122
Aa123
Aa124
Aa125
Aa126
Aa127
Aa128
Aa129
Aa130
Aa131
Ar133
Aa134
Aa135
Ar136
Aa137
Aa138
Aa139
Aa140/Ar140
Aa141
Aa142
Aa143
Aa144
Aa145
Ar146
Aa147
Ar148
Aa149
Ar150
Aa151
Aa152
Aa153
Aa154
Aa155
Aa156
Aa157
Aa158
Aa159
Aa160/Ar160
Aa161
Aa162
Aa163
Aa164
Aa165
Ar166



```
Aa56          ....G..G
Aa57          .....GA.
Aa58          ........
Aa59          ....C.A.
Aa60          .....GA.
Aa61          .....GA.
Aa62          .....GA.
Aa63          ........
Aa64          ........
Aa65          ......A.
Aa66/Ar66     ........
Aa67          ........
Aa68          ....G..G
Aa69          .....GA.
Aa70/Ar70     ........
Ar71          ........
Aa72          ........
Aa73          .....GA.
Aa74          ........
Aa75          .....GA.
Aa76          ........
Aa77          ........
Ar78          .....G..
Aa79          ........
Aa80/Ar80     ........
Aa81          .....GA.
Aa82/Ar82     ........
Aa83          ........
Aa84          .....GA.
Aa85          .....GT.
Aa86/Ar86     ....G..G
Aa87          ....G..G
Aa88          .....GA.
Aa89          .....GA.
Aa90          ........
Aa91          ....G..G
Aa92          ......A.
Aa93          ........
Aa94          ........
Aa95          .....GA.
Aa96          ....G..G
Aa97          ........
Aa98          .....GA.
Aa99          ........
Aa100/Ar100   ........
Aa101         ....G..G
Aa102         .....GA.
Aa103         ........
Aa104/Ar104   ....G...
Aa105         ....GGA.
Aa106         ....G..G
Aa107         ....G..G
Aa108         ....G..G
Ar109         ........
Aa110         .....G..
Aa111         ....G..G
Aa112         .....GA.
Aa113         ........
Aa114/Ar114   .....GA.
Aa115         ........
Aa116         .....GA.
Ar117         ........
Aa118         .....GA.
Ar119         .......G
Aa120         ....G..G
Aa121         ........
Aa122         ........
Aa123         .....GA.
Aa124         .....G..
Aa125         ........
Aa126         ....G...
Aa127         .....GA.
Aa128         .C...GA.
Aa129         ........
Aa130         .....GA.
Aa131         .....GG.
Ar133         ........
Aa134         ....G...
Aa135         .....GA.
Ar136         .....GA.
Aa137         .....GA.
Aa138         ........
Aa139         .....GA.
Aa140/Ar140   ........
Aa141         .....GA.
Aa142         ........
Aa143         ....G..G
Aa144         ........
Aa145         .....GA.
Ar146         .....GA.
Aa147         ....G..G
Ar148         ........
Aa149         ........
Ar150         ........
Aa151         ....G..G
```

```
Aa152          .....GA.
Aa153          .....GA.
Aa154          ....C.A.
Aa155          .....GA.
Aa156          .......
Aa157          .....GA.
Aa158          .....GA.
Aa159          .....GA.
Aa160/Ar160    .......
Aa161          ....G..G
Aa162          ......A.
Aa163          ....G..G
Aa164          ....G..G
Aa165          .......
Ar166          .......
Aa167          ....G..G
Aa168          .......
Ar169          .......
Aa170          .....GA.
Aa171/Ar171    .....GA.
Aa172          ....G..G
Aa173          .......
Aa174          ....G...
Aa175          .....GA.
Ar176          .......
Ar177          .....GA.
Aa178          .......
Aa179          .......
Ar180          .......
Aa181          .......
Aa182          .....GA.
Ar184          .......
Aa185          .......
Aa186          .......
Aa187          ......G
Ar188          ....G..G
Aa189          ......G
Aa190          ....G..G
Aa191          .....GA.
Aa192          .......
Aa193          .......
Aa194          .....GA.
Aa195          .......
Aa196          .....GA.
Aa197          .....GA.
Aa198          .....GA.
Aa202          .......
Aa203          .....GA.
Aa204          .....GA.
Aa205          .......
Aa210          .......
```

Pairwise differentiation of MHC IIB pools among populations

	Adour 2010	Adour 2011	Adour 2012	Burrishoole glass eels	Bann Lower	Bann Toome	Burrishoole adult eels	Denmark	Finland	Glynn Lagoon	Larne Lagoon	Larne Lough	Portugal	Quoile	Boretree	Lough Comber	Sweden	Germany
Adour 2010		0.617	0.396	0.300	0.238	0.380	0.191	0.656	0.453	0.428	0.521	0.893	0.811	0.677	0.945	0.198	0.526	0.714
Adour 2011	-0.006		0.700	0.883	0.344	0.167	0.153	0.972	0.472	0.256	0.510	0.916	0.846	0.301	0.872	0.319	0.603	0.996
Adour 2012	0.001	-0.011		0.762	0.296	0.251	0.061	0.707	0.526	0.172	0.247	0.739	0.872	0.517	0.783	0.167	0.537	0.788
Burrishoole glass eels	0.011	-0.040	-0.026		0.123	0.060	0.194	0.243	0.248	0.055	0.092	0.302	0.568	0.095	0.302	0.051	0.098	0.990
Bann Lower	0.022	0.009	0.015	0.050		0.571	0.590	0.157	0.420	0.552	0.373	0.507	0.057	0.520	0.117	0.688	0.392	0.420
Bann Toome	0.005	0.024	0.014	0.051	-0.007		0.284	0.269	0.638	0.481	0.569	0.738	0.188	0.369	0.622	0.777	0.290	0.519
Burrishoole adult eels	0.018	0.021	0.035	0.020	-0.011	0.009		0.111	0.371	0.173	0.416	0.600	0.112	0.312	0.113	0.594	0.426	0.419
Denmark	-0.013	-0.045	-0.017	0.018	0.033	0.014	0.028		0.141	0.574	0.971	0.639	0.398	0.115	0.934	0.487	0.277	0.935
Finland	0.000	-0.001	-0.004	0.016	0.005	-0.012	0.003	0.026		0.526	0.597	0.364	0.729	0.646	0.403	0.546	0.245	0.663
Glynn Lagoon	0.002	0.012	0.019	0.057	-0.006	-0.001	0.020	-0.009	-0.002		0.378	0.420	0.287	0.205	0.542	0.807	0.389	0.548
Larne Lagoon	-0.005	-0.005	0.015	0.043	0.006	-0.009	0.001	-0.037	-0.009	0.004		0.998	0.615	0.375	0.891	0.542	0.556	0.559
Larne Lough	-0.043	-0.052	-0.027	0.011	-0.005	-0.023	-0.013	-0.020	0.002	0.002	-0.068		0.882	0.506	0.977	0.214	0.689	0.992
Portugal	-0.028	-0.033	-0.034	-0.013	0.065	0.022	0.032	0.001	-0.021	0.013	-0.013	-0.040		0.206	0.918	0.125	0.208	0.998
Quoile	-0.019	0.014	-0.008	0.052	-0.002	0.005	0.010	0.051	-0.017	0.028	0.004	-0.004	0.028		0.251	0.159	0.409	0.738
Boretree	-0.041	-0.033	-0.025	0.013	0.048	-0.013	0.030	-0.040	0.005	-0.005	-0.032	-0.060	-0.042	0.020		0.333	0.547	0.993
Lough Comber	0.023	0.014	0.031	0.070	-0.019	-0.023	-0.011	-0.002	-0.006	-0.028	-0.009	0.025	0.047	0.035	0.013		0.434	0.556
Sweden	-0.004	-0.009	-0.007	0.043	0.007	0.011	0.001	0.014	0.016	0.006	-0.007	-0.021	0.022	0.003	-0.009	0.002		0.728
Germany	-0.012	-0.041	-0.012	-0.065	0.004	-0.002	0.002	-0.036	-0.011	-0.004	-0.006	-0.068	-0.067	-0.021	-0.056	-0.005	-0.017	

R values of pairwise differentiation are given below the diagonal, corresponding p values above