

Phylogeographic Analysis of the Major Histocompatibility
Complex Class II b Alleles of the Trinidadian Guppy, *Poecilia
reticulata*

by

Brandon Fox

Submitted for a Combined Honours Degree in the faculty of
Environment Sustainability and Society and Biochemistry
and Molecular Biology

at

Dalhousie University Halifax, Nova Scotia

Copyright by Brandon Fox, 2016

Supervisor Dr. Paul Bentzen

DALHOUSIE UNIVERSITY

DATE: April 18th, 2016

TITLE: Phylogeographic Analysis of the Major Histocompatibility Complex Class
II b Alleles of the Trinidadian Guppy, *Poecilia reticulata*

AUTHOR: Brandon Fox

DEPARTMENT OR SCHOOL: College of Sustainability

DEGREE; Bachelor of Arts Convocation: May, 2016 Environment,

Sustainability and Society and Biochemistry and Molecular Biology

Permission is herewith granted to Dalhousie University to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions. I understand that my thesis will be electronically available to the public.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

The author attests that permission has been obtained for the use of any copyrighted material appearing in the thesis (other than the brief excerpts requiring only proper acknowledgement in scholarly writing), and that all such use is clearly acknowledged.

Brandon Fox Signature of Author

Table of Contents

List of Figures	IV
Abstract	VII
List of abbreviations used.....	VIII
Chapter 1	1
Introduction	1
The guppy, <i>Poecilia reticulata</i>	1
Major histocompatibility complex.....	2
Trinidad.....	3
Overview of thesis.....	6
Chapter 2	7
Literature Review	7
Chapter 3	9
Methodology.....	9
Samples	9
Molecular methods	11
Data analysis	11
Chapter 4	12
Results	12
Phylogenetic trees.....	12
GenGIS phylogeographic maps	29
Chapter 5	39
Discussion	39
Bibliography	41

List of Figures

Figure 1. The main watersheds of northern Trinidad, the Caroni in blue, Oropouche in orange and the Northern watershed in green.

Figure 2. Locations from southern Trinidad where samples were collected

Figure 3. Locations from Northern Trinidad where guppy samples were collected

Figure 4. Locations from Paria and Marianna of northern Trinidad where guppy samples were collected

Figure 5. Non Trinidadian locations where guppy samples were collected.

Figure 6. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST1 alleles, with bootstrap values.

Figure 7. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST2 alleles, with bootstrap values.

Figure 8. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST3 alleles, with bootstrap values.

Figure 9. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST4 alleles, with bootstrap values.

Figure 10. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST5 alleles, with bootstrap values.

Figure 11. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST6 alleles, with bootstrap values.

Figure 12. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST7 alleles, with bootstrap values.

Figure 13. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST8 alleles, with bootstrap values.

Figure 14. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST9 alleles, with bootstrap values.

Figure 15. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST10 alleles, with bootstrap values.

Figure 16. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST11 alleles, with bootstrap values.

Figure 17. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST12 alleles, with bootstrap values.

Figure 18. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST13 alleles, with bootstrap values.

Figure 19. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST14 alleles, with bootstrap values.

Figure 20. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST15 alleles, with bootstrap values.

Figure 21. Maximum likelihood phylogenetic tree of all guppy MHC class IIb alleles, with bootstrap values.

Figure 22. Geophylogeny of guppy MHC class IIb ST3 alleles, derived from maximum likelihood phylogenetic tree.

Figure 23. Geophylogeny of guppy MHC class IIb ST4 alleles, derived from maximum parsimony phylogenetic tree.

Figure 24. Geophylogeny of guppy MHC class IIb ST7 alleles, derived from maximum likelihood phylogenetic tree.

Figure 25. Geophylogeny of guppy MHC class IIb ST7 alleles, derived from maximum parsimony phylogenetic tree.

Figure 26. Geophylogeny of guppy MHC class IIb ST10 alleles, derived from maximum parsimony phylogenetic tree.

Figure 27. Geophylogeny of guppy MHC class IIb ST12 alleles, derived from maximum parsimony phylogenetic tree.

Figure 28. Geophylogeny of guppy MHC class IIb ST14 alleles, derived from maximum likelihood phylogenetic tree.

Figure 39. Geophylogeny of guppy MHC class IIb ST9 alleles, derived from maximum likelihood phylogenetic tree.

Abstract

The Major Histocompatibility Complex (MHC) are a set of vertebrate immune genes, that act in defense from cellular parasites. In *Poecilia reticulata* MHC is shown to also affect coloration and mate choice. The guppy acts as a model system in studying MHC. Having multiple copies of the MHC gene, large populations, and short generation time are ideal for studying MHC diversity. Samples were collected from Trinidad as Trinidad's water system offers an ideal living lab with many genetically distinct populations separated by watersheds, waterfalls, and different predation levels. MHC may be separated into groups, known as super types, based on like binding properties of the peptide binding region. Through the construction of maximum likelihood trees super types were compared to the phylogeny of all alleles together. These trees were then aligned with the geography of Trinidad to look for correlation between clades of the tree and clusters of sample sites. When looking at the full tree of MHC alleles, significant correlation between phylogenetic tree and sample sites were found, however only super types 3, 4, 7, 10, and 14 were found to have this correlation. I discuss the role of selection, both natural and sexual on the presence and absence of MHC alleles, the roll of the Guppy's colonization history in Trinidad, and the role of temporal separation due to the age of the MHC genes themselves.

List of abbreviations used

MHC	Major histocompatibility complex
PBR	Peptide binding region
ST	Super type

Chapter 1

Introduction

Genes of the Major Histocompatibility Complex (MHC) are found in all vertebrates and code for proteins that function in cellular defense against parasites (Herdegen, Babik, & Radwan, 2014).

This immune function allows MHC to be used as a proxy to estimate an organism's or a population's ability to adapt to the local environment (Lighten, 2015). In some species, such as the Trinidadian Guppy, *Poecilia reticulata*, variation in MHC genes has been shown to be correlated with various traits, including those involved in sexual selection. The interactions between parasite dependent natural selection and mate choice dependent sexual selection both acting on the MHC genes are difficult to model. In this thesis I hope to aid in the modeling of MHC selection by looking at relationship of MHC protein binding region phylogeny and the geography of Trinidad.

The guppy, *Poecilia reticulata*

Poecilia reticulata, the Trinidadian guppy, is a hardy species of freshwater fish indigenous to Central America and northern South America. It is known to be able to survive in brackish and polluted waters (Baillie, 2012; Lighten, 2015). This hardiness is attributed in part to its short generation times and high level of genetic variability (Lighten, 2015). The guppy is a sexually selective species, with the females preferring ornate males (Nicoletto, 1993). Ornaments of the male guppy include coloured spots, especially orange and black, structural pigments, and elongated fins. The coloured spots of the guppy require pigments that it is unable to produce itself and must acquire through dietary intake (Grether, Hudon, & Millie, 1999). Ornaments are thought to be honest indicators correlated to various measures of fitness in the guppy, as

pigmentation requires resources that are also limiting in immune function (Grether, Hudon, & Millie, 1999; Lighten, 2015; Nicoletto, 1993). Correlation between ornament and MHC allele diversity has also been reported (Lighten, 2015).

Major histocompatibility complex

Genes of the MHC are present in all jawed vertebrates, and code for proteins involved in the immune response (Herdegen, Babik, & Radwan, 2014). MHC proteins are involved in antigen presenting, binding to pathogen-derived antigens to elicit an immune response (Herdegen, Babik, & Radwan, 2014). The MHC has two distinct classes: class I, which binds to intracellular pathogens antigens, and class II, binding to extracellular pathogen antigens (Fraser, Ramnarine, & Neff, 2010; Herdegen, Babik, & Radwan, 2014). The MHC genes are the most polymorphic genes found in vertebrates, and the Red Queen Hypothesis is thought to account for this polymorphism found in MHC genes. (Herdegen, Babik, & Radwan, 2014; Lighten, 2015). The Red Queen hypothesis describes the co-evolutionary trends between host and pathogen relations, when one evolves an adaptation against the other, the other than must evolve a counter adaptation to survive (Van Valen, 1973). The large amount of variability in the MHC can be attributed to the fast rate of evolution of pathogens (Fraser, Ramnarine, & Neff, 2010; Lighten, 2015; Herdegen, Babik, & Radwan, 2014). With the MHC needing to be able to respond to multiple types of variable pathogens, many species, the guppy included, exhibit multiple copies of the genes, and the number of copies can vary among individuals (copy number variation; CNV) (Sommer, 2005; Herdegen, Babik, & Radwan, 2014; Lighten, 2015). Individuals may be heterozygous at any MHC locus, and this may contribute to an increase in fitness due to heterozygous advantage or negative frequency dependent selection (NDS) (Lighten, 2015; Herdegen, Babik, & Radwan, 2014; Fraser, Kunstner, Reznick, Dreyer, & Weigel, 2015).

Heterozygous advantage results from heterozygotes, having different MHC alleles, being able to respond to a broader range of pathogens than homozygotes, having only one allele (Fraser, Ramnarine, & Neff, 2010; Lighten, 2015; Herdegen, Babik, & Radwan, 2014). NDS may benefit individuals with rare MHC alleles in the population; due to the allele being rare there is little pressure for pathogens to evolve an adaptation against the allele compared to common alleles (Fraser, Ramnarine, & Neff, 2010; Herdegen, Babik, & Radwan, 2014; Lighten, 2015). Conversely MHC diversity may also be selected against in populations with exposure to only a few species of pathogens, reducing the benefits of heterozygosity and diverse MHC alleles in the population (Fraser, Ramnarine, & Neff, 2010; Lighten, 2015). Lighten (2015) found that the guppy may carry 1-6 MHC class IIb loci. In describing MHC phenotypes it has been proposed to group alleles into super types (ST), based on like antigen binding in by the peptide-binding region (PBR) of the gene (Lighten, 2015; Reche & Reinherz, 2007). While grouping into STs has been shown to be an effective way of categorizing human MHC alleles, it is not yet confirmed for non-human species.

Trinidad

The river systems of Trinidad offer a variety of different local environments, often with physical barriers such as waterfalls, between (Herdegen, Babik, & Radwan, 2014; Baillie, 2012; Barson, Cable, & Van Oosterhout, 2009; Barson, Cable, & Van Oosterhout, 2009). Previous studies have shown that the changes in the local environment have contributed to changes in guppy behaviour, male colour, reproduction cycles, morphology, and neutral and selective genetic markers between populations (Magurran, 2005). Behavioural, morphological, and reproductive differences have largely been attributed to differences in predation between local environments (Magurran, 2005). Large predatory fish that feed on adult *P. reticulata* are found in downstream

areas but are unable to pass barriers such as waterfalls to get upstream; in response, downstream guppies tend to be small, shoal tightly together, and invest in early reproduction with large clutch sizes, as they are unlikely to survive long in the predator filled waters (Magurran, 2005; Reznick, Butler IV, & Rodd, 2001). Upstream predators are closer in size to *P. reticulata* and feed mostly on the young, which selects for larger guppies that invest in long term reproductive success with smaller clutch sizes of large young (Magurran, 2005 (Reznick, Butler IV, & Rodd, 2001).

Colouration of male *P. reticulata* was thought to result from a combination of natural selection, predators being able to spot bright guppies more easily, and sexual selection, females preferring bright males (Reznick et al, 2001). Lighten (2015) has suggested the male *P. reticulata* colouration is linked with variation at the Major Histocompatibility Complex (MHC), as an honest indicator of immuno-competence. The northern part of Trinidad has three major watersheds, the Caroni in the west, the Oropouche to the east, and the north shore (Baillie, 2012). Baillie (2012) has shown, via mitochondrial DNA (mtDNA) samples and microsatellites, that the populations the Caroni and Oropouche are genetically distinct from each other and likely populated the island in two different events, from two different source populations. Some exceptions can be seen in these watersheds to the origins such as the population in the Turure.

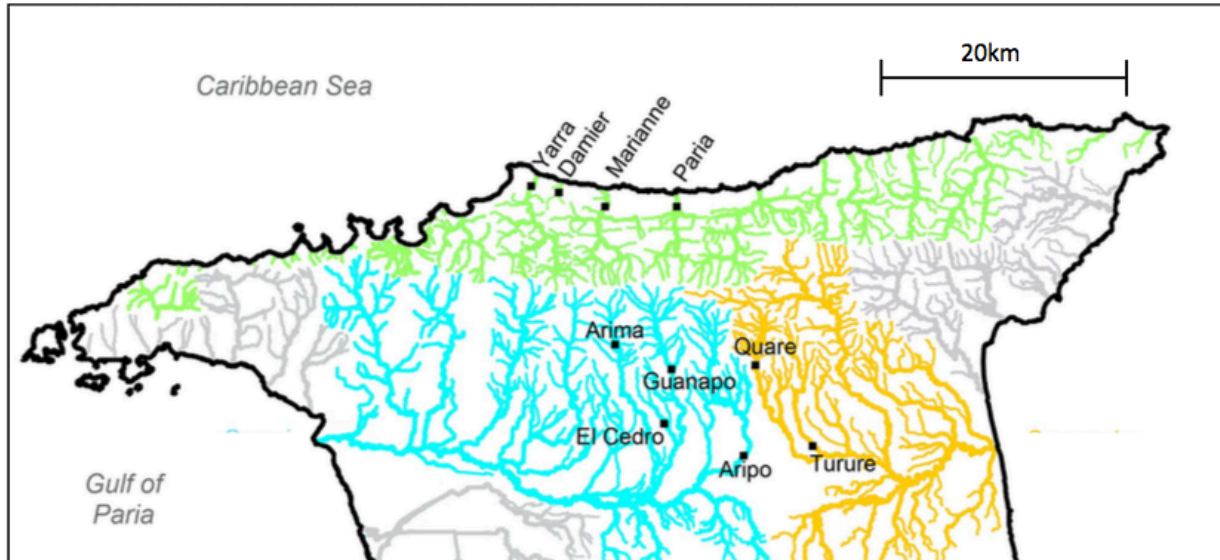


Figure 1. The main watersheds of northern Trinidad, the Caroni in blue, Oropouche in orange and the Northern watershed in green.

The MHC is of particular interest in conservation of vertebrates, and in agricultural and aquaculture settings. While MHC is primarily studied for its immune function and disease resistance, it also plays a role in mate choice, success of fertilization, kin selection and autoimmune disease, with populations with a diverse MHC gene pool generally being more successful (Sommer, 2005). Human activity often fragments and reduces the total habitat available for wild populations, reducing total population size and preventing gene flow between populations (Sommer, 2005). Reduction in population size, and gene flow may lead to decreased diversity in all genes, including MHC, due to increased risk of genetic drift and inbreeding (Sommer, 2005). Diversity in MHC alleles is positively correlated with recovery of bottlenecked populations and decreasing disease susceptibility (Sommer, 2005). The decrease in disease susceptibility also holds true in agricultural and aquaculture populations, as seen in cattle and salmon, with high levels of MHC diversity reducing disease prevalence, and need for antibiotics

and other treatments (Sommer, 2005). The guppy and the watersheds of Trinidad offer a model system, acting as a natural laboratory, for looking at the change of MHC diversity in populations, and its role in disease resistance, behaviour, and reproductive success (Sommer, 2005; Herdegen, Babik, & Radwan, 2014). The structure of Trinidad's watersheds offers sites that may experience non-negligible migration between populations, but are still genetically, morphologically, and behaviourally distinct due to different selective pressures in downstream sites and between different rivers or watersheds (Herdegen, Babik, & Radwan, 2014; Sommer, 2005). This project hopes to add the modeling of MHC diversity in a natural system, and in that aid in conservation, population management efforts.

Overview of thesis

This study uses the MHC class II b PBR sequence data and genotyping from Lighten (2015).

Lighten (2015) has organized PBR alleles into ST groups based on the targets parasite, with PBR's that bind to similar parasites being grouped together. Allelic presence of MHC class IIb region have been linked to the location of the population and mapped out using GIS to infer gene flow. Maximum likelihood and maximum parsimony phylogenetic trees have been constructed from the sequence data of alleles found in Lighten (2015). The phylogenetic trees were linked to location to observe how allelic presence changes in terms of ST.

This thesis focuses on allelic level distribution in contrast to Lighten (2015) who looked only at the MHC supertype phylogeographic distribution. Lighten's study was chosen due to the large amount of population and sequence data for guppy MHC available from it. This study hopes to contribute to a working model on MHC evolution and distribution between populations for the use in future studies.

Chapter 2

Literature Review

Fraser (et al, 2010) was one of the first studies to look at the MHC in natural populations of guppies, and showed that MHC is under strong selection compared to neutral markers. Lighten's (2015) study focused on mapping groups of functionally similar MHC phenotypes, based on the protein binding region, as supertypes (15 in all for *P. reticulata*). Parasite load of the guppies was then compared to the supertypes found in the population. Lighten (2015) found several associations between STs and parasite prevalence, i.e. ST 2, ST 11, and *Gyrodactylus*, a common parasite of *P. reticulata*. Lighten's study does have a drawback in that the grouping of MHC alleles into supertypes for *P. reticulata* is only a theory based on modeling of human MHC diversity. Supertyping in humans examines the peptide-binding region of the MHC (Reche, 2010). The peptide-binding region is important for the immune function of the MHC; it binds to an antigen marker and begins to illicit an immune response (Reche, 2010). In humans supertyping is determined grouping together peptide binding regions that bind to like antigens (Reche, 2010). The method of super typing does not take into account the genetic descent of alleles, only phenotypic similarity, and may group together genetically distinct alleles (Reche, 2010). Lighten (2010) used a guppy specific peptide binding region, which has yet to be confirmed if it follows the same pattern as human peptide binding region. The methods used in sequencing the MHC and interpreting the data are quite novel (Babik, Taberlet, Ejsmond, & Radwan, 2009; Lighten, 2015).

The use of Next Generation Sequencing (NGS), while considered more practical than previously used methods, such as Sanger sequencing, for gathering large amounts of sequence data, is not without drawbacks (Babik, Taberlet, Ejsmond, & Radwan, 2009; Lighten, 2015).

NGS is prone to pick up on small errors that may be made in the processing of the sequence, i.e. error in DNA synthesis during PCR that may give null alleles or false positives (Babik, Taberlet, Ejsmond, & Radwan, 2009; Lighten, 2015). To compensate for PCR related errors many studies placed high restrictions on what they allow to be considered an allele, i.e. having to show up on x % of reads, which likely lead to false negatives in the counting of alleles (Lighten, 2015). The Ultra Deep Illumina sequencing used by Lighten (2015) looks to more accurately sequence by increasing the reads of the sequence.

Chapter 3

Methodology

Samples

P. reticulata samples were taken between 2008 and 2012. 59 populations were sampled from throughout Trinidad. Samples were euthanized and preserved in ethanol (Lighten, 2015).

Figures 2-5 show the sites where guppy samples were obtained from.

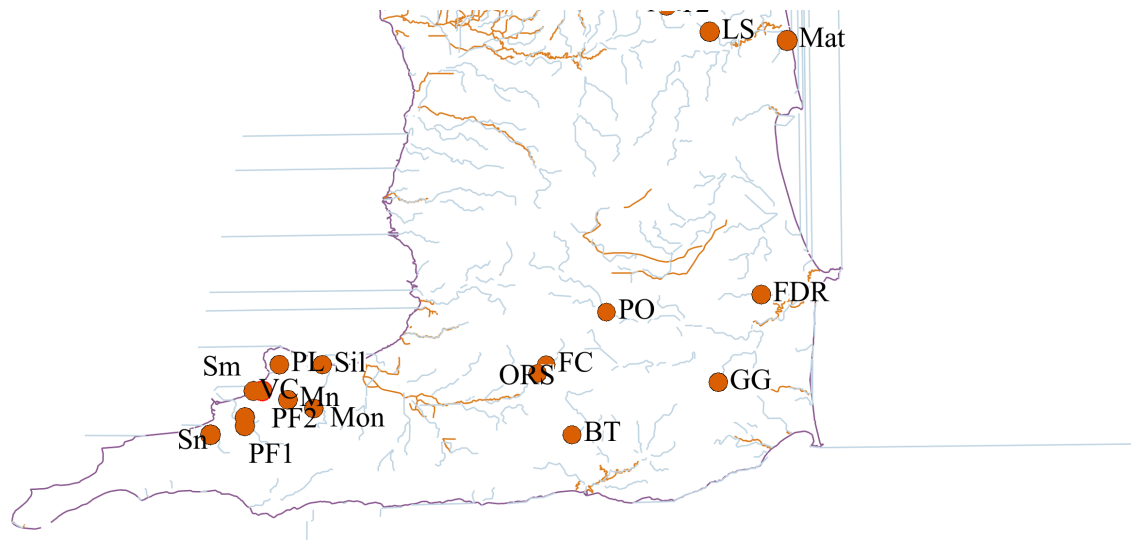


Figure 2. Locations in southern Trinidad where guppy samples were taken. Abbreviated locations: Snake River (Sn), Stollmeyer (Sm), Vance (Vc), Morne (Mn), Mondesir (Mon), Pitch Lake (PL), Fifth Company (FC), Point Fortin (PF), Silver Stream (Sil), Oropouche South (ORS), Guayguare (GG), Basse-Terre (BT), Poole (PO), Faecal Dump Rd. (FDR), Matura (MAT), and Las Seiva (LS).

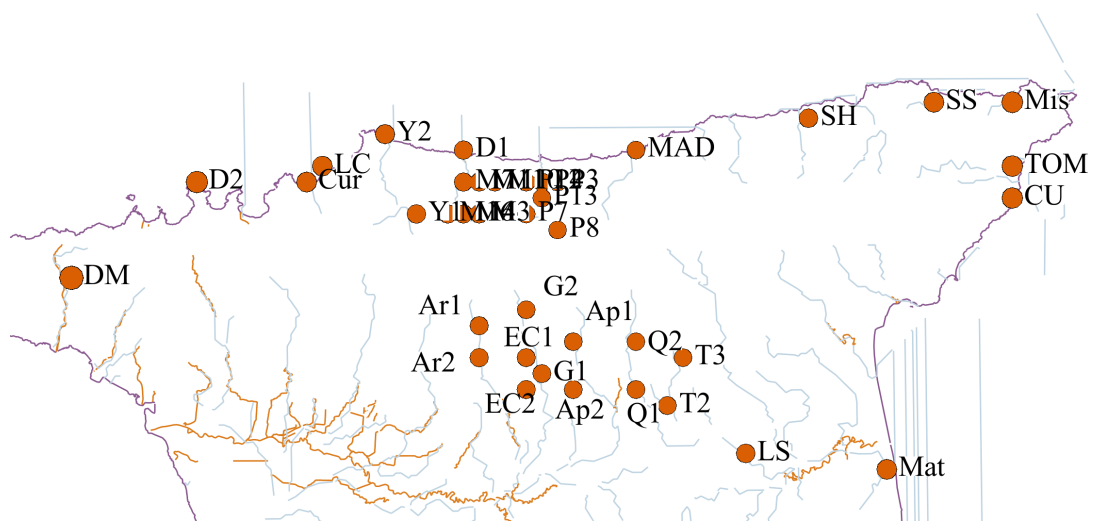


Figure 3. Locations in Northern Trinidad where guppy samples were collected. Abbreviated locations; Diego Martin (DM), Damier (D), Curaguata (Cur), Las Cuevas (LC), Yara (Y), Madamas (MAD), Turure (T), Arima (AR), San Souci (SS), Shark (SH), Mission (Mis), Cumana (CU), Tompire (TOM), Guanapo (G), Aripo (Ap), Quare (Q), and El Cedro.

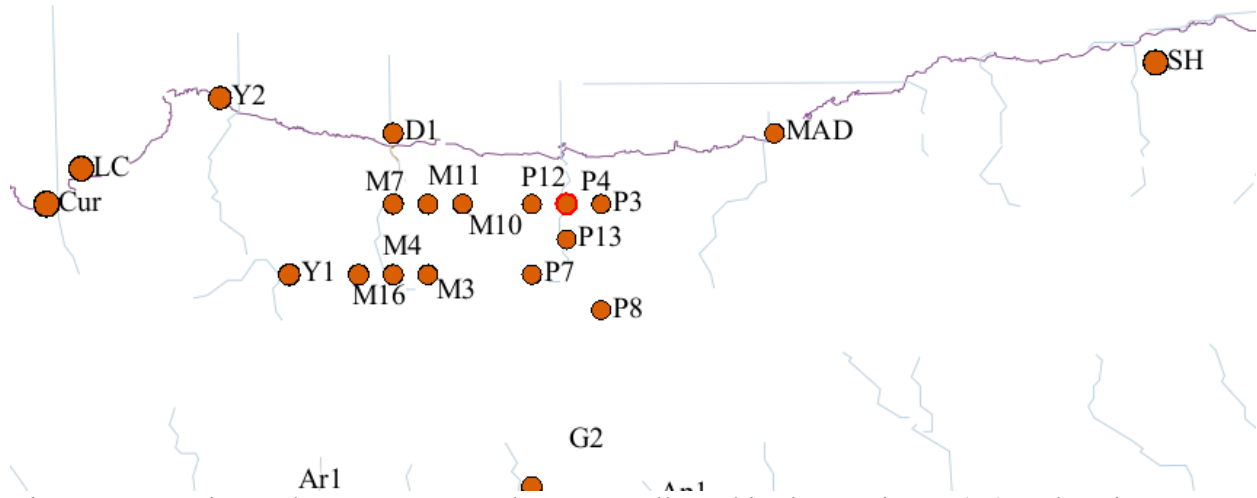


Figure 4. Locations where guppy samples were collected in the Marianne (M) and Paria (P) River systems of northern Trinidad.

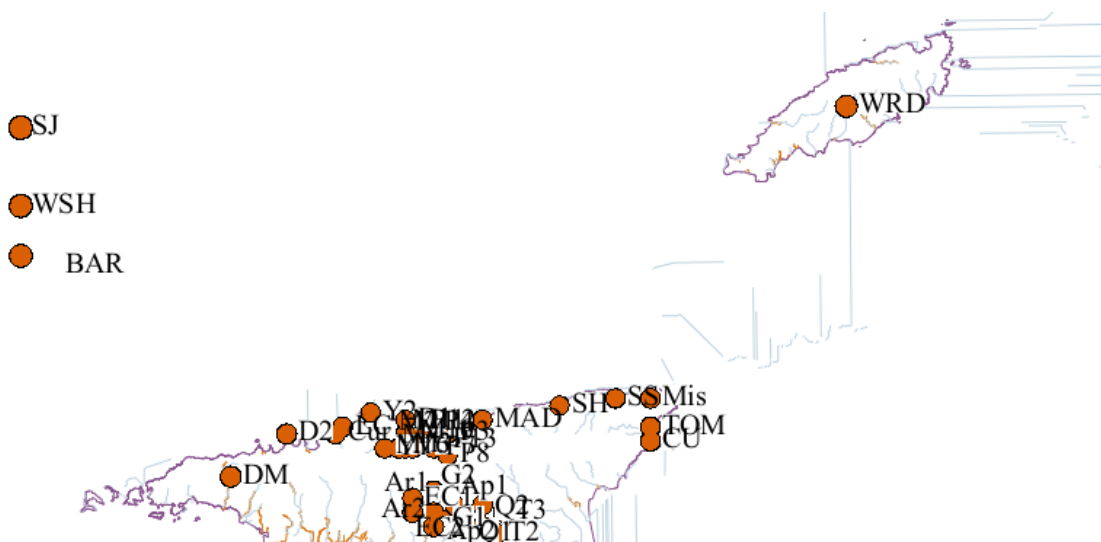


Figure 5. Non Trinidadian locations where samples, or genetic data were available. Locations abbreviated; Windward, Tobago (WRD), Barbados (BAR), and Hawaii (WSH). Samples of the related swamp guppy, *Poecilia picta*, were also collected from St. Joseph (SJ), Trinidad.

Molecular methods

A 209 base-pair portion of the peptide-binding region of the MHC II b was amplified via PCR, in either a Mastercycler Eppgradient S or a ep384 thermocycler in the following cycles: 98 °C for 3 min; 30 cycles of 98 °C 15 s, 57 °C 40 s, 72 °C 60 s; 10 min at 72 °C, then held at 10 °C (Lighten, 2015). Samples were then sequenced with an Illumina MiSeq sequencer. Individuals were then genotyped, and validated by Lighten (2015).

Data analysis

CNV in the MHC prevents the assignment of individual alleles to loci, to compensate for this all alleles are looked at as though they are dominant, if they are present in an individual they are expressed (Herdegen, Babik, & Radwan, 2014). Allelic nucleotide sequences were used to construct maximum likelihood and parsimony phylogenetic trees using the phangorn package on R, and bootstrapped via the boot package (Schliep, 2011; Canty, 2002). Trees were constructed for all 15 supertypes, as well as a tree containing all alleles used in Lighten's (2015) study. Trees were edited with FIGTREE (www.tree.bio.ed.ac.uk/software/figtree/). Tree structure was then

compared to the geography of Trinidad, comparing allelic groups to sample sites where they are found using GIS program GenGIS. Lighten's (2015) also included MHC alleles from populations in Tobago, Barbados, Hawaii, and from a Trinidadian population the related swamp guppy *Poecilia picta* for comparison, these were retained in this study.

Chapter 4

Results

Phylogenetic trees

Maximum likelihood trees were constructed for the 15 supertypes defined by Lighten (2015) (Figures 6-20). An additional maximum likelihood tree was constructed for all alleles together (Figure 21). Bootstrap values were calculated and reported for all trees. Maximum parsimony trees were also constructed, but found to have smaller bootstrap values than the maximum likelihood trees; they are shown in the appendix. Branches on the maximum parsimony trees with high bootstrap values frequently conformed to the structures found in the maximum likelihood trees.

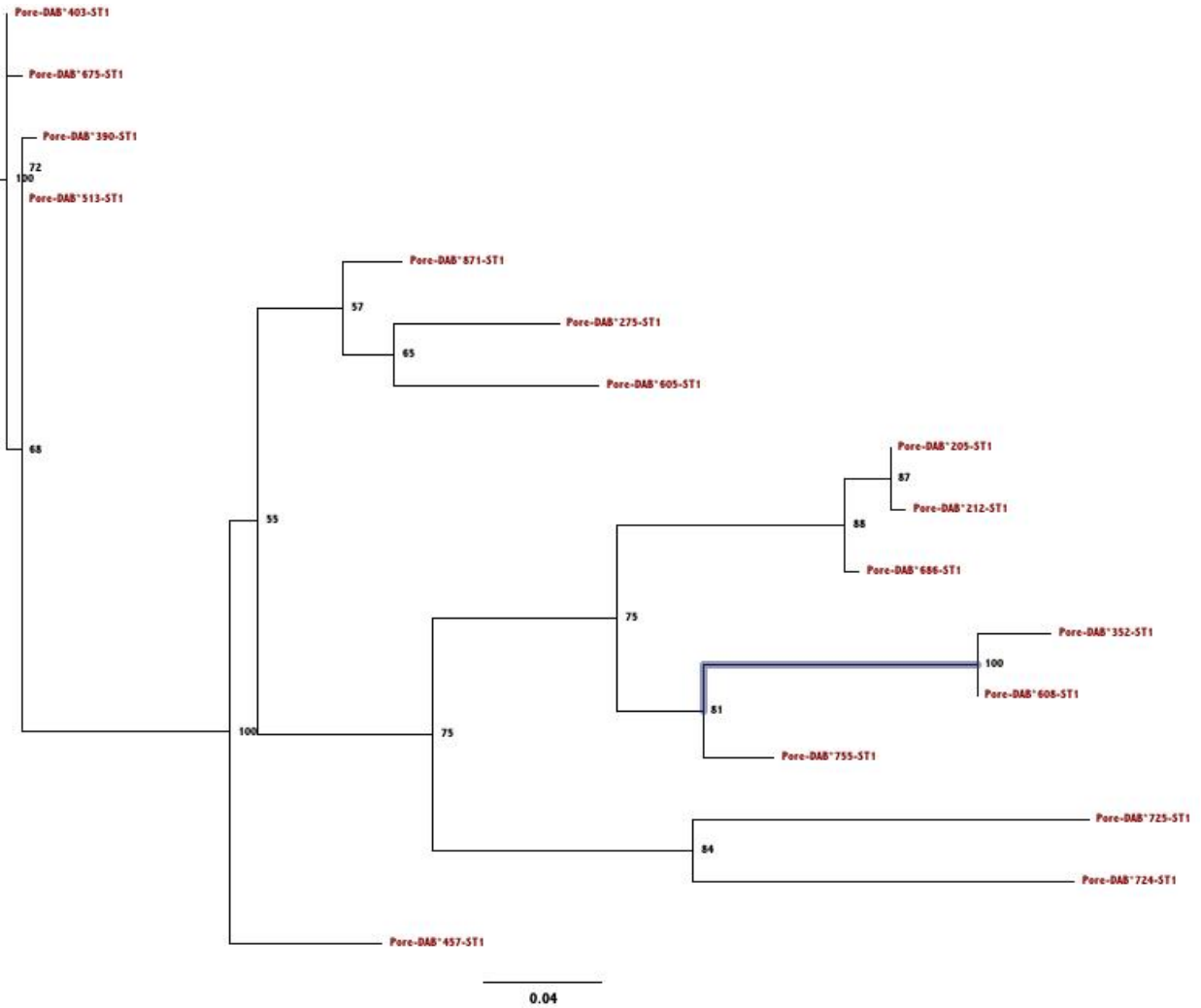


Figure 6. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST1 alleles, with bootstrap values.

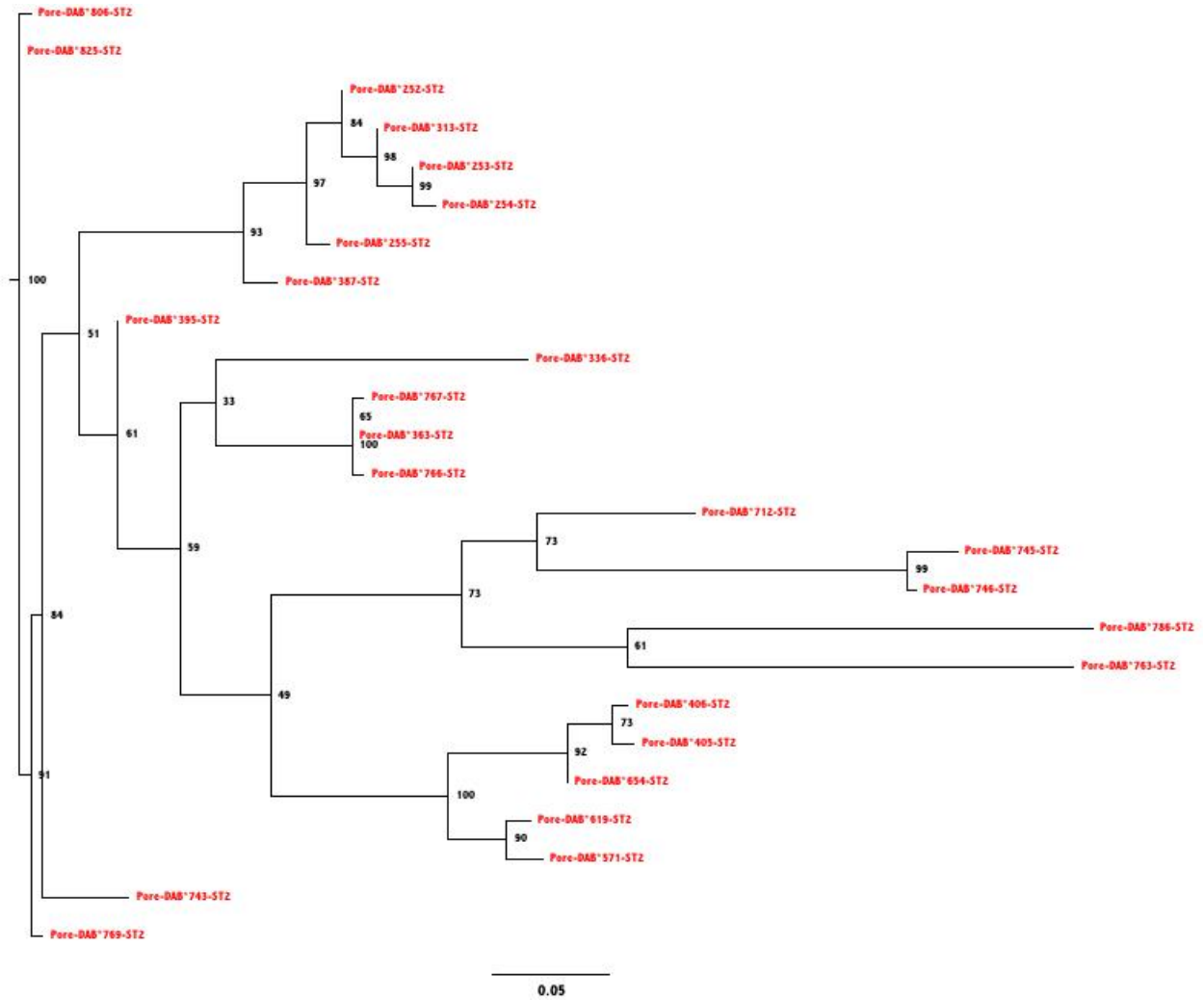


Figure 7. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST2 alleles, with bootstrap values.

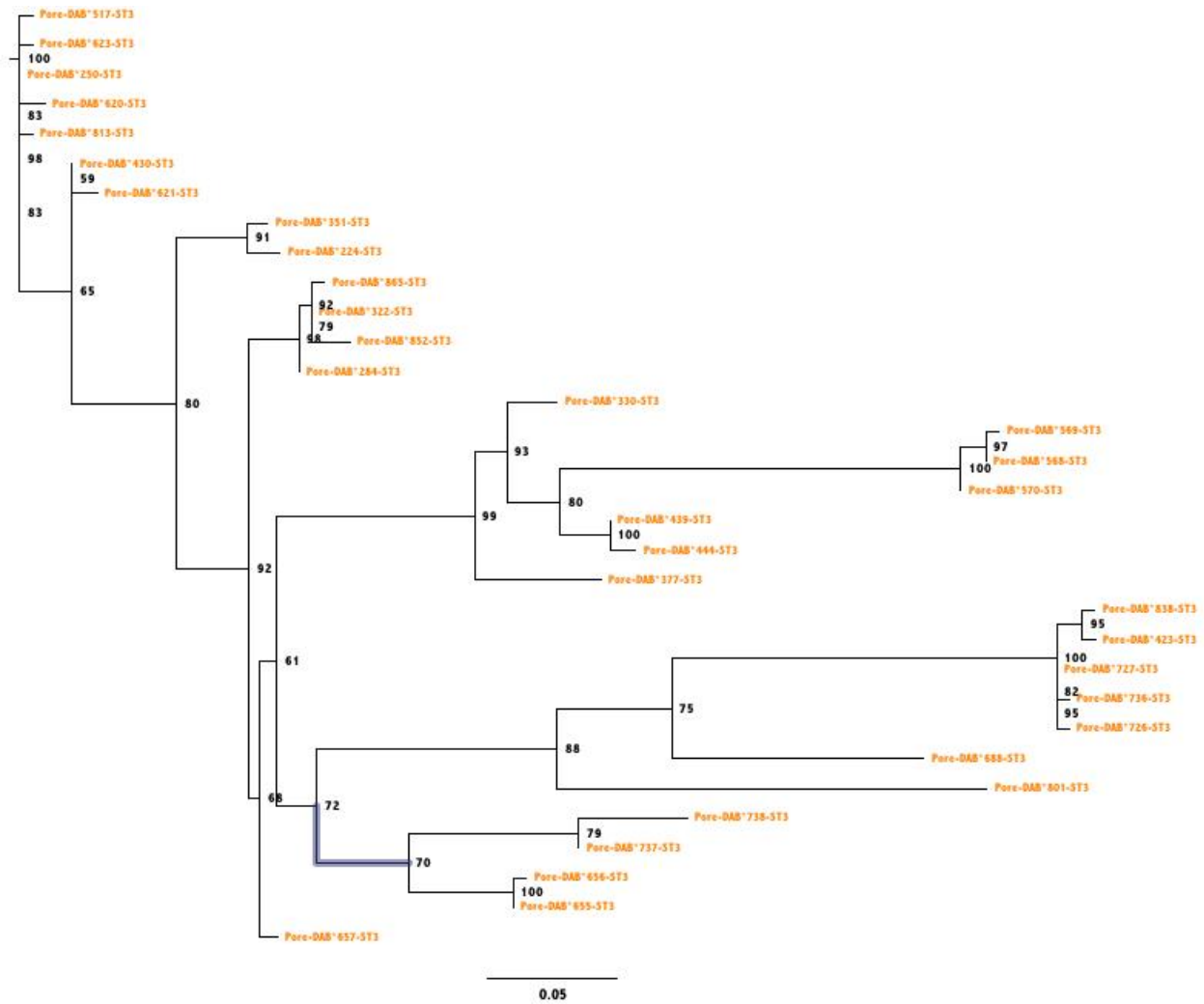


Figure 8. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST3 alleles, with bootstrap values.

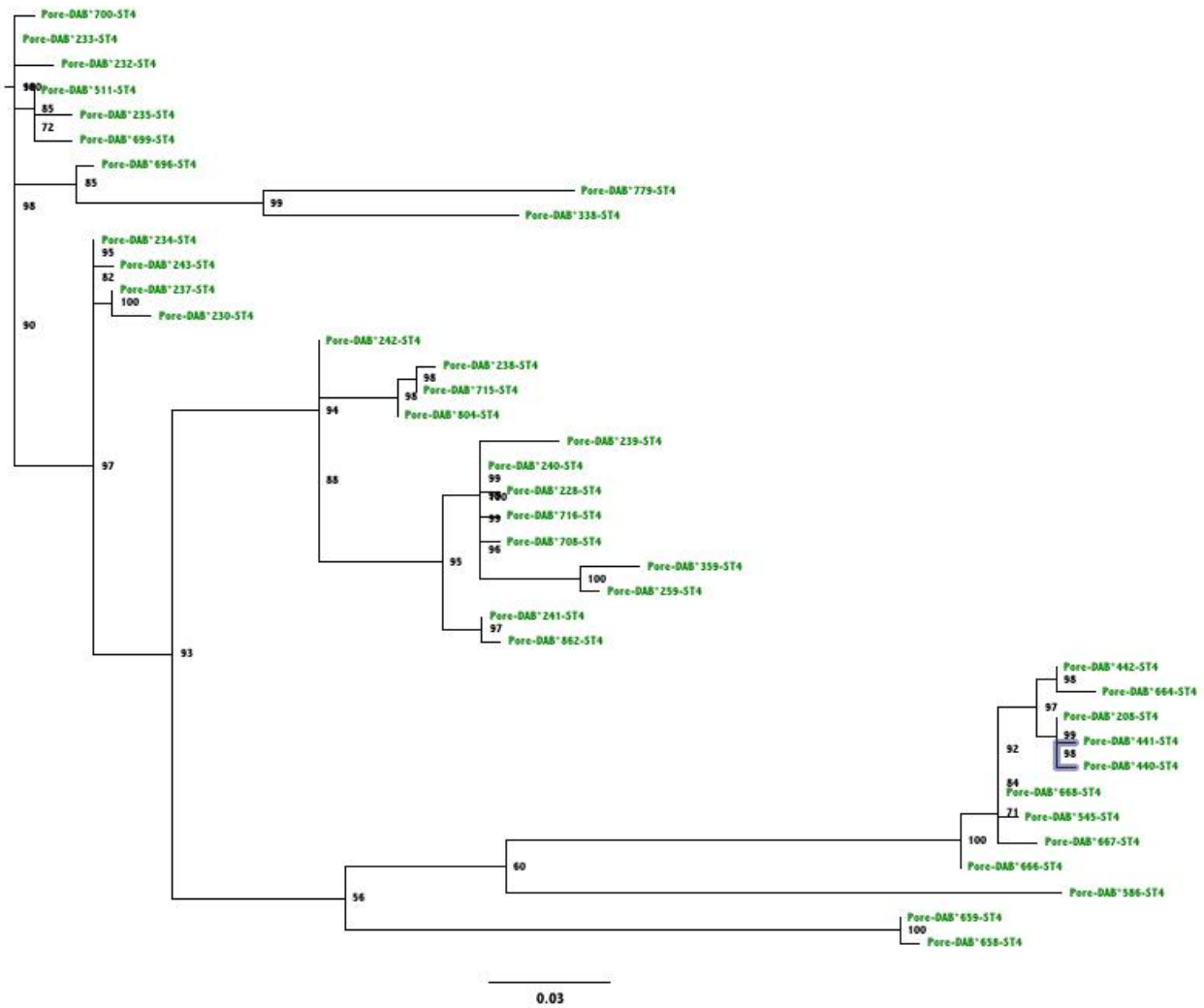


Figure 9. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST4 alleles, with bootstrap values.

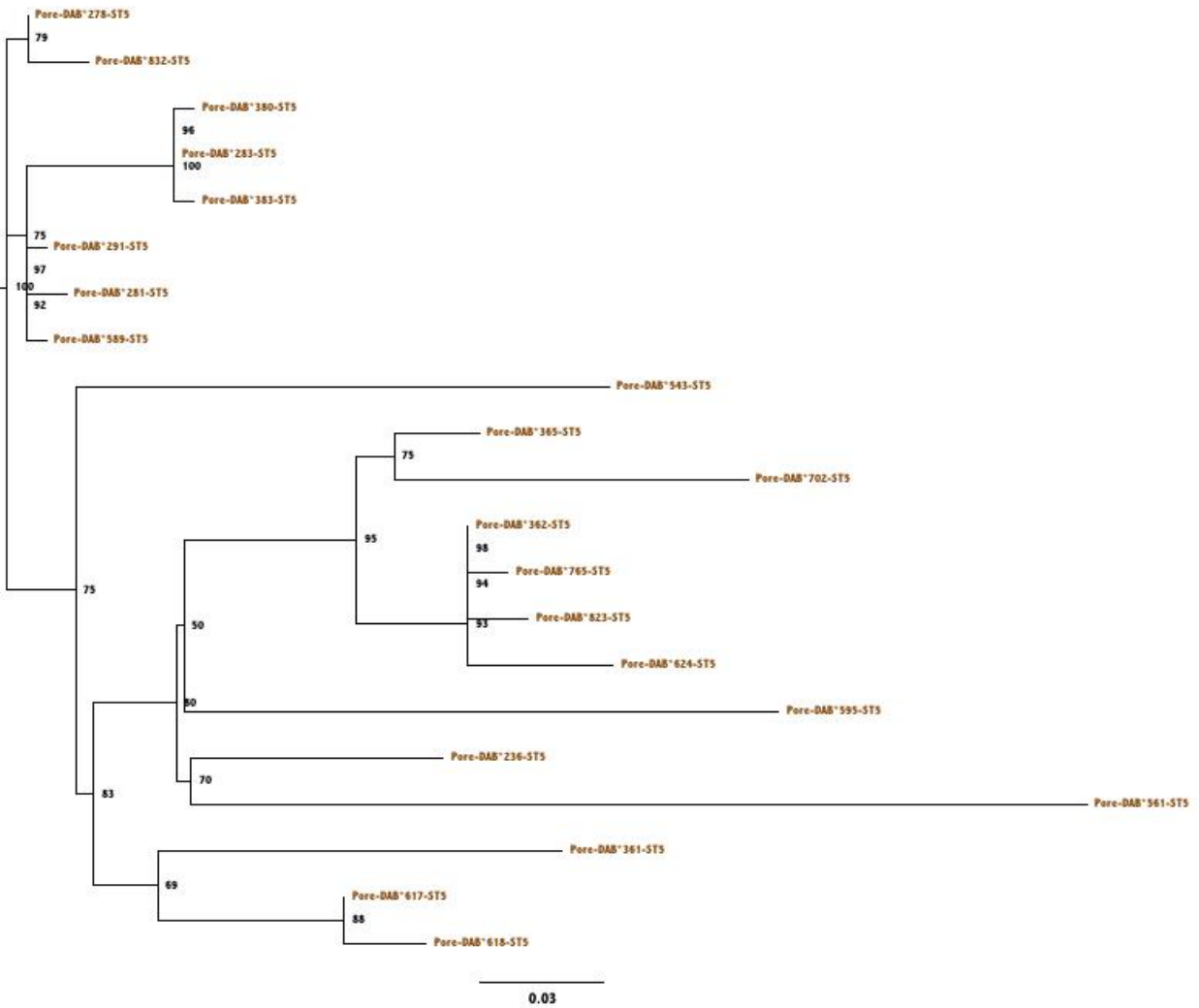


Figure 10. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST5 alleles, with bootstrap values.

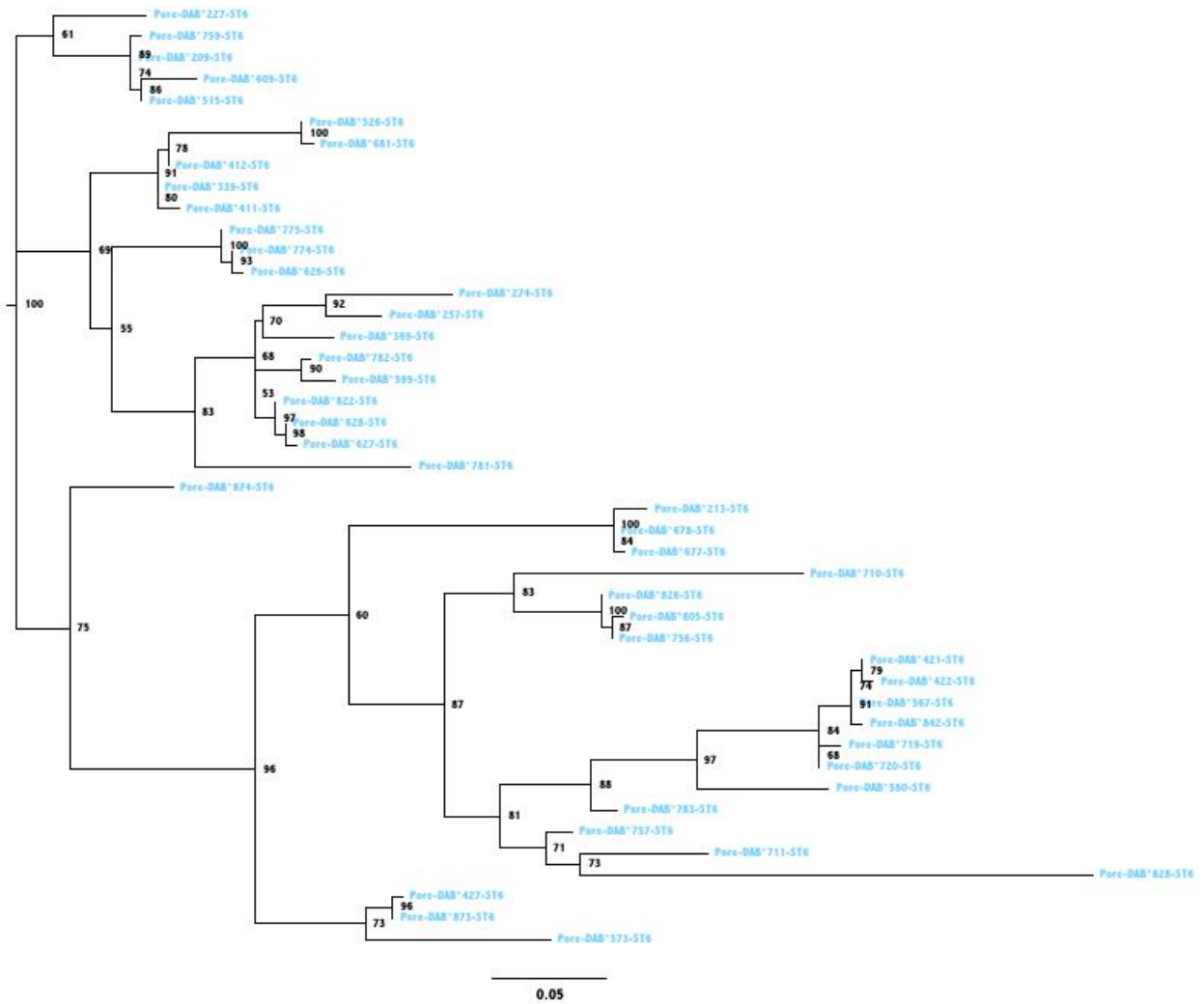


Figure 11. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST6 alleles, with bootstrap values.

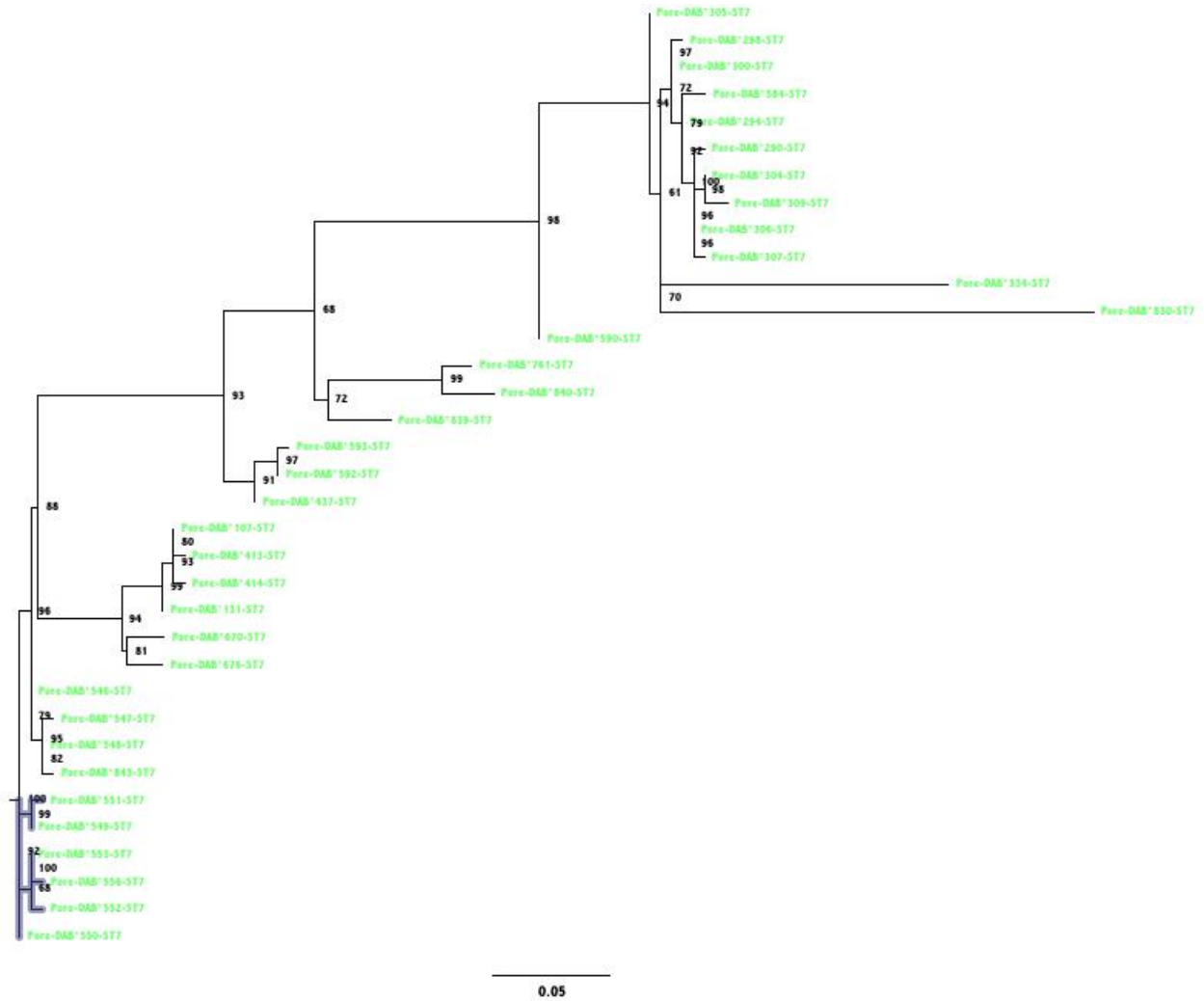


Figure 12. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST7 alleles, with bootstrap values.

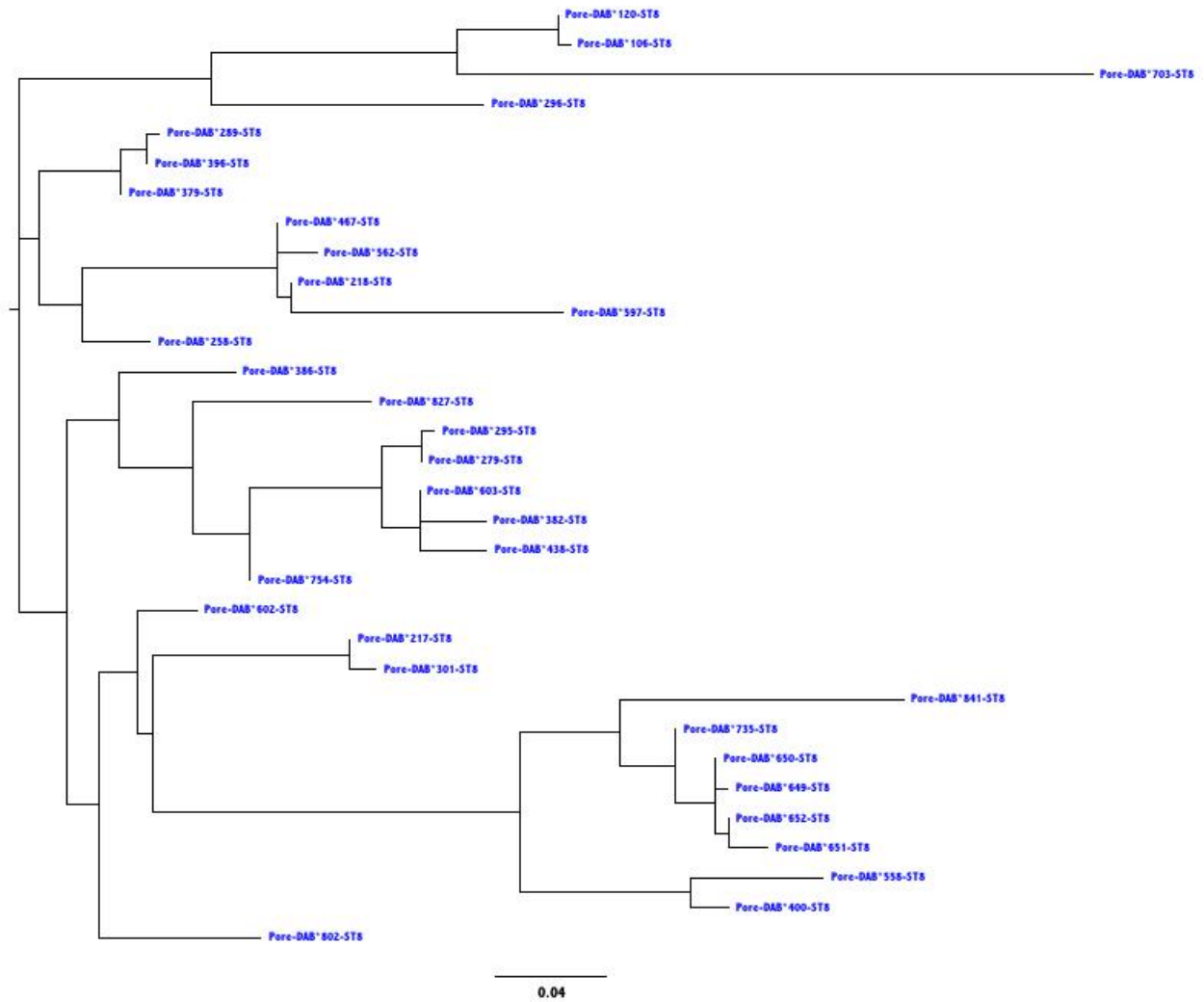


Figure 13. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST8 alleles, with bootstrap values.

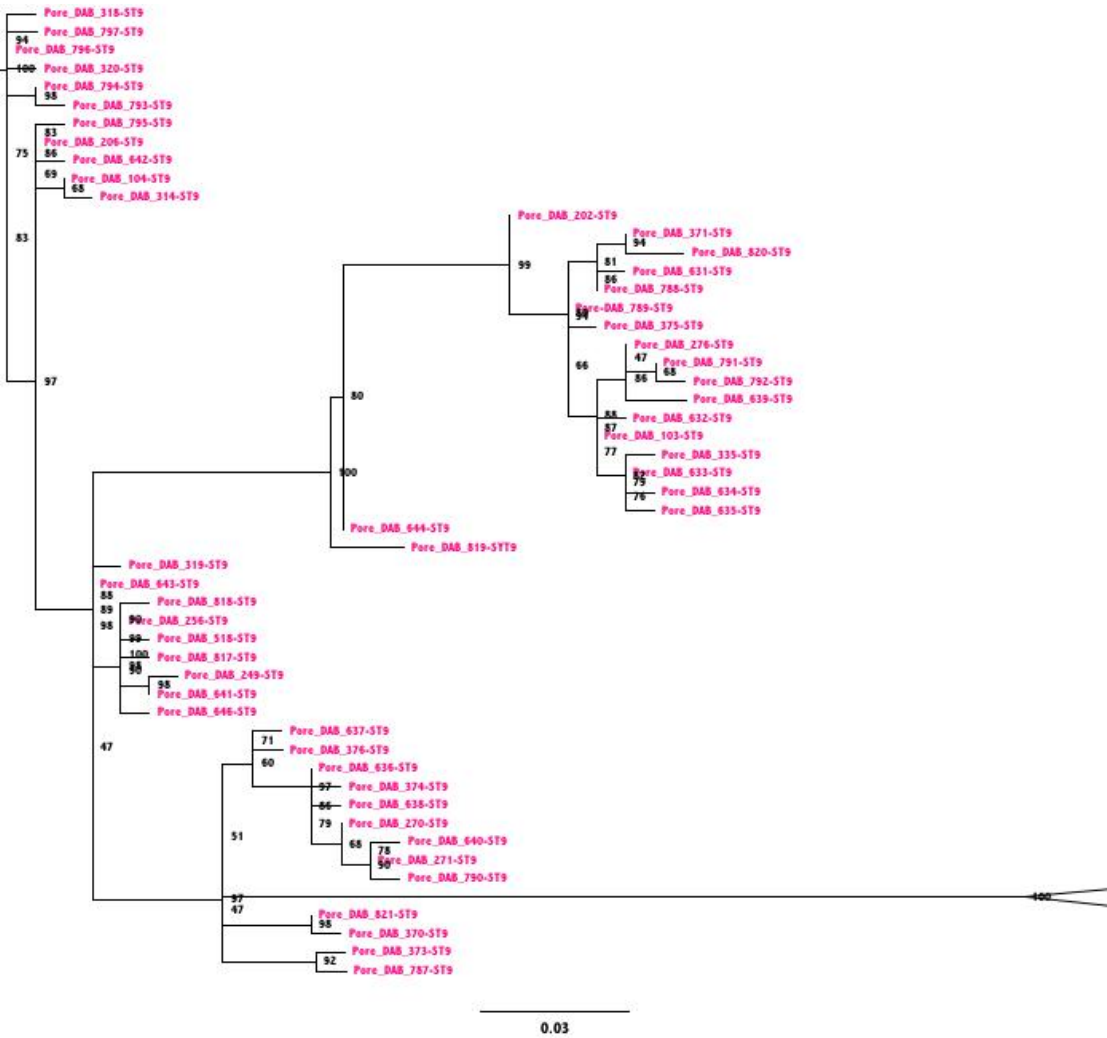


Figure 14. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST9 alleles, with bootstrap values.

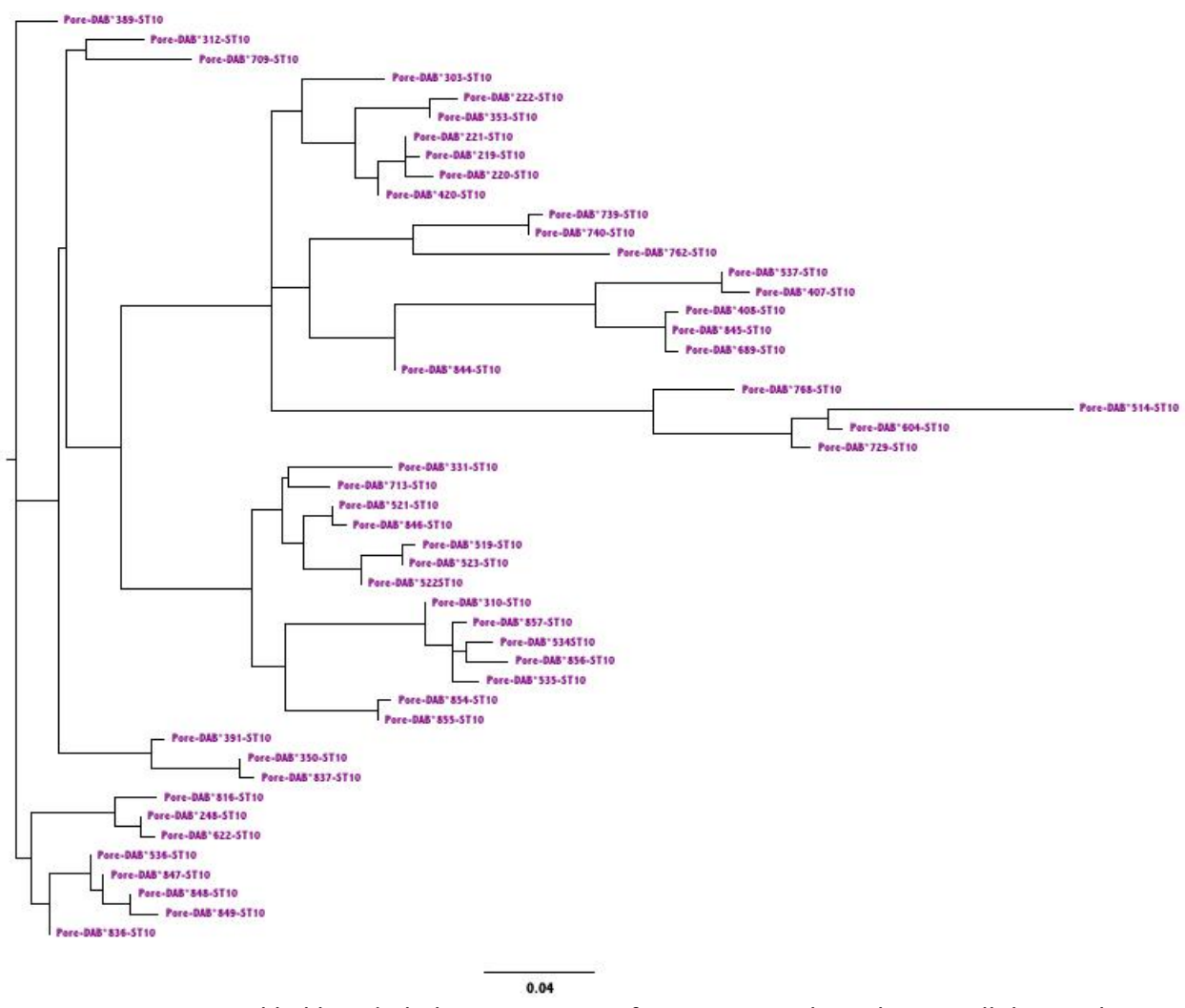


Figure 15. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST10 alleles, with bootstrap values.

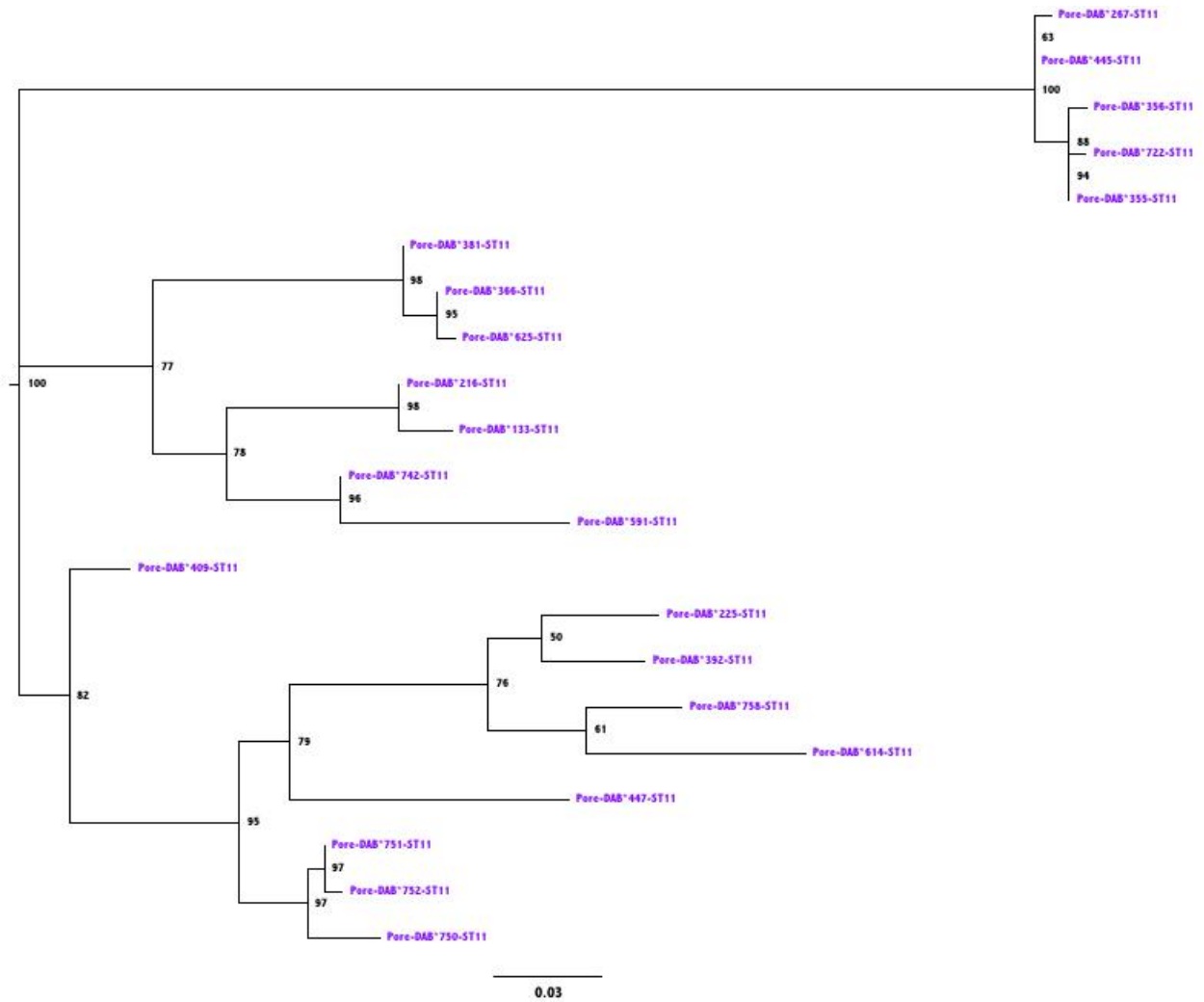


Figure 16. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST11 alleles, with bootstrap values.

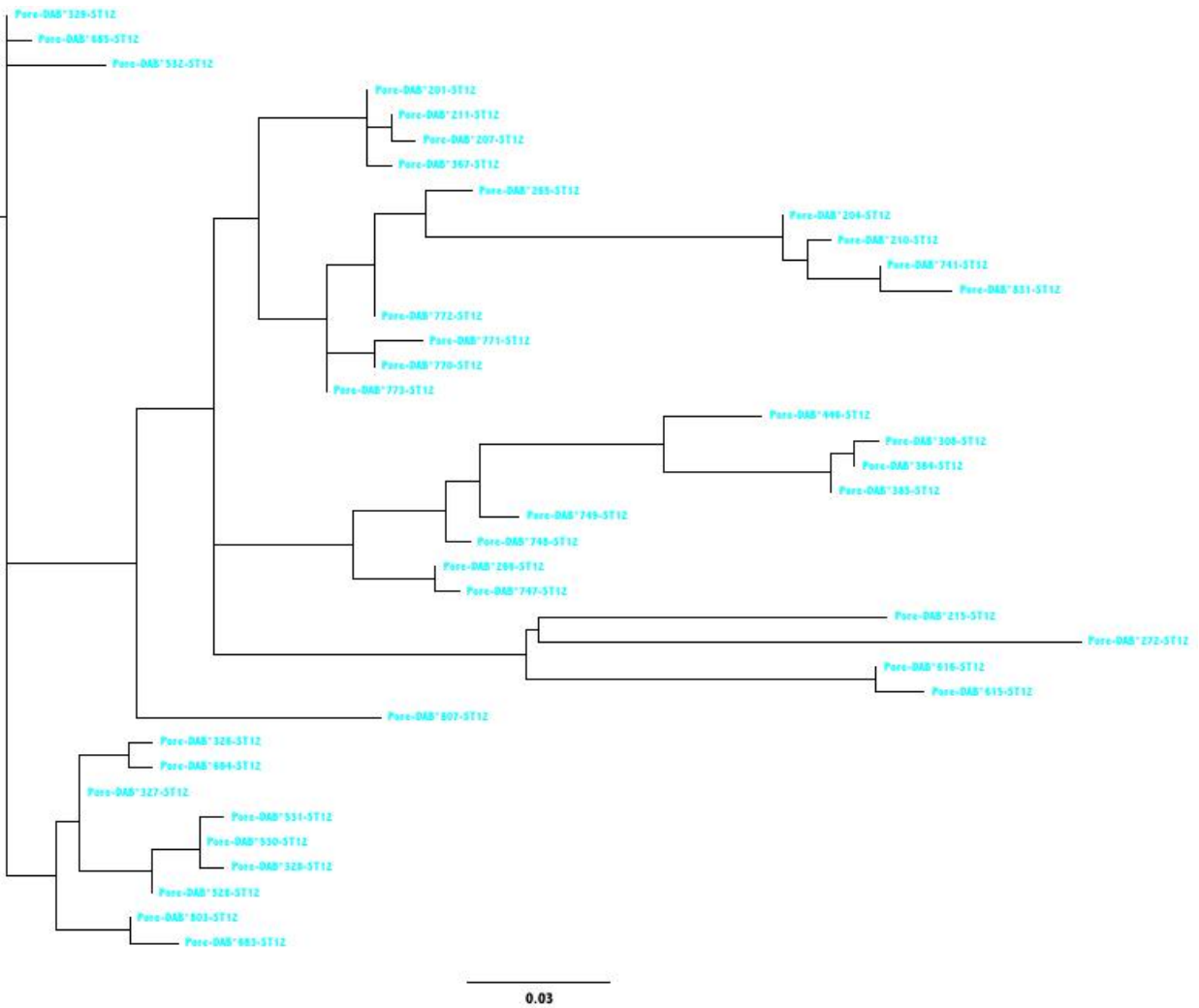


Figure 17. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST12 alleles, with bootstrap values.

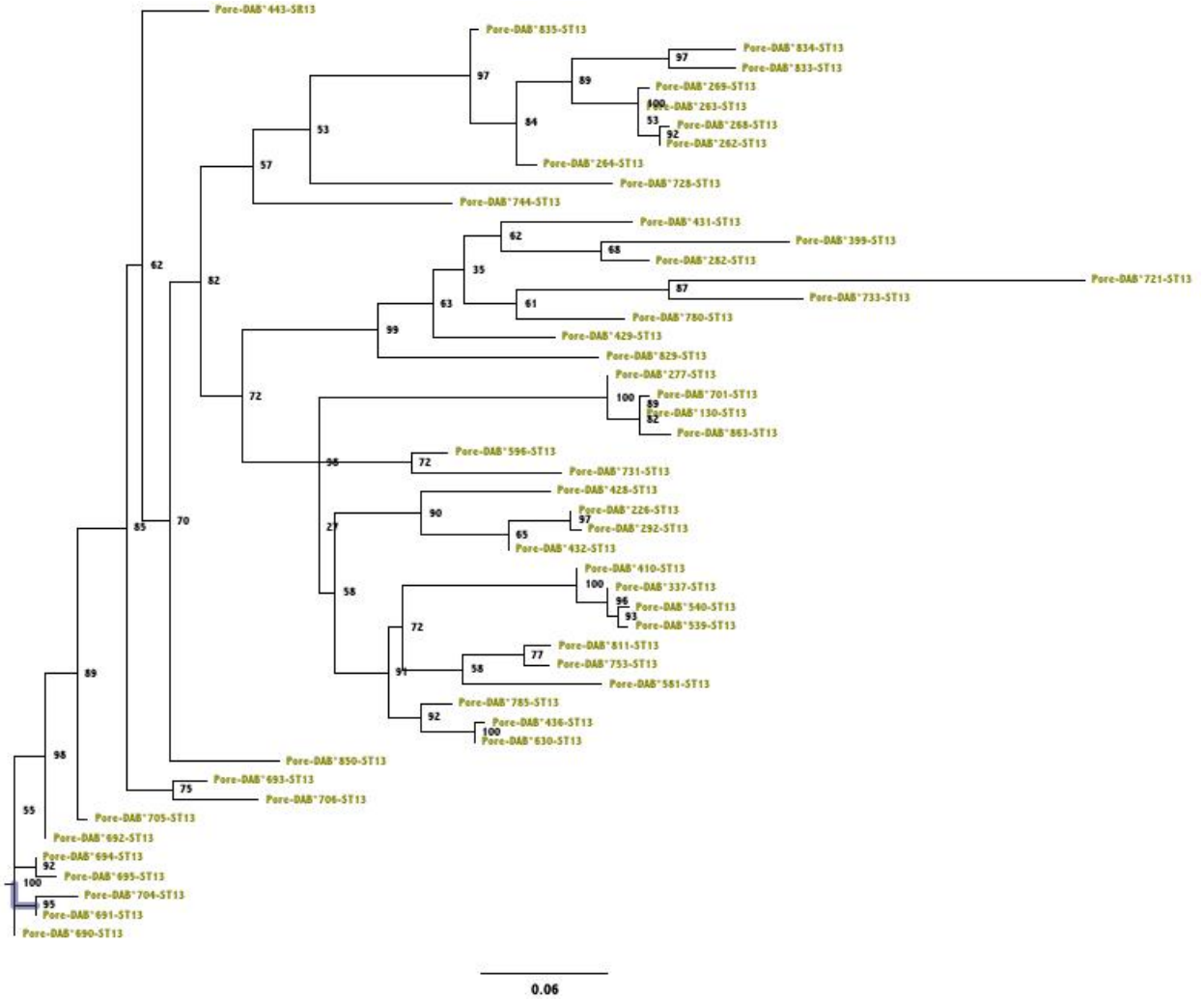


Figure 18. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST13 alleles, with bootstrap values.

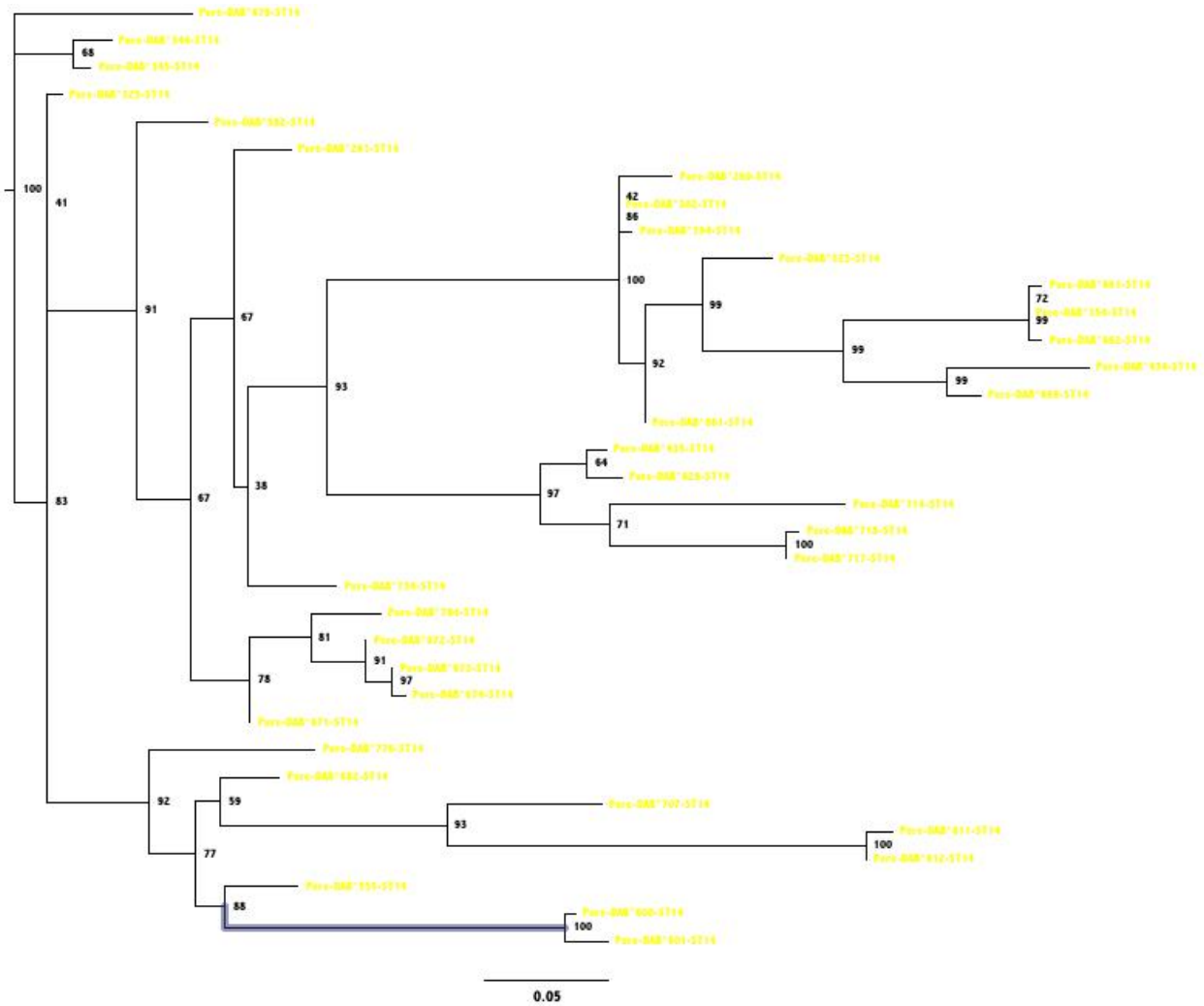


Figure 19. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST14 alleles, with bootstrap values.

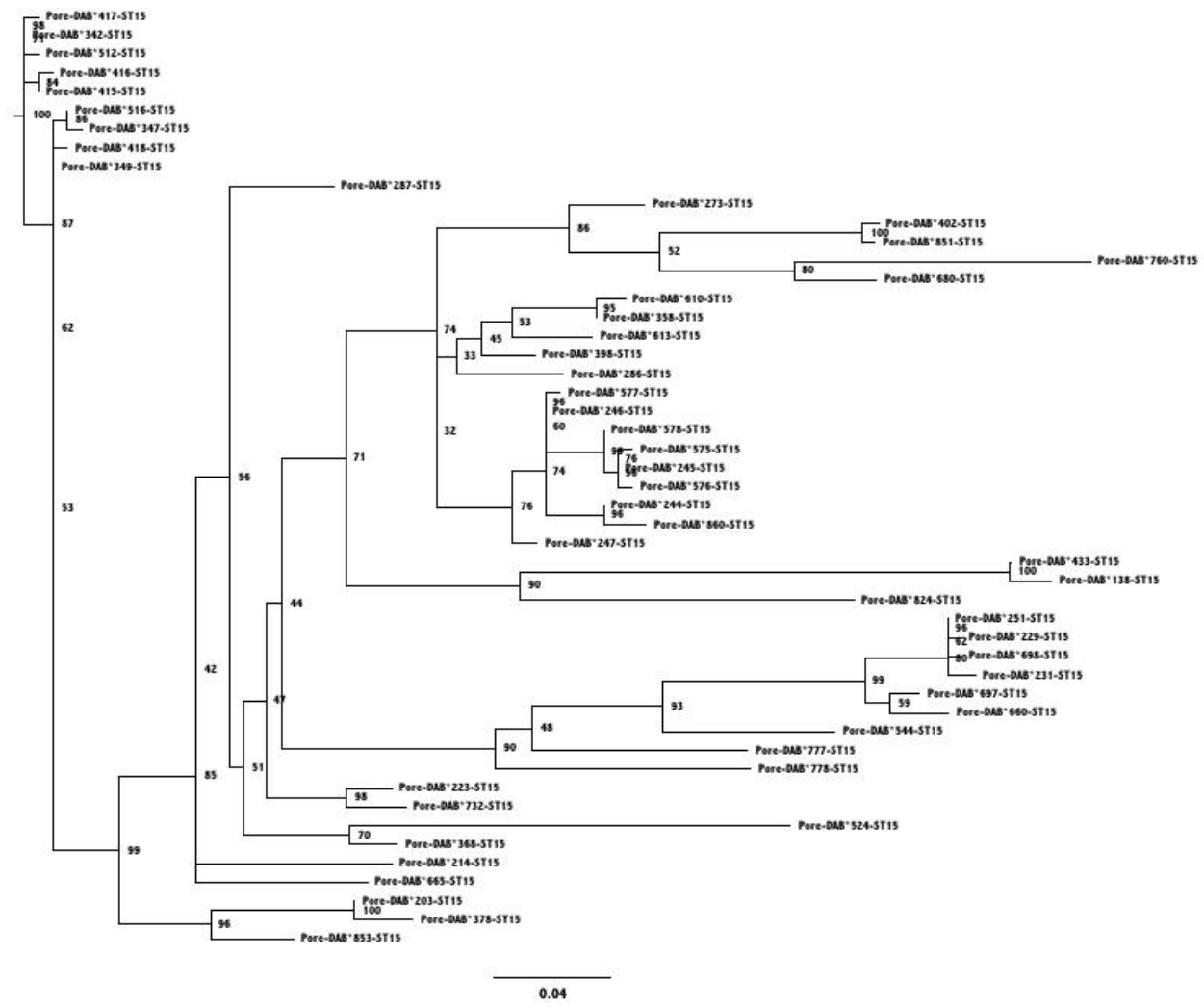


Figure 20. Maximum likelihood phylogenetic tree of guppy MHC class IIb ST15 alleles, with bootstrap values.

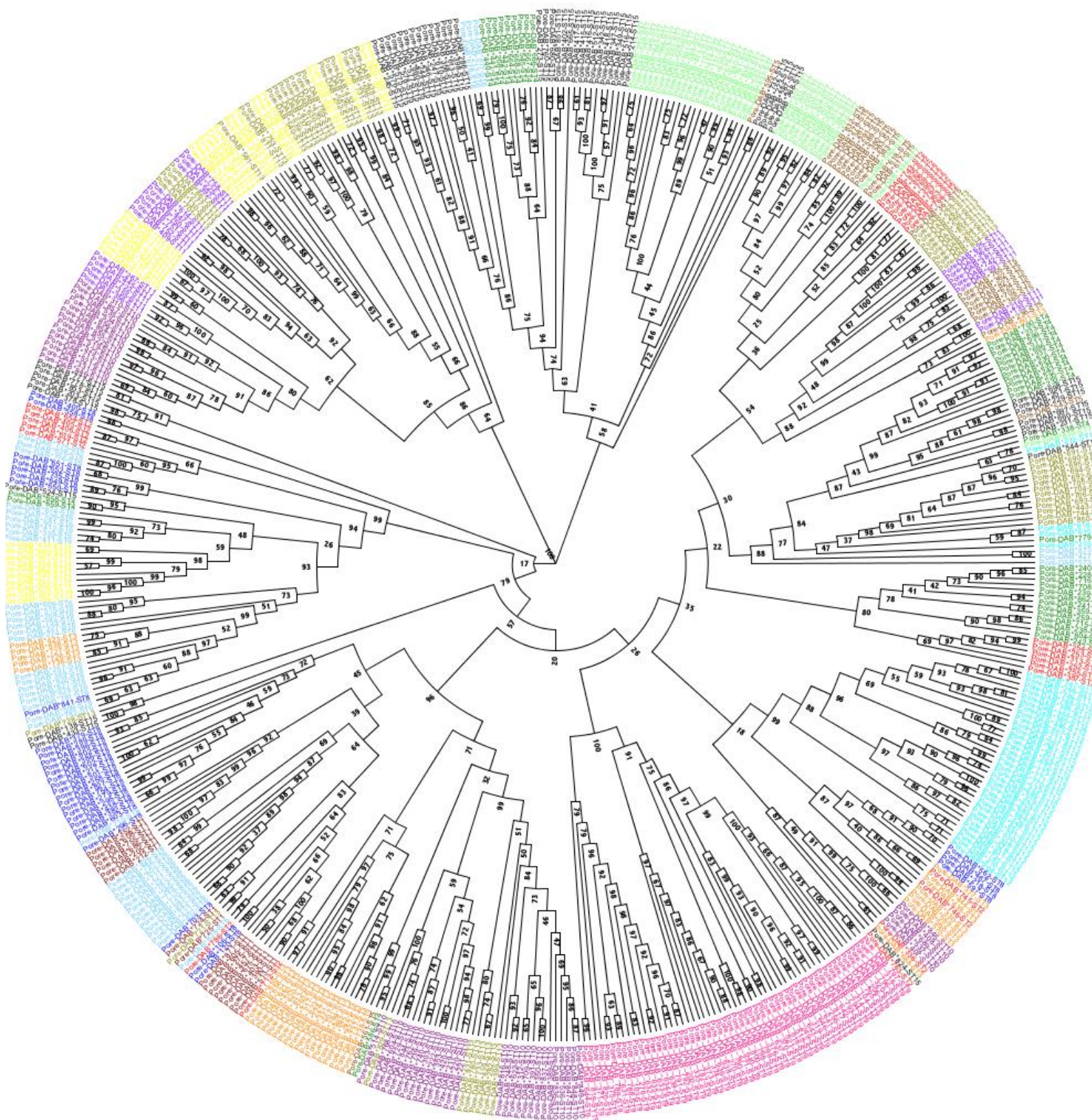


Figure 21. Maximum likelihood phylogenetic tree of guppy all available MHC class IIb alleles for this study, with bootstrap values. Colour coded by ST; ST1 burgundy, ST2 red, ST3 Orange, ST4 dark green, ST5 brown, ST6 powder blue, ST7 lime green, ST8 blue, ST9 pink, ST10 magenta, ST11 neon purple, ST12 neon blue, ST13 olive, ST14 yellow, ST15 Black.

GenGIS phylogeographic maps

The maximum likelihood and parsimony trees were overlaid with maps of Trinidad connecting node of tree to sample sites where allele can be found. Monte Carlo permutation tests were performed to determine if the relationship between tree topology and geographic location of the sample was significant (p -value < 0.05). Correlation between tree topology and geographic location is determined based on the number of times individual end nodes cross over each other when comparing them to location, greater number of crossing suggests that there is no correlation. Site locations are coloured by their watershed; Caroni, North east (NE), North slope, Oropouche, South west, and South east. The maximum likelihood of ST3 had significant correlation between tree topology and geographic location, with a p -value of 0.001 (figure 22). The maximum parsimony tree of ST4 had significant correlation between tree topology and geographic location, with a p -value 0.027(figure 23). Both the maximum likelihood and parsimony trees of ST7 had significant correlation between tree topology and geographic location, p -values of 0.014 and 0.007 respectively (figures 24, 25). The maximum parsimony tree of ST10 had significant correlation between tree topology and geographic location, with a p -value 0.002(figure 26). The maximum parsimony tree of ST12 had significant correlation between tree topology and geographic location, with a p -value of 0.032 (figure 27). The maximum likelihood tree of ST14 had significant correlation between tree topology and geographic location, with a p -value of 0.010 (figure 28). The maximum likelihood tree that included all alleles of all STs also found significant relation to tree topology and geographic location of sample site, p -value0.019. ST9 appears to be distinct from other types in that it is separate from all others in the tree containing all alleles (figure 21), and as such the maximum

likelihood tree compared to the geography has been included (figure 29) despite not having significant correlation with geography, p-value 0.850.

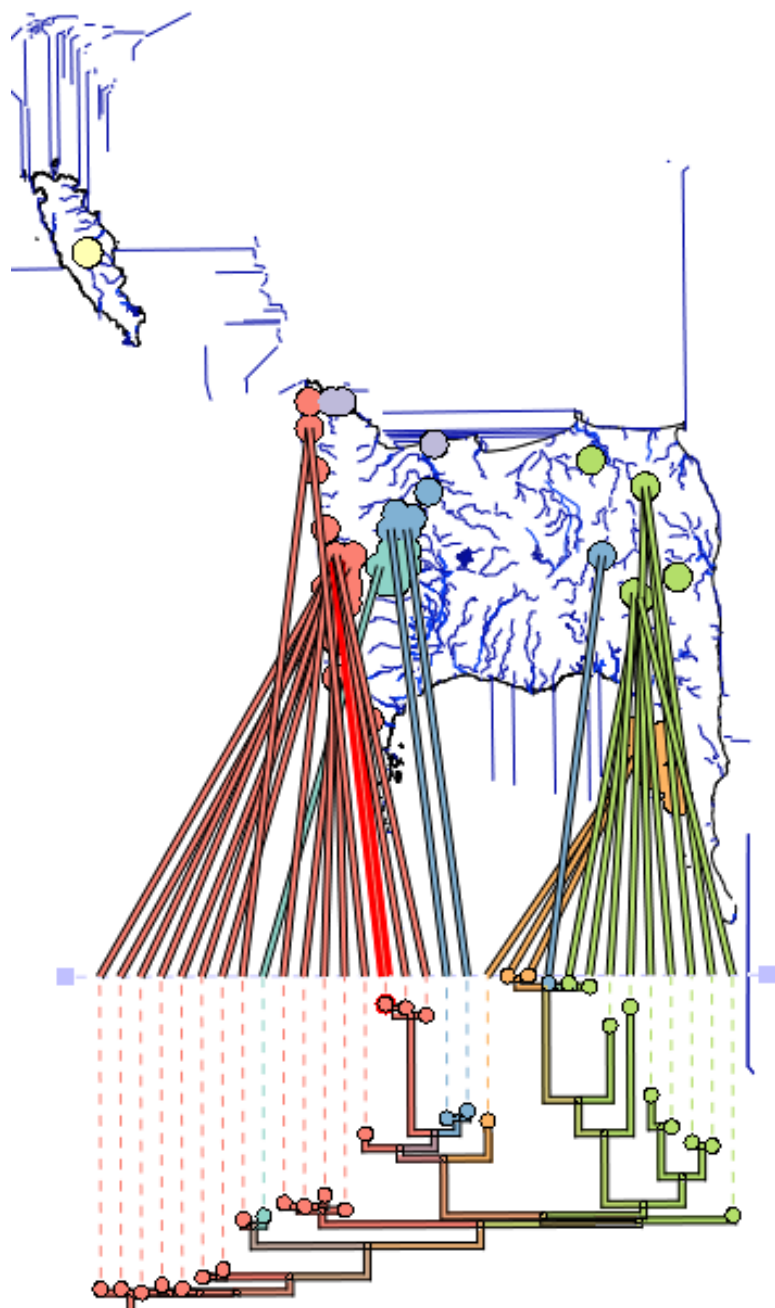


Figure 22. Geophylogeny of guppy MHC class IIb ST3 alleles, derived from maximum likelihood phylogenetic tree.

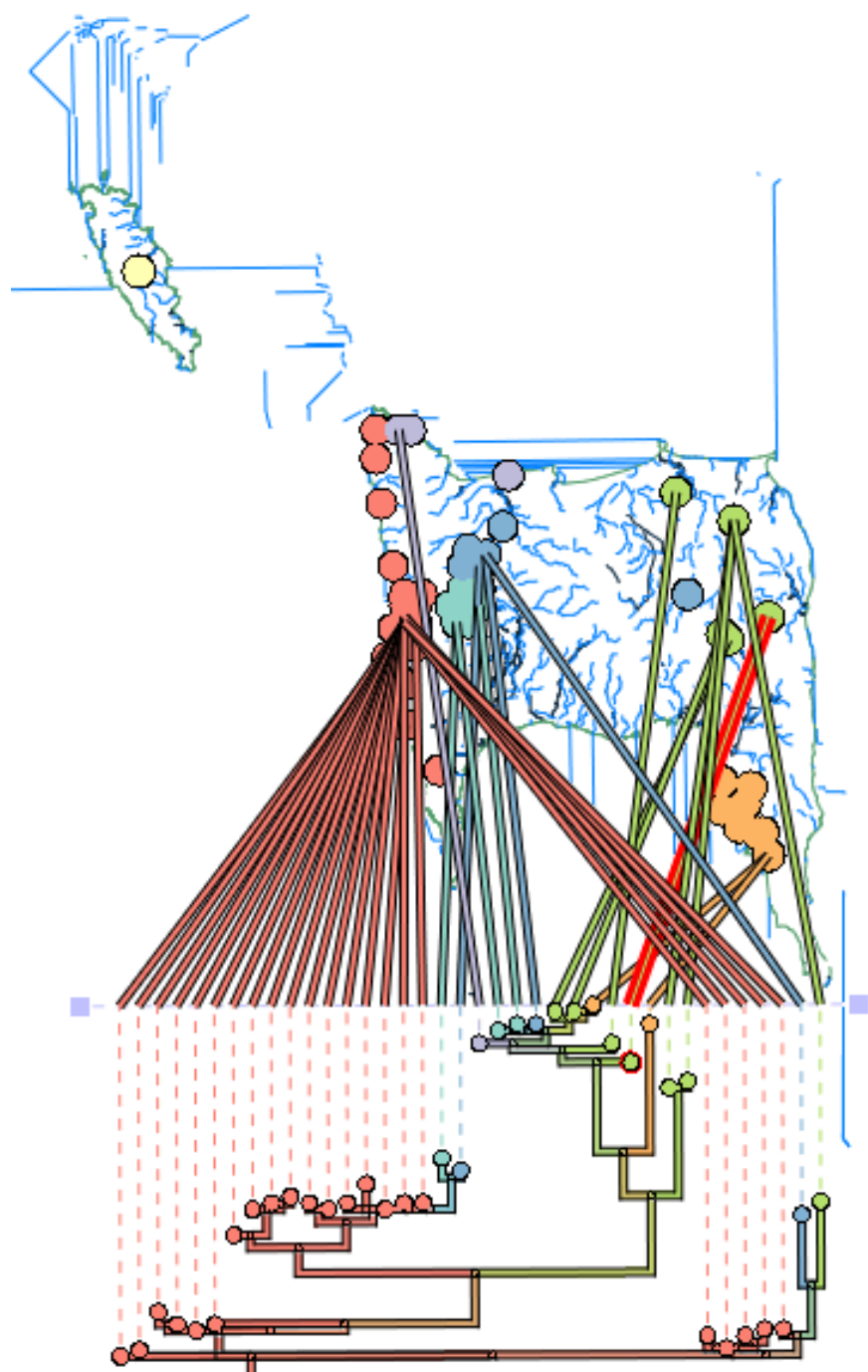


Figure 23. Geophylogeny of guppy MHC class IIb ST4 alleles, derived from maximum parsimony phylogenetic tree.

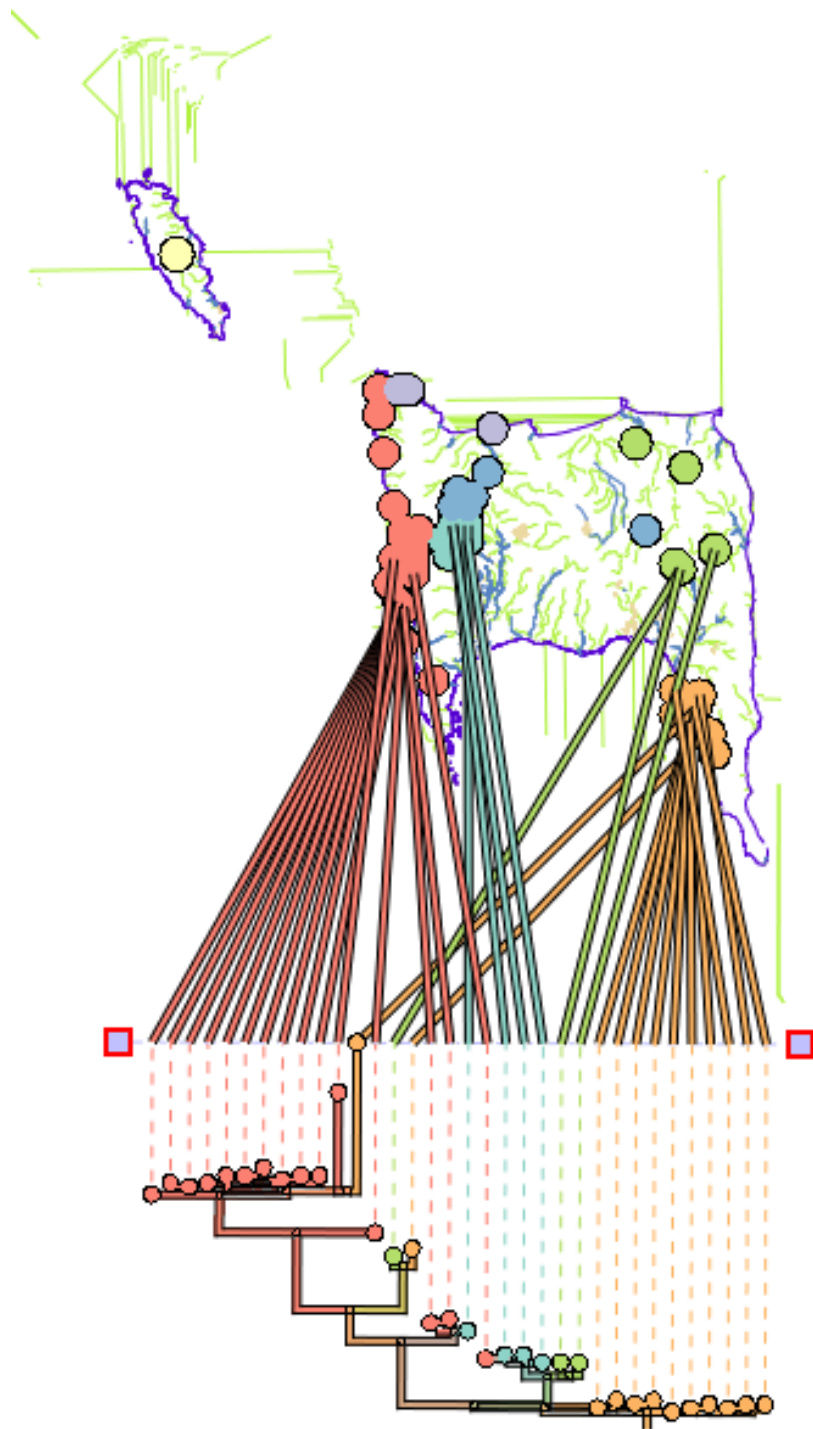


Figure 24. Geophylogeny of guppy MHC class IIb ST7 alleles, derived from maximum likelihood phylogenetic tree.

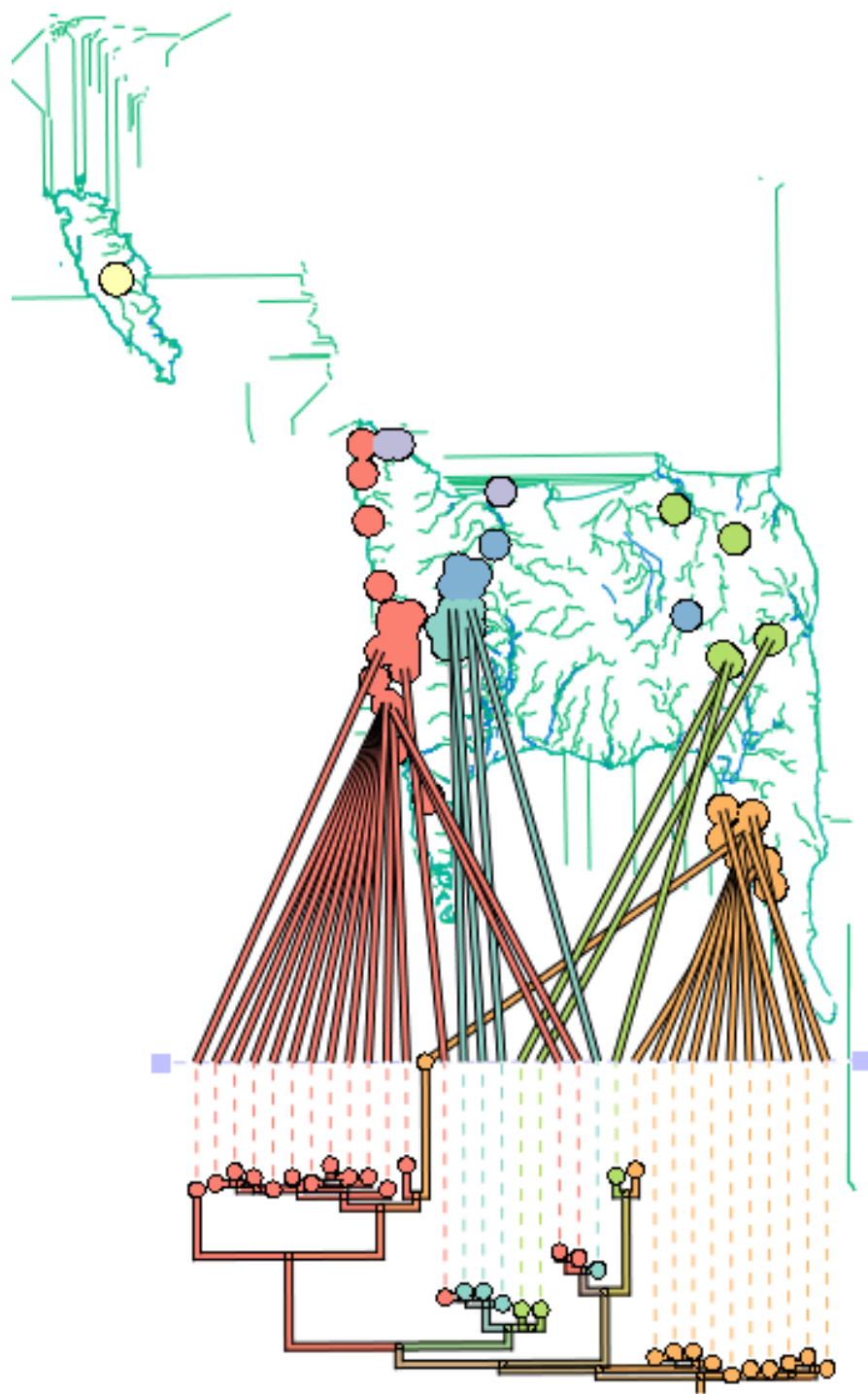


Figure 25. Geophylogeny of guppy MHC class IIb ST7 alleles, derived from maximum parsimony phylogenetic tree.

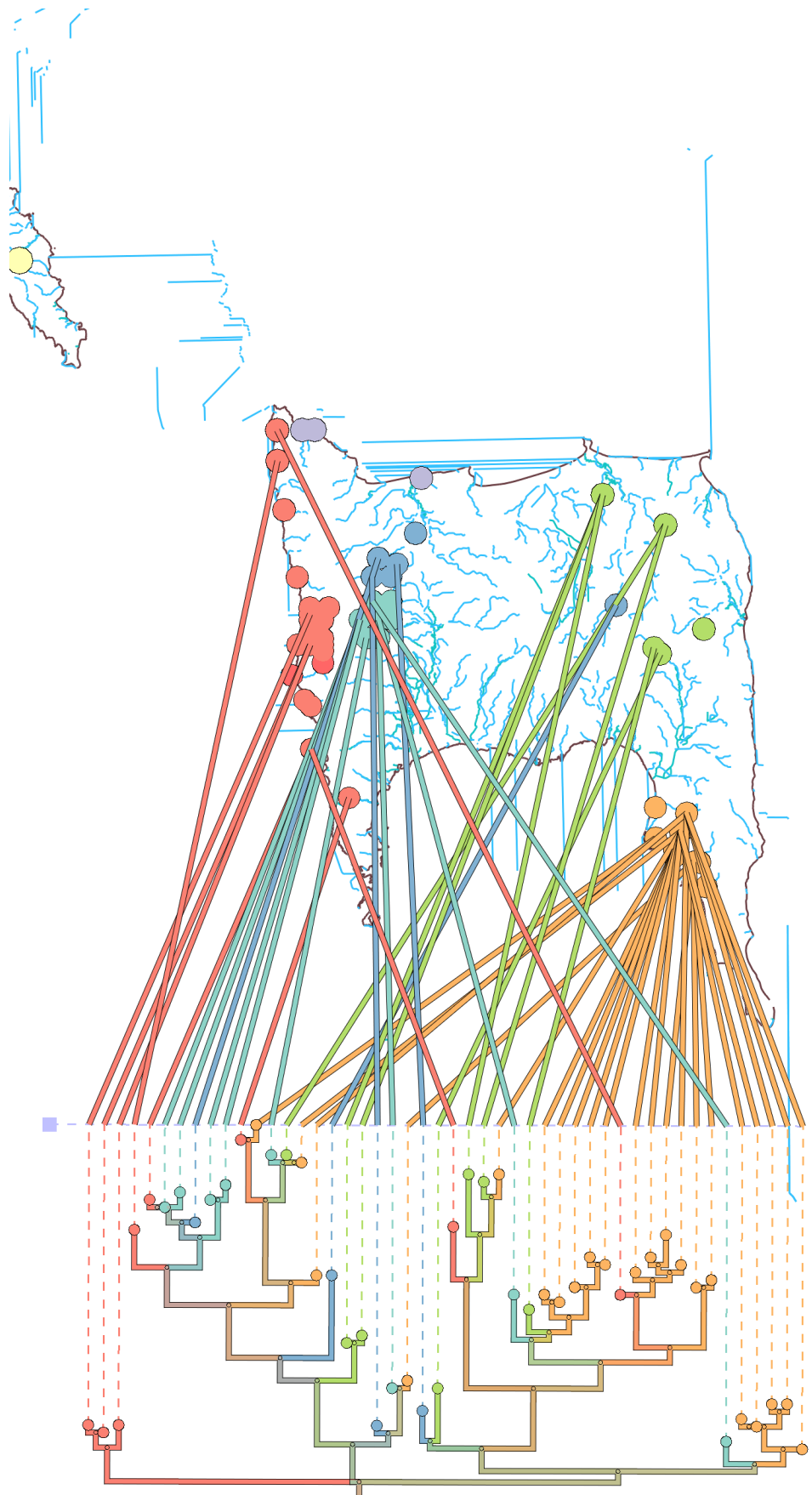


Figure 26. Geophylogeny of guppy MHC class IIb ST10 alleles, derived from maximum parsimony phylogenetic tree.

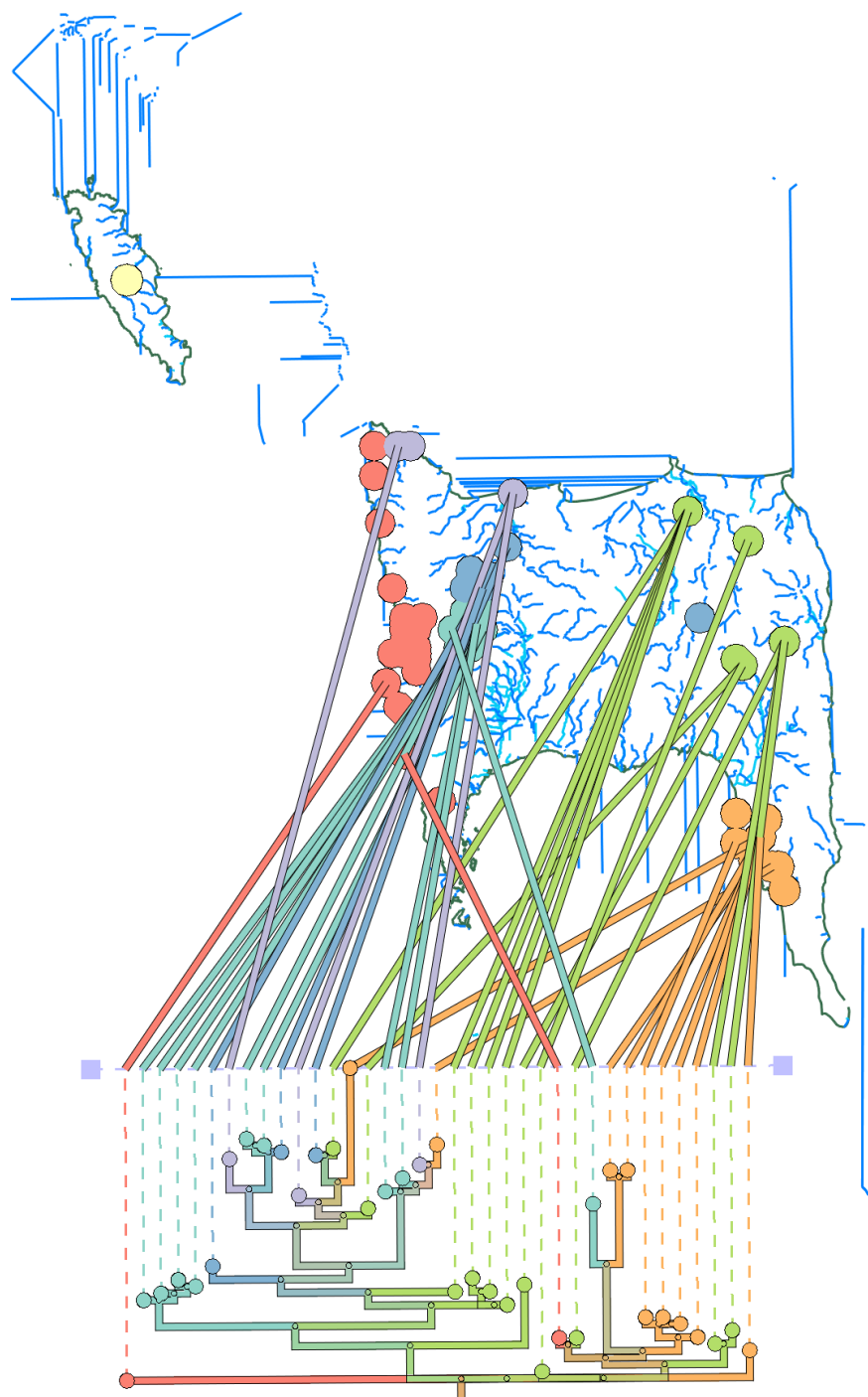


Figure 27. Geophylogeny of guppy MHC class IIb ST12 alleles, derived from maximum parsimony phylogenetic tree.

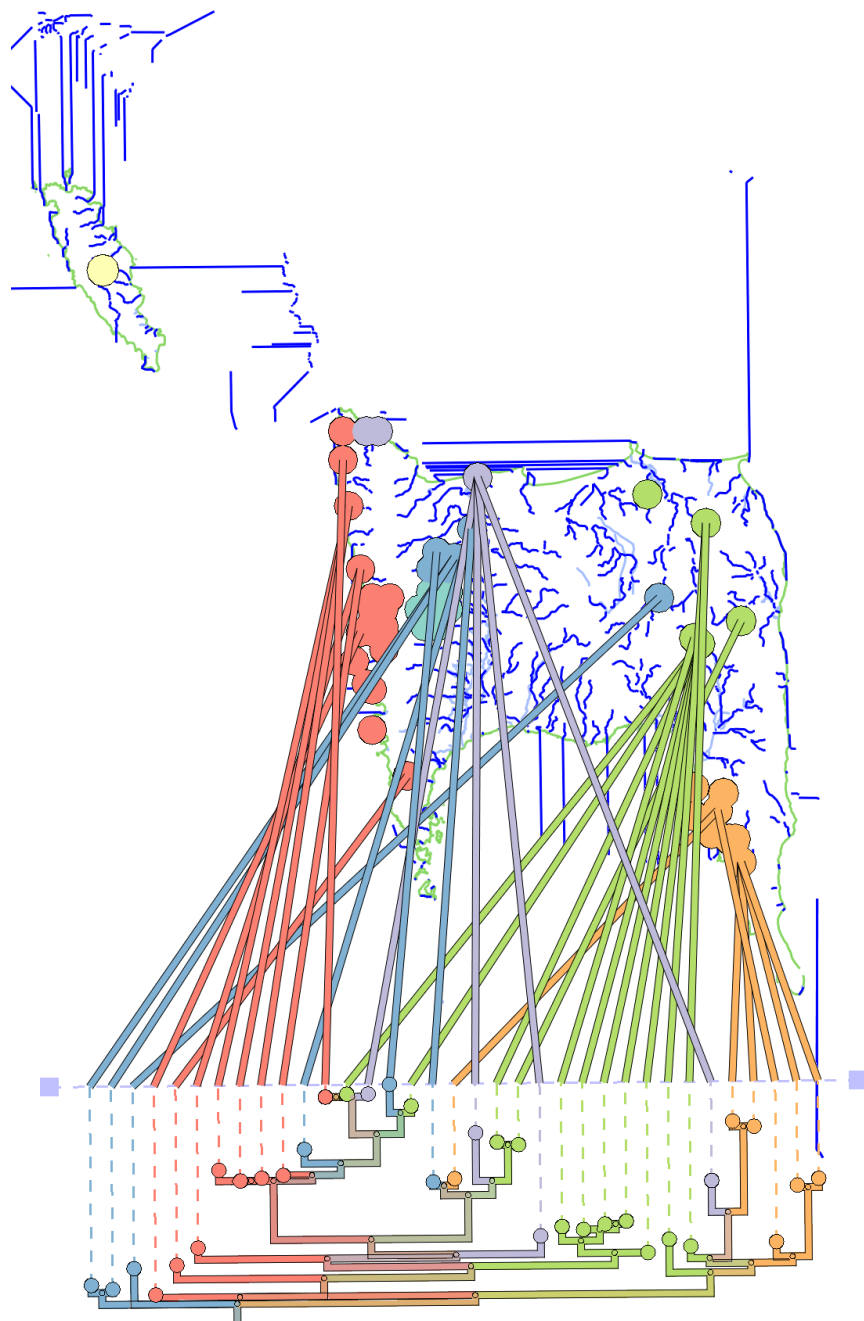


Figure 28. Geophylogeny of guppy MHC class IIb ST14 alleles, derived from maximum likelihood phylogenetic tree.

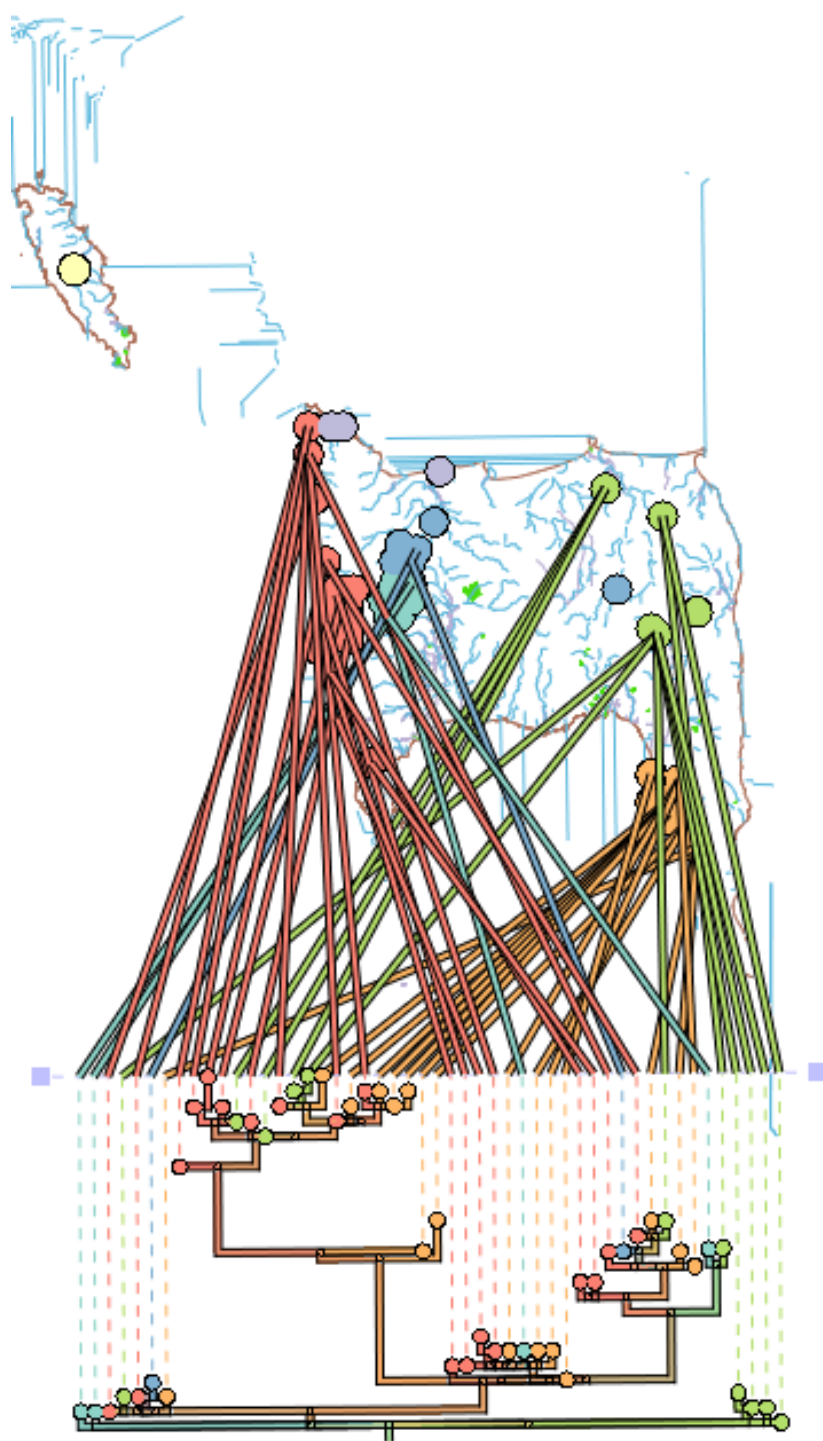


Figure 29. Geophylogeny of guppy MHC class IIb ST9 alleles, derived from maximum likelihood phylogenetic tree.

Chapter 5

Discussion

In looking at the total diversity of MHC we see that there is a significant correlation between the phylogeny of the tree and the geography of Trinidad. This could be resultant of isolation by distance; however, this correlation is not present in all individual supertypes. Looking at the phylogenetic tree for all MHC alleles we see that STs are frequently interspersed throughout the tree rather than being found in specific clade, with the exception of ST 9. The classification of MHC STs as used by Lighten (2015), is based on the translated protein binding protein properties, which potentially allows for the separation of supertypes when looking at them the nucleotide level, due to the redundancy of the genetic code. The breaking up of STs in the full tree may also allow for the correlation between geography and phylogeny, with clusters located in different branches being exclusive to clusters of populations, i.e. ST 3 with the north south divide and two large clusters in the full tree. Despite this ST 9 is found as a single clade of the phylogenetic tree (figure 21), and is found in almost all sites across Trinidad however it shows no correlation with the geography of Trinidad (figure 29). ST 9 being a separate clade that is wide spread may suggest that it is a precursor or the founding supertype on the island, or somehow separated from the other MHC alleles.

Baillie (2012) has suggested that guppy populations in Trinidad are the result of multiple colonisations from different source populations. Differences in the source populations' MHC patterns may provide some insight on the structure of full tree in terms of supertype, and the correlation with geography. Different founding colonies separated by source, geographic and temporal distance may have had to adapt to local parasites using different source pools of MHC.

This could result in similar phenotype despite different sequences, and would be especially localized to founding sites.

MHC are old genes in terms of evolution, being found in all jawed vertebrates (Herdegen, Babik, & Radwan, 2014). The rapid rate of change in MHC sequences combined with their age may explain some of the. Some ST maybe older than others resulting in greater distribution, especially when looking at the role of MHC in immunity and the high rate of change. ST, or precursors of, that may have been present during colonization of Trinidad would be expected to have a greater distribution than a ST that arose in in the mostly isolated populations in Trinidadian rivers. STs 4, 10, and 7 all found significant correlation to geography with their maximum parsimony trees, may be indicative of this. Parsimony trees are constructed by conjoining branches with the fewest amounts of changes between nucleotides, this may suggest recent divergences (Nakjang, et al., 2013).

The primary role of the MHC is in parasitic immunity, and as such the absence and presence of the parasitic target, of the PBR are expected to influence the distribution of alleles. The role of secondary effects of MHC on sexual selection, i.e. through increased pigmentation, may retain alleles that are not effect for immune function but otherwise increase factors for mate choice (Lighten, 2015). This retention could result in increased distribution by not allowing for the same amount of genetic drift that would exist from just natural selection (Lighten, 2015). Further studies looking at HMC and specific target immunity may be taken to address this.

Bibliography

- Babik, W., Taberlet, P., Ejsmond, M. J., & Radwan, J. (2009). New generation sequences as a tool for genotyping of highly polymorphic multilocus MHC system. *Molecular Ecology*, 9(3), 713-719.
- Baillie, L. (2012). *Genetic population structure of the Trinidadian Guppy (Poecilia reticulata) across Trinidad and Tobago*. Master's Thesis, Dalhousie University, Biology, Halifax.
- Barson, N. J., Cable, J., & Van Oosterhout, C. (2009). Population genetic analysis of microsatellite variation of guppies (*Poecilia reticulata*) in Trinidad and Tobago: evidence for a dynamic source-sink metapopulation structure, founder events and population bottlenecks. *Journal of evolutionary biology*, 22(3), 485-497.
- Canty, A. J. (2002). Resampling methods in R: the boot package. *R News* 2.3, 2-7.
- Fraser, B. A., Kunstner, A., Reznick, D. N., Dreyer, C., & Weigel, D. (2015). Population genomics of natural and experimental populations of guppies (*Poecilia reticulata*). 24(2), 389-408.
- Fraser, B. A., Ramnarine, I. W., & Neff, B. D. (2010). Temporal variation at the MHC class IIB in wild populations of the guppy (*Poecilia reticulata*). *Evolution*, 64(7), 2086-2096.
- Grether, G. F., Hudon, J., & Millie, D. F. (1999). Carotenoid limitation of sexual coloration along an environmental gradient in guppies. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1426), 1317-1322.
- Herdegen, M., Babik, W., & Radwan, J. (2014). Selective pressures on MHC class II genes in the guppy (*Poecilia reticulata*) as inferred by hierarchical analysis of population structure. *Journal of Evolutionary Biology*, 27, 2347-2359.
- Lighten, J. (2015). *Elucidating patterns of Major Histocompatibility Complex polymorphisms in the Trinidadian guppy (Poecilia reticulata) using Next Generation Sequencing*. PhD Thesis, Dalhousie University, Philosophy, Halifax, Nova Scotia.
- Magurran, A. E. (2005). *Evolutionary ecology: the Trinidadian guppy*. ((. s. module)., Ed.) Oxford: Oxford University Press.
- Nicoletto, P. F. (1993). Female sexual response to condition-dependent ornaments in the guppy, *Poecilia reticulata*. *Animal Behaviour*, 46(3), 441-450.
- Parks, D. H., Mankowski, T., Zangooei, S., Porter, M. S., Armanini, D. G., Baird, D. J., . . . Beiko, R. G. (2013). GenGIS 2: Geospatial analysis of traditional and genetic biodiversity with new gradient algorithms and an extensible plugin in framework. *PLoS One*.
- Reche, P. A., & Reinherz, E. L. (2007). Definition of MHC supertypes through clustering of MHC peptide-binding repertoires. In D. R. Flower (Ed.), *Immunoinformatics* (pp. 163-173). Humana Press.
- Reznick, D., Butler IV, M. J., & Rodd, H. (2001). Life-history evolution in guppies. VII. The comparative ecology of high-and low-predation environments. *The American Naturalist*, 157(2), 126-140.
- Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, 27(4), 592-593.
- Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, 2(1).
- Van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, 1, 1-30.