Journal of Heredity, 2022, **113**, 525–537
https://doi.org/10.1093/jhered/esac030
Advance access publication 11 June 2022
Original Article

Original Article

Sequence Diversity and Differences at the Highly Duplicated MHC-I Gene Reflect Viral Susceptibility in Sympatric Pinniped Species

Alayna K. Gigliotti, W. Don Bowen, Michael O. Hammill, Wendy B. Puryear, Jonathan Runstadler, Frederick W. Wenzel, and Kristina M. Cammen

From the School of Marine Sciences, University of Maine, Orono, ME 04469, US (Gigliotti and Cammen); Bedford Institute of Oceanography, Dartmouth, NS B2Y 4A2, Canada (Bowen); Fisheries and Oceans Canada, Maurice Lamontagne Institute, Mont-Joli, QC G5H 3Z4, Canada (Hammill); Department of Infectious Disease and Global Health, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA 01536, US (Puryear and Runstadler); and Protected Species Branch, NOAA, NMFS, Northeast Fisheries Science Center, Woods Hole, MA 02543, USA (Wenzel).

Address correspondence to K. M. Cammen at the address above, or e-mail: kristina.cammen@maine.edu

Corresponding Editor: C. Scott Baker

Abstract

Differences in disease susceptibility among species can result from rapid host–pathogen coevolution and differences in host species ecology that affect the strength and direction of natural selection. Among 2 sympatric pinniped species that differ in sociality and putative disease exposure, we investigate observed differences in susceptibility through an analysis of a highly variable, duplicated gene family involved in the vertebrate immune response. Using high-throughput amplicon sequencing, we characterize diversity at the 2 exons that encode the peptide binding region of the major histocompatibility complex class I (*MHC-I*) gene in harbor ($N = 60$) and gray ($N = 90$) seal populations from the Northwest Atlantic. Across species, we identified 106 full-length exon 2 and 103 exon 3 sequence variants and a minimum of 11 duplicated *MHC-I* loci. The sequence variants clustered in 15 supertypes defined by the physiochemical properties of the peptide binding region, including a putatively novel Northwest Atlantic *MHC-I* diversity sublineage. Trans-species polymorphisms, dN/dS ratios, and evidence of gene conversion among supertypes are consistent with balancing selection acting on this gene. High functional redundancy suggests particularly strong selection among gray seals at the novel Northwest Atlantic *MHC-I* diversity sublineage. At exon 2, harbor seals had a significantly greater number of variants per individual than gray seals, but fewer supertypes. Supertype richness and private supertypes are hypothesized to contribute to observed differences in disease resistance between species, as consistently, across the North Atlantic and many disease outbreaks, gray seals appear to be more resistant to respiratory viruses than harbor seals.

Key words: gray seal, harbor seal, major histocompatibility complex, MHC supertype

Host defense mechanisms have long been recognized as under strong selective pressure (Haldane 1949) and more recently as a common target of rapid evolution (Thompson 1998). Selective pressures exerted by the pathogen on its host, and vice versa, can drive rapid host–pathogen coevolution (sensu the red queen hypothesis) (van Valen 1973; Woolhouse et al. 2002; Papkou et al. 2019). This process of coadaptation motivates elevated rates of molecular evolution in host immunity genes (Paterson et al. 2010) and promotes genetic diversity (Lively and Apanius 1995). At any point in time during this dynamic process of evolution, a host's genetic makeup significantly influences their disease susceptibility. It thus follows that differences in disease susceptibility among host species may be a product of the factors that affect disease exposure and outcome, and ultimately evolution on host defense genes. Here, we investigate the molecular ecology of variation in disease susceptibility between 2 sympatric pinniped species through a comparison of genetic diversity and evidence of selection at a highly variable, duplicated gene family involved in the vertebrate immune response.

In general, our understanding of marine mammal host–pathogen dynamics is limited compared with terrestrial mammals (Kim et al. 2005). Yet, infectious disease outbreaks have increased among marine mammals in the past several decades (Gulland and Hall 2007; Simeone et al. 2015). Among marine mammals, harbor seals (*Phoca vitulina*) have experienced the greatest number of mass mortality events associated with infectious disease in the past half-century (Sanderson and Alexander 2020). In contrast, over the same period, sympatric gray seal (*Halichoerus grypus*) populations have experienced fewer mass mortality events and dramatically lower mortality rates during events that affect both species (Puryear et al. 2021).

Harbor and gray seal mass mortality events across the North Atlantic have been attributed to viral outbreaks of

Received February 18, 2022; Accepted June 8, 2022

influenza A virus (IAV) and phocine distemper virus (PDV). In 1988, European populations of harbor seals experienced a regionally variable death rate from PDV of 10–60%, affecting approximately 18 000 individuals over a 16-week span (Heide-Jørgensen and Härkönen 1992; Härkönen et al. 2006). Gray seals were also affected but to a much lesser extent, totaling about 10% of total reported deaths, with most gray seals testing seropositive for the virus but remaining asymptomatic (Pomeroy et al. 2005). A second European PDV outbreak occurred in 2002, following similar patterns to that of 1988 (Harding et al. 2002). The first documented European outbreak of IAV in seals occurred in 2014, also affecting predominantly harbor seals, with few documented cases of gray seal infection (Zohari et al. 2014; Bodewes et al. 2015). During this period, several PDV and IAV outbreaks were also documented in the Northeast United States (Duignan et al. 1993; Earle et al. 2011; Anthony et al. 2012; Siembieda et al. 2017). Most recently, in 2018 and 2019, outbreaks of both PDV and IAV were codetected in this region, resulting in the reported deaths of over 3000 individuals, primarily harbor seals infected with PDV (NOAA Fisheries 2021).

Consistently, across many outbreaks and geographic regions, gray seals appear to be more resistant to respiratory viruses than harbor seals. It has been suggested that Arctic phocid seals, such as the harp seal (*Pagophilus groenlandicus*), may act as a primary host while gray seals act as reservoirs that episodically transmit diseases to more susceptible harbor seal populations (Markussen and Have 1992; Hall et al. 2006; Puryear et al. 2016; Jo et al. 2018; Puryear et al. 2021). The evolution of such variable disease resistance in sympatric seal species is plausible given differences in species ecology that probably affect frequency of disease exposure, in particular sociality, which affects contact rates and thus disease transmission potential (Sanderson et al. 2013). Disease transmission within seals is likely linked to seal haul-out sites, where the animals aggregate on land and both interspecies and intraspecies interactions can occur. These aggregations, which occur during breeding/pupping and molting seasons, are typically larger for gray seals than harbor seals (Katona et al. 1993). The 2 species also vary in the seasonality of these important life history events, such that they may have different exposure risks and be experiencing different levels of physiological stress, depending on when a pathogen enters the population. Population monitoring of seal haul-outs across the North Atlantic over the past few decades has revealed dramatic changes in density, with the most common trend being one of gray seal increases and harbor seal declines (Carroll et al. 2020; den Heyer et al. 2020; Wood et al. 2020). In the Northwest Atlantic, these local patterns reflect regional trends of ongoing rapid gray seal population growth and recent declines in harbor seal population growth rates (Sigourney et al. 2021), following historical bottlenecks and contemporary recovery of both species (Cammen et al. 2019).

The immunogenetic variation of the harbor and gray seals has been previously investigated through a candidate gene study of 8 genes involved in cellular interactions with morbilliviruses and full-length sequencing of the major histocompatibility complex (MHC) class I genes. McCarthy et al. (2011) compared candidate gene diversity across European harbor seal populations that differed in mortality rates during the 1988 and 2002 PDV epidemics, but found little association between genetic variation and disease status. Hammond et al. (2012) characterized variation at the *MHC-I* genes in a small sample size of harbor and gray seals from the United Kingdom and found preliminary evidence to suggest that gray seals are more diverse in terms of gene content variation and allelic polymorphism. In this study, we used high-throughput amplicon sequencing to compare levels of *MHC-I* diversity between gray and harbor seals on a population scale in the Northwest Atlantic.

The MHC is a polymorphic, multi-gene family that encodes cell-surface receptors integral in antigen presentation, an early step in the vertebrate immune response to foreign bodies such as viruses (Piertney and Oliver 2006). Extensive polymorphism has been documented in MHC proteins, in particular, within the peptide binding region (PBR) where antigens bind to the MHC receptor for presentation on the cell surface (Hughes and Nei 1989; Neefjes et al. 2011). This diversity within the PBR is often observed to be conserved across species (i.e., trans-species polymorphisms) as well as within species among alleles (i.e., gene conversion), highlighting its functional importance and selective maintenance (Klein et al. 2007; Spurgin et al. 2011). Across many systems, MHC diversity is shaped by heterozygote advantage, a form of balancing selection, where an increase in fitness results from having 2 different alleles at a given locus (Doherty and Zinkernagel 1975; Osborne et al. 2015). Natural selection for variation in this gene family has furthermore resulted in frequent gene duplications, that provide the potential benefit of expressing more than 2 variants of MHC receptors. Diversity in MHC haplotypes driven by a combination of allelic polymorphism and gene content variation is common among nonhuman mammalian species (Rada et al. 1990; Boyson et al. 1996; Ballingall et al. 2008; Tanaka-Matsuda et al. 2009; Tallmadge et al. 2010).

Previous studies in pinnipeds have primarily focused on *MHC-II*, the class of MHC molecules that recognize primarily extracellular pathogens. These studies report variable levels of genetic diversity across species and populations and attribute this variation to factors including demographic history (Hoelzel et al. 1999; Weber et al. 2004; Lau et al. 2015), population structure, habitat substrate, and pathogen presence (Cammen et al. 2011). Relatively fewer studies in pinnipeds have focused on *MHC-I*, the class of MHC molecules that recognize primarily intracellular pathogens such as viruses, and those that do have reported low genetic diversity compared to terrestrial mammals (Slade 1992). For example, the Hawaiian monk seal (*Monachus schauinslandi*), a species that has experienced severe historical population declines, exhibits low *MHC-I* sequence variation and allelic richness (Aldridge et al. 2006).

In this study, we characterized sequence diversity of the PBR-containing exons of *MHC-I* in harbor and gray seal populations from the Northwest Atlantic. We assessed multiple lines of evidence of selection, including 1) the presence of trans-species polymorphisms; 2) the ratio of nonsynonymous to synonymous mutations (dN/dS); 3) gene conversion and functional redundancy of supertypes; and 4) levels of genetic differentiation at putative MHC alleles and supertypes among sampling sites. We compare our findings between gray and harbor seals and consider the implications of *MHC-I* diversity for disease resistance among these pinniped species that experience variable mortality rates during episodic outbreaks of respiratory viruses.

Methods

Samples

Tissue samples previously collected from gray seals (*Halichoerus grypus atlantica*) and harbor seals (*Phoca vitulina vitulina*) sampled at several sites across the Northwest Atlantic were provided for this study by the National Oceanic and Atmospheric Administration (NOAA) Northeast Fisheries Science Center (NEFSC), Tufts University, and Department of Fisheries and Oceans, Canada (DFO). Tissue biopsies from weaned gray seal pups were collected during the winter breeding seasons of 2013–2016 on Muskeget Island in Massachusetts, United States of America (41.334°N, 70.294°W, N = 30), Sable Island in Nova Scotia, Canada (43.934°N, 59.915°W, N = 30) and Pictou Island in the Gulf of St. Lawrence, Canada (45.815°N, 62.562°W, N = 30). Live harbor seal pups were similarly sampled in 2016 on Bic Island in the St. Lawrence Estuary, Canada (48.399°N, 68.861°W, N = 30). Harbor seal samples from the Northeast United States (N = 30) were collected by the Northeast Fisheries Observer Program from individuals by caught in commercial fisheries between 2013–2015.

Molecular Methods

Genomic DNA was extracted from ~10 to 20 mg of seal skin from each individual using the Qiagen DNeasy Blood & Tissue Extraction kit and the manufacturer's spin-column protocol for Purification of Total DNA from Animal Tissues, with minor modifications. 20 µL of 1 M DTT was added during the digestion phase and tissues were incubated at 56 °C and 850 rpm overnight to fully digest. Following digestion, 4 µL of RNAse (100 mg/mL) was added to remove RNA from the targeted genomic DNA.

Independent PCRs were performed to amplify exons 2 and 3 of *MHC-I*, which comprise the peptide binding region (PBR). PCR primers were designed from an alignment of gray and harbor seal *MHC-I* sequences (GenBank Accession No: JX218867–JX218936 from Hammond et al. (2012)) and evaluated using primer3web v4.1.0 (Untergasser et al. 2012). Forward primer PvLAex2F (5′-GGCTCCCACTCCMTGARGT-3′) from Hammond et al. (2012) and a new reverse primer PvLAex2Rb (5′-GKCCTCGCTYTGGTTGTAG-3′) amplified a 272-bp fragment of exon 2. Primers PvLAex3F (5′-GGC GGGGCCAGGGTCT-3′) and PvLAex3R (5′-CCGCGGCCC CTGGTA-3′) amplified a 304-bp fragment of exon 3. Illumina Nextera adapter sequences were added onto the 5′ end of the primers to facilitate downstream sequencing.

PCRs were conducted in a total volume of 20 µL that contained 2 µL of DNA, 0.2 mM dNTPs, 0.3 µM of each primer, 3% DMSO, 0.4 units of Phusion polymerase (Life Technologies), and 1× Phusion HF Buffer (7.5 mM $MgCl_2$ [1.5 mM at 1× dilution]). The PCR profile for exon 2 consisted of an initial denaturation step at 98 °C for 1 min followed by 28 cycles of 98 °C for 10 s, 68 °C for 25 s, 72 °C for 10 s, and a final extension step of 72 °C for 10 min. The PCR profile for exon 3 consisted of an initial denaturation step of 100 °C for 2 min, followed by 28 cycles of 100 °C for 10 s, 70 °C for 25 s, 72 °C for 10 s, and a final extension step of 72 °C for 10 min.

Following PCR, the KAPA Pure Beads (KR1245-v3.16) magnetic bead protocol was performed on all samples using Beckman Counter Agencourt AMPure XP beads to remove any PCR impurities such as small, unwanted fragments, and primer-dimers. A 1.8× bead-to-sample volumetric ratio was used for size selection of our ~300 bp products. Agarose gel electrophoresis was performed on all samples to confirm the size of the final product and approximate concentration prior to sequencing. Dilutions were performed to products as necessary to achieve uniform concentration. Samples were sequenced by the Hubbard Center for Genome Studies in New Hampshire, USA using rapid, paired-end sequencing (2 × 250) on an Illumina HiSeq2500. High-throughput short-read sequencing is a cost-effective and efficient approach for population scale studies of MHC (Sommer et al. 2013). However, with this approach, we cannot associate exon 2 and exon 3 variants that are sequenced individually so the 2 sets of sequences must be analyzed separately.

Genotyping Pipeline

Sequencing reads were first processed using the program USEARCH v11 (Edgar 2010) and then a pipeline adopted from Sommer et al. (2013) was implemented in R to generate a list of putative alleles per individual. Specific details for the genotyping pipeline are provided in the Supplementary Methods, and briefly summarized here. Within USEARCH, the UPARSE-OTU algorithm was used to identify clusters of identical sequences (i.e., putative alleles) from least to most frequent, and UNOISE was used to denoise the filtered set of amplicon reads. Reads were then further processed in R to identify a list of putative alleles, putative artifacts, and unclassified variants for each exon using a decision tree modified from Sommer et al. (2013) (Supplementary Figure S1). Finally, a 1% filter was implemented on the entire set of remaining reads so that only putative alleles present in at least 1% of the reads per individual were retained for subsequent analyses. Because these putative alleles do not meet the minimum requirement for naming a new *MHC-I* allele (must include complete sequence of exon 2 and exon 3 [Marsh et al. 2010]), we refer to these primarily as *MHC-I* sequence variants throughout the remainder of the paper.

The genotyping pipeline was assessed through an analysis of technical replicates and allele discovery saturation curves. Technical replicates included 24 Northeast US harbor seals, and 8 Muskeget Island, 8 Sable Island, and 8 Gulf of St. Lawrence gray seals, which were each amplified and sequenced twice in independent reactions. Weighted and unweighted sequence overlap among replicates were calculated following Wen et al. (2017).

To test if the read depth achieved was sufficient to reach saturation in allele discovery, we compared individual allele saturation curves to a simulated individual with 240 000 reads and equal read proportions representing 24 alleles (greater than the maximum number of putative alleles observed per individual). Curves were generated by drawing 100 random subsamples of 2^{-x} times the total number of reads per individual up to x = 11 for all locations except harbor seals from the St. Lawrence Estuary, and up to x = 15 for this latter population, which had a significantly larger number of reads than all other locations. The saturation curve from the simulated data was compared to the individual curves from each sample to determine the minimum read depth for saturation of allele discovery. An additional filter for read depth was applied to exon 3 based on this analysis; any sample with fewer than 400 reads was removed from the dataset.

Analysis of Genetic Diversity

Putative allele sequences from each exon were aligned using CodonCode Aligner v8.0.1. Existing full-length gray and harbor seal *MHC-I* gene sequences (GenBank Accession No: JX218867–JX218936) were trimmed to match our exon 2 and exon 3 regions and added to each alignment. FASTA files of the alignments were then exported and further analysis was conducted in MEGA v10.1.8 (Kumar et al. 2018) and R.

Evidence for trans-species polymorphisms was assessed through a comparison of sequence diversity between harbor and gray seals. The relationship among sequence variants was explored using the maximum likelihood method and the Tamura–Nei (Tamura and Nei 1993) or Jones–Taylor–Thornton (Jones et al. 1992) model for nucleotide and amino acid sequences, respectively. Nodes of the final tree were evaluated as the proportion of trees among 500 bootstrap replicates that showed similar clustering of tips. Newick tree files were exported for visualization using the ggtree package in R (Yu 2020).

dN/dS ratios among sequence variants were calculated for each exon within each species using the Nei–Gojobori method (Nei and Gojobori 1986) with a Jukes-Cantor correction. Sites in the PBR were determined as in Aldridge et al. (2006) assuming homology with the human *HLA-A2* molecule (Saper et al. 1991), and dN/dS ratios were evaluated within and outside the PBR.

MHC supertypes were defined as in Lighten et al. (2017) by grouping sequence variants with similar physiochemical properties of amino acids in the PBR. Physiochemical similarity was evaluated based on hydrophobicity (*z1*), steric bulk (*z2*), polarity (*z3*), and electronic effects (*z4* and *z5*), characterized for each amino acid by Sandberg (1998). Sequence variants were clustered based on their PBR amino acid scores across these 5 metrics using discriminant analysis of principle components (DAPC), implemented in the *adegenet* package in R (Jombart 2008). DAPC relies on selection of the optimal number of clusters, *k*, based on Bayesian information criterion (BIC) values associated with each value of *k*. In practice, one typically selects the minimum *k* after which BIC increases, or the elbow in a BIC plot. However, because our BIC plots had wide elbows (Supplementary Figure S2), we evaluated group assignment across a range of *k* from 10 to 20 and manually assigned sequence variants to supertypes based on consensus following visual inspection of the DAPC cluster assignments. The functional redundancy (S_r) of MHC supertypes was calculated by dividing the number of sequence variants associated with each supertype by the number of its unique PBR amino acid sequences.

Genetic differentiation among sampling sites within species was estimated as pairwise Jost's D_{est} calculated using MHC sequence variant and supertype frequencies in the SuperTypeR package, as described in Lighten et al. (2017). This method, which compares population-level allelic frequencies, is appropriate for data like MHC sequences generated by next-generation sequencing that cannot be assigned to loci (Jost 2008).

Sequence variants were named for each exon according to their supertype assignment and using nomenclature similar to Hammond et al. (2012), modified from Marsh et al. (2010). Harbor seal supertypes were named *Phvi-ex2ST*01-ST*15* and *Phvi-ex3ST*01-ST*15*, and gray seal supertypes were named *Hagr-ex2ST*01-ST*15* and *Hagr-ex3ST*01-ST*15*. Distinct sequences within each supertype that differ in the amino acid sequence are designated with 2 digits following a colon (e.g., *Phvi-ex2ST*01:01* and *Phvi-ex2ST*01:02*); sequences that differ only by synonymous mutations are further indicated with an additional 2 digits following a colon (e.g., *Phvi-ex2ST*02:01:01* and *Phvi-ex2ST*02:01:02*).

Results

Assessing the Genotyping Pipeline

Technical replicates and simulations were used to test the efficacy of our genotyping pipeline, and the results are presented in detail in the Supplementary Results. We found that all populations showed high overlap in both exon 2 and exon 3 genotypes between technical replicates in weighted and unweighted comparisons of alleles (Supplementary Figure S3), supporting precision of the genotyping pipeline. Our simulations revealed apparent saturation in allele discovery by 400 reads (Supplementary Figure S4). This saturation point was also observed in the actual sequencing data for most individuals for both exon 2 (Supplementary Figure S5) and exon 3 (Supplementary Figure S6). A read depth of 400 was therefore used as a minimum threshold for inclusion in subsequent analyses, and 15 exon 3 samples with a read depth below 400 reads were excluded. Our sequencing resulted in variable read depth across seal sampling locations, between species, and between exons (Supplementary Figure S7), probably due to differences in sample quality and reaction optimization. However, these findings suggest that our pipeline is robust to differences in read depth, above a minimum read depth of 400, and support our ability to compare allelic diversity among populations and between species.

Analysis of Genetic Diversity

Following the filtering steps described above, we identified a total of 106 full-length (232 bp) exon 2 sequence variants and 103 full-length (272 bp) exon 3 sequence variants. One shorter exon 2 sequence variant (187 bp) was also identified, but is assumed to originate from a pseudogene due to multiple large deletions. Ten exon 2 variants were shared by both species; 44 additional variants were observed among 86 gray seals and 52 additional variants were observed among 58 harbor seals. Only one exon 3 variant was shared by both species; 65 additional variants were observed among 76 gray seals and 37 additional variants were observed among 45 harbor seals.

The total number of sequence variants per individual ranged from 7 to 21 for both exons 2 (mean: 13.05 ± 0.22 SE) and 3 (mean: 11.37 ± 0.19 SE), suggesting a minimum of 4–11 duplicated MHC class I loci per individual. Individuals had a greater average number of exon 2 variants than exon 3 variants for both harbor (exon 2: mean = 13.95 ± 0.36 SE, range = 9–21; exon 3: mean = 11.04 ± 0.29 SE, range = 7–15, $t = 12.00$, $P < 1 \times 10^{-14}$) and gray seals (exon 2: mean = 12.47 ± 0.25 SE, range = 7–16; exon 3: mean = 11.57 ± 0.25 SE, range = 7–21, t = 4.14, $P < 1 \times 10^{-4}$). The maximum number of sequence variants per individual was greater for harbor seals for exon 2, but greater for gray seals for exon 3. Harbor seals were also observed to have a significantly greater average number of exon 2 variants per individual than gray seals ($t = -3.37$, $P < 0.01$) (Figure 1A). There

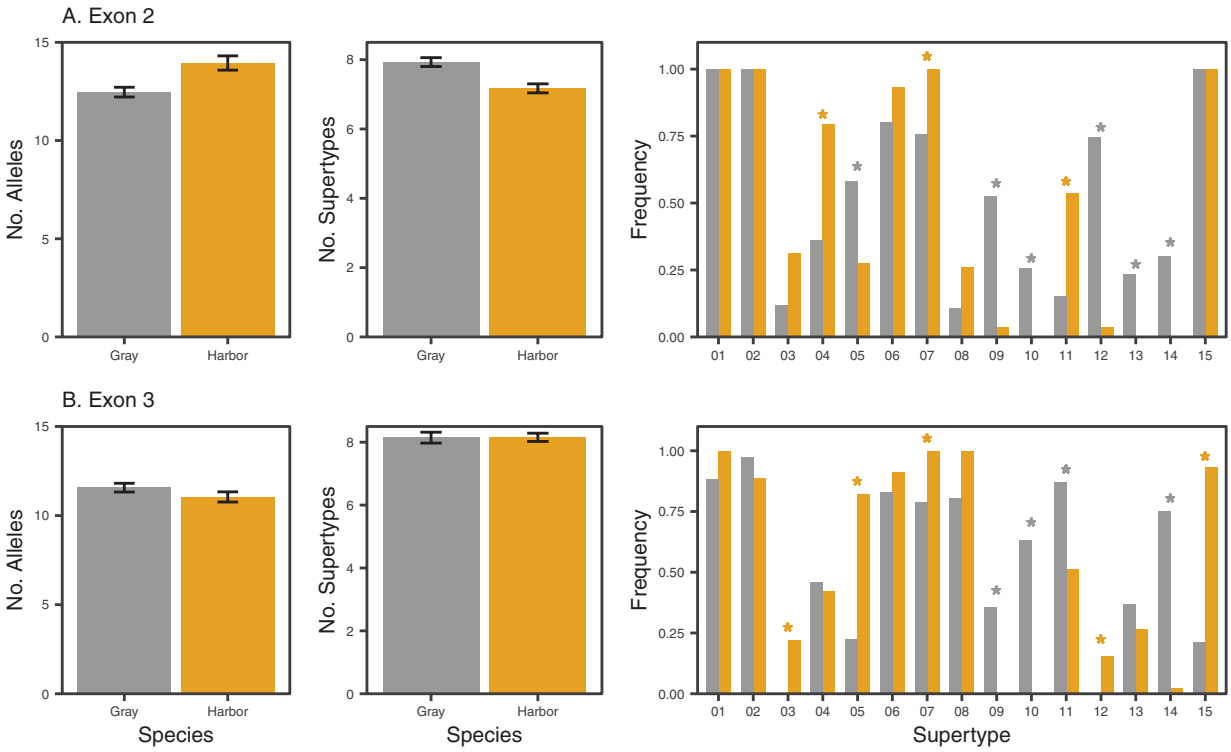

Figure 1. Comparison of the average number of *MHC-I* (**A**) exon 2 and (**B**) exon 3 alleles and supertypes per individual, as well as supertype frequency, between gray and harbor seals. For exon 2, harbor seals had a significantly greater average number of alleles ($t = -3.37$, $P < 0.01$) and fewer supertypes ($t = 4.15$, $P < 0.0001$) per individual than gray seals. No significant difference was observed between species for exon 3 (alleles: $t = 1.38$, $P = 0.17$; supertypes: $t = -0.45$, $P = 0.66$). Stars indicate a supertype observed in only one species or found in significantly different proportions between species, following a Bonferorri correction for multiple tests (See online version for color figure).

was no significant difference between species in the average number of exon 3 variants per individual ($t = 1.38$, $P = 0.17$) (Figure 1B).

Maximum likelihood trees revealed some monophyletic clades for both exon 2 and 3, though most clades contain both gray and harbor seal sequences, suggestive of trans-species polymorphisms (Figure 2). The dN/dS ratio was found to be significantly greater than 1 within, but not outside, the PBR for exon 2 in both species (Table 1), suggestive of balancing selection at this region of the gene that is critical to antigen recognition. The dN/dS ratio was highest within the PBR for exon 3 as well, though the ratio was not significantly greater than 1 for either species.

To further characterize functional variation in our sequences, DAPC was used to cluster sequence variants based on the physicochemical properties of amino acids in the PBR. Among the variants identified in Northwest Atlantic seals in this study, this process identified 15 exon 2 supertypes (*ex2ST*01-15*) and 15 exon 3 supertypes (*ex3ST*01-15*) (Figure 3). Overall, supertypes corresponded well with the putative MHC loci and lineages previously identified by Hammond et al. (2012) using the degree of sequence similarity across the complete *MHC-I* gene; in other words, the majority of supertypes contained sequences from only 1 putative MHC locus per species as previously described by Hammond et al. (2012) (Table 2).

Hammond et al. (2012) defined 3 lineages among the full-length *MHC-I* sequences characterized in their study: A pseudogene lineage; the N*01 lineage, which they classified as a candidate nonclassical *MHC-I* gene; and 2 diversity sublineages. From our study, *ex2ST*01* and *ex3ST*01* corresponded to the N*01 lineage, *ex2ST*02* and *ex3ST*02* included primarily loci from diversity sublineage 1, and all other supertypes, with the exception of *ex2ST*15* (see below), included loci from diversity sublineage 2. Despite general one-to-one correspondence, some supertypes associated with the diversity sublineage 2 contained multiple Hammond loci, and alleles from some loci were found distributed across multiple supertypes. Furthermore, for only half of the supertypes were the loci associated with the exon 2 supertype linked to the loci associated with the exon 3 supertype in a one-to-one relationship (e.g., *PhviN*07* and *HagrN*07* associate with ex2ST*07 and ex3ST*07) (Table 2). This inability to consistently link exon 2 and exon 3 data even at the level of supertype suggests at least partially independent evolution of these 2 segments encoding the PBR. It further supports our treatment of the 2 datasets in separate analyses, as our read length precludes phasing the haplotypes of this duplicated loci, which exhibits interindividual copy number variation and a tendency for gene conversion.

One supertype, *ex2ST*15*, included only new sequences identified in the Northwest Atlantic and did not cluster with any of the sequences previously described by Hammond et al. (2012). Most individuals also exhibited more than 2 variants from this supertype, suggesting that it represents multiple loci and perhaps a new Northwest Atlantic *MHC-I* lineage. This supertype had notably high functional redundancy among gray seals (Table 2). High functional redundancy indicates a

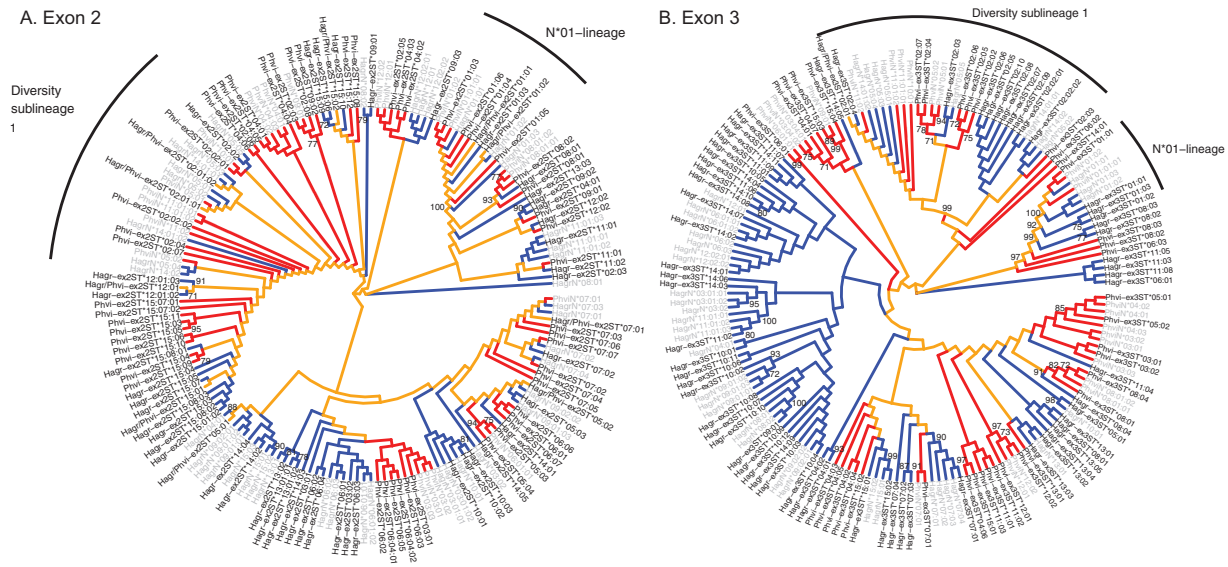

Figure 2. Maximum likelihood trees depicting the relationship among *MHC-I* (**A**) exon 2 and (**B**) exon 3 nucleotide sequences observed among gray and harbor seals in the Northwest Atlantic (black tip labels, current study) and Northeast Atlantic (gray tip labels, Hammond et al. (2012)). Branches are colored by species (blue: Gray seal [Hagr]; red: Harbor seal [Phvi]; purple: Both species). Bootstrap values above 70 are shown (See online version for color figure).

Table 1. Ratio of nonsynonymous (dN) to synonymous (dS) mutations within the *MHC-I* exon 2 and exon 3 sequences for Northwest Atlantic gray and harbor seals. *P* values are associated with a codon-based *Z* test of selection for deviation from neutrality (dN/dS = 1)

	Gray seals				Harbor seals			
	dN	dS	dN/dS	*P*	dN	dS	dN/dS	*P*
Exon 2	0.13	0.11	1.18	0.56	0.14	0.10	1.40	0.21
PBR	0.61	0.17	3.59	< 0.01	0.58	0.17	3.41	< 0.01
Non-PBR	0.08	0.11	0.73	0.42	0.09	0.09	1.00	0.92
Exon 3	0.07	0.06	1.10	0.59	0.08	0.06	1.33	0.42
PBR	0.38	0.26	1.46	0.33	0.40	0.29	1.38	0.40
Non-PBR	0.03	0.04	0.75	0.42	0.03	0.04	0.75	0.56

large proportion of the unique variants within this supertype had identical PBR amino acid sequences and suggests that this supertype's functionality has been preserved over time as new alleles were generated by mutations that do not affect the functionality of the PBR. Low functional redundancy (1–2) was observed among all other supertypes across both exons and both species.

Maximum likelihood trees showed that most supertypes are not monophyletic, meaning that sequences from a given supertype were found distributed across the tree built using amino acid sequences (Supplementary Figures S8 and S9). The only exceptions were supertypes with very few sequence variants and supertypes that correspond to the N*01 lineage identified by Hammond et al. (2012). For exon 3, one other exception was *ex3ST*02*, which corresponds to Hammond et al.'s (2012) diversity sublineage 1.

The total number of supertypes per individual ranged from 5 to 11 for both exon 2 (mean: 7.62 ± 0.10 SE) and exon 3 (mean: 8.21 ± 0.12 SE). For both exons, the number of supertypes per individual was significantly correlated with the number of sequence variants per individual (exon 2: $F(1,142) = 73.63$, $P < 1 \times 10^{-13}$, $R^2 = 0.34$; exon 3: $F(1,119) = 45.46$, $P < 1 \times 10^{-9}$, $R^2 = 0.28$). Every individual contained at least one variant from the supertype associated with the diversity sublineage 1 for both exons (*ex2ST*02* and *ex3ST*02*) and from the newly proposed Northwest Atlantic *MHC-I* lineage, *ex2ST*15*. Similarly, all individuals for exon 2, and almost all individuals for exon 3, contained at least one variant from the supertype associated with the nonclassical *MHC-I* lineage (*ex2ST*01* and *ex3ST*01*). Finally, all individuals also exhibited at least 2 variants from a supertype associated with the diversity sublineage 2. Among harbor seals, for both exons, this included at least 1 variant from *ex2ST*07* and *ex3ST*07*, which was associated with Hammond et al.'s (2012) *HagrN*07* and *PhviN*07* loci.

Harbor seals exhibited on average fewer exon 2 supertypes per individual than gray seals (harbor: mean = 7.17 ± 0.13 SE, range = 6–10; gray: mean = 7.93 ± 0.13 SE, range = 5–11; $t = 4.15$, $P < 0.0001$), but there was no significant difference in the average number of exon 3 supertypes per individual between species (harbor: mean = 8.27 ± 0.13 SE, range = 7–10; gray: mean = 8.17 ± 0.17 SE, range = 5–11; $t = -0.45$, $P = 0.66$) (Figure 1). Several supertypes were observed in significantly different frequency between species. Three exon 2 supertypes and 2 exon 3 supertypes were observed only in gray seals, while 2 exon 3 supertypes were observed only in harbor seals.

Figure 3. Relative amino acid composition across *MHC-I* (**A**) exon 2 and (**B**) exon 3 variants comprising 15 harbor and gray seal supertypes (ST1–ST15) shown for residues within the peptide binding region. Colors denote the types of amino acid as follows, basic (H, K, R): Blue; acidic (D, E): Red; polar (C, N, Q, S, T, Y): Yellow; hydrophobic (A, F, G, I, L, P, M, V, W): Green. Additional information about these supertypes can be found in Table 2 (See online version for color figure).

Little difference was observed in supertype or variant frequency among the sampling sites for either species (Figure 4). Among gray seals, pairwise genetic differentiation, estimated as Jost's D_{est} using sequence variant frequencies, was lower for both exons between the Gulf of St. Lawrence and Sable Island than between either Canadian sampling site and the US sampling site, consistent with isolation by distance (Table 3). A similar pattern was observed in calculations of Jost's D_{est} using supertype frequencies for exon 2, but not exon 3. Across both exons and both species, pairwise comparisons of variant frequencies showed consistently higher estimates of genetic differentiation than supertype frequencies.

Discussion

We report on sequence variation at a highly duplicated and diverse *MHC-I* gene family among 2 sympatric pinniped species, gray and harbor seals, that vary in resistance to viral outbreaks. As expected, both species show evidence of gene duplication and balancing selection, particularly in regions of the gene that encode the antigen-recognition groove, with putative implications for immune response. We compare our population scale amplicon sequencing data from the Northwest Atlantic to a previous study involving full-length sequencing of the *MHC-I* genes in a small sample size of harbor and gray seals from the United Kingdom (Hammond et al. 2012). This comparison revealed similar lines of evidence suggesting greater functional *MHC-I* variation among gray seals than harbor seals across the North Atlantic, but also identified a putatively novel Northwest Atlantic *MHC-I* diversity sublineage. The comparison of *MHC-I* diversity between species and across an ocean basin provides insights into how evolution of the immune system may shape contemporary responses to episodic outbreaks of respiratory viruses across the North Atlantic.

Despite its extraordinarily high degree of genetic diversity, studies that compare MHC diversity among species often report shared genetic variants, even among evolutionarily divergent species (Klein et al. 2007). Such long-lasting trans-species polymorphisms can result from balancing selection for functionally important variants (Arden and Klein 1982), and they are common not only among the MHC loci, but also in other immune genes (Těšický and Vinkler 2015). Accordingly, in this study, we find some components of genetic diversity to be shared between gray and harbor seals, while others are unique. About 10% of the variants that we identified at exon 2 were found in both gray and harbor seals, though only a single exon 3 variant was shared by both species. The proportion of supertypes, defined based on similarities in physiochemical properties of the PBR, shared by both species was

Table 2. Harbor and gray seal *MHC-I* supertypes defined through Discriminant Analysis of Principal Components using the physiochemical properties of the amino acids in the peptide binding region. Associated loci and lineages refer to a comparison of the inferred supertypes with putative loci and lineages (candidate nonclassical N*01 lineage, diversity sublineage 1 and 2) identified by Hammond et al. (2012). The table includes the total number of sequence variants associated with each supertype, the range and mean number of variants per individual, the percent of individuals with variants from that supertype, and the functional redundancy (S_r) of each supertype

			Gray seal				Harbor seal			
Supertype	Assoc. loci	Assoc. lineage	Total no. variants	Variants per individ (mean)	% Individs present	S_r	Total no. variants	Variants per individ (mean)	% Individs present	S_r
*ex2ST*01*	*HagrN*01, PhviN*01*	N*01	3	1–2 (1.65)	100.00	1.00	6	1–3 (1.21)	100.00	1.00
*ex2ST*02*	*HagrN*05, HagrN*14, PhviN*05, PhviN*09*[a], *PhviN*10, PhviN*11, PhviN*12*[a]	Div. 1	4	1–3 (1.66)	100.00	2.00	10	1–5 (3.22)	100.00	2.00
*ex2ST*03*	*PhviN*03*	Div. 2	2	0–1 (0.12)	11.63	1.00	1	0–1 (0.31)	31.03	1.00
*ex2ST*04*	*PhviN*04*	Div. 2	1	0–1 (0.36)	36.05	1.00	6	0–3 (1.07)	79.31	2.00
*ex2ST*05*	*HagrN*04, HagrN*09, PhviN*06*	Div. 2	3	0–3 (1.14)	58.14	1.00	4	0–2 (0.31)	27.59	1.33
*ex2ST*06*	*HagrN*06*	Div. 2	4	0–3 (1.14)	80.23	1.00	8	0–4 (1.91)	93.10	1.60
*ex2ST*07*	*HagrN*07, PhviN*07*	Div. 2	2	0–2 (0.77)	75.58	2.00	7	1–3 (2.02)	100.00	1.40
*ex2ST*08*	*PhviN*08*	Div. 2	1	0–1 (0.10)	10.47	1.00	2	0–2 (0.29)	25.86	2.00
*ex2ST*09*	*HagrN*08*	Div. 2	3	0–2 (0.53)	52.32	1.00	2	0–1 (0.03)	3.45	1.00
*ex2ST*10*	*HagrN*03, HagrN*10*	Div. 2	3	0–1 (0.26)	25.58	1.50	0	0	0.00	N/A
*ex2ST*11*	*HagrN*11*	Div. 2	2	0–1 (0.15)	15.12	1.00	1	0–1 (0.53)	53.45	1.00
*ex2ST*12*	*HagrN*12*	Div. 2	4	0–2 (0.81)	74.42	2.00	2	0–1 (0.03)	3.45	1.00
*ex2ST*13*	*HagrN*13*	Div. 2	4	0–2 (0.24)	23.26	1.33	0	0	0.00	N/A
*ex2ST*14*	*HagrN*15*	Div. 2	5	0–2 (0.33)	30.23	1.25	0	0	0.00	N/A
*ex2ST*15*	none	none	13	2–5 (3.20)	100.00	3.25	13	1–7 (2.98)	100.00	1.30
*ex3ST*01*	*HagrN*01, PhviN*01*	N*01	3	0–3 (0.97)	88.16	1.50	1	1 (1.00)	100.00	1.00
*ex3ST*02*	*HagrN*05, HagrN*14, PhviN*05, PhviN*10, PhviN*11*	Div. 1	11	1–7 (1.93)	100	1.37	7	1–3 (1.93)	100	1.40
*ex3ST*03*	*HagrN*03:02*	Div. 2	0	0	0	N/A	2	0–2 (0.24)	22.22	1.00
*ex3ST*04*	*PhviN*03:04*	Div. 2	2	0–1 (0.46)	46.05	1.00	7	0–4 (0.51)	42.22	1.00
ex3ST*05	*HagrN*10, PhviN*03:01, PhviN*04:01, PhviN*04:02*	Div. 2	1	0–1 (0.22)	22.37	1.00	2	0–2 (0.87)	82.22	1.00
*ex3ST*06*	*PhviN*06*	Div. 2	1	0–1 (0.83)	82.89	1.00	3	0–3 (1.13)	91.11	1.00
*ex3ST*07*	*HagrN*07, PhviN*07*	Div. 2	4	0–2 (0.95)	78.95	1.00	1	1 (1.00)	100	1.00
*ex3ST*08*	*PhviN*08*	Div. 2	4	0–3 (1.11)	80.26	1.33	4	1–2 (1.62)	100	1.00
*ex3ST*09*	*HagrN*08*	Div. 2	1	0–1 (0.36)	35.53	1.00	0	0	0	N/A
*ex3ST*10*	*HagrN*04, HagrN*09*	Div. 2	12	0–8 (1.49)	63.16	1.33	0	0	0	N/A
*ex3ST*11*	*HagrN*11*	Div. 2	9	0–4 (1.62)	86.84	1.12	3	0–2 (0.60)	51.11	1.00
*ex3ST*12*	*PhviN*12*	Div. 2	0	0	0	N/A	2	0–1 (0.16)	15.56	2.00
*ex3ST*13*	*HagrN*13*	Div. 2	5	0–3 (0.42)	36.84	1.00	1	0–1 (0.27)	26.67	1.00
*ex3ST*14*	*HagrN*03, HagrN*06, HagrN*12*	Div. 2	11	0–3 (0.99)	75.00	1.37	1	0–1 (0.02)	2.22	1.00
*ex3ST*15*	*HagrN*15, PhviN*04:03, PhviN*09*	Div. 2	2	0–1 (0.22)	21.06	1.00	4	0–3 (1.69)	93.33	1.00

[a]Member of diversity sublineage 2.

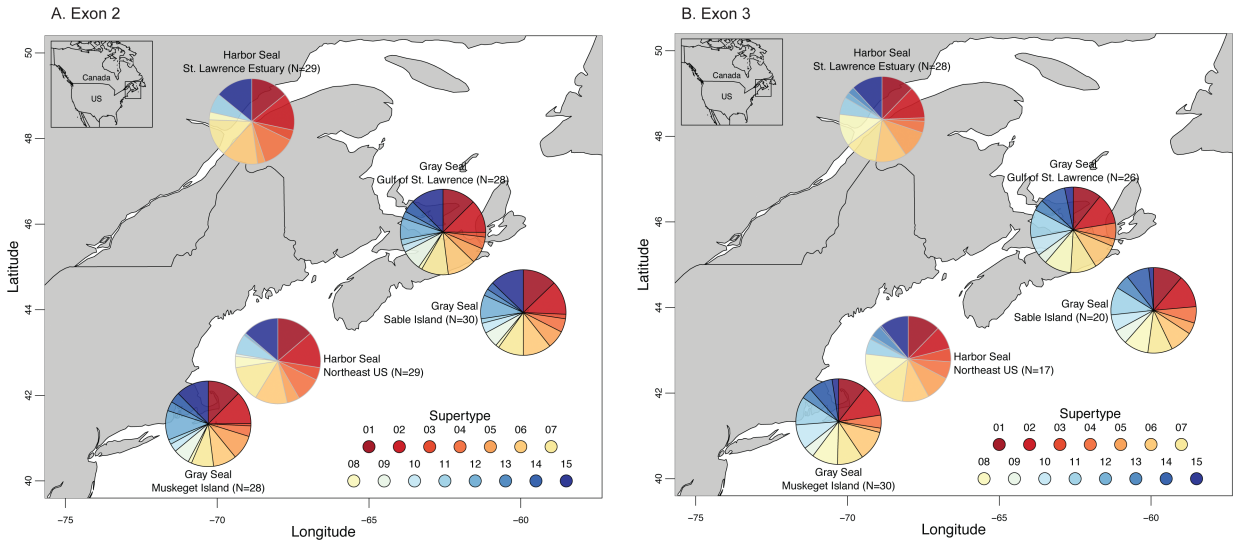

Figure 4. Proportion of gray and harbor seal individuals exhibiting each of the *MHC-I* (**A**) exon 2 and (**B**) exon 3 supertypes at 5 sampling locations in the Northwest Atlantic.

Table 3. Pairwise genetic differentiation, estimated as Jost's D_{est}, of *MHC-I* sequence variants and supertypes among gray and harbor seal sampling sites. Jost's D_{est} calculated by comparing population-level variant frequencies with 1000 bootstraps. Sampling site abbreviations as follows: Gulf of St. Lawrence (GSL), Muskeget Island (MI), Northeast United States (NEUS), Sable Island (SI), St. Lawrence Estuary (SLE)

	exon 2 variant D_{est}	exon 2 supertype D_{est}	exon 3 variant D_{est}	exon 3 supertype D_{est}
Gray seal				
GSL vs. SI	0.017 ± 0.020	0.003 ± 0.010	0.035 ± 0.027	0.015 ± 0.021
MI vs. GSL	0.047 ± 0.026	0.012 ± 0.014	0.050 ± 0.026	0.008 ± 0.014
MI vs. SI	0.031 ± 0.023	0.011 ± 0.013	0.050 ± 0.030	0.024 ± 0.022
Harbor seal				
SLE vs. NEUS	0.019 ± 0.028	0.011 ± 0.011	0.049 ± 0.041	0.029 ± 0.024

much higher, particularly for exon 2, suggesting that convergent selection for certain PBR epitopes has occurred. Our reconstruction of the evolutionary relationships among *MHC-I* variants revealed few clades that were monophyletic within a species, consistent with the evolutionary theory underlying trans-species polymorphisms (Figure 2). A lack of monophyletic supertypes further suggests the possible occurrence of gene conversion, where alleles with divergent sequence variants have retained or evolved similar functionality in antigen recognition (Supplementary Figures S8 and S9). Gene conversion has been previously proposed as a mechanism by which bottlenecked populations can rapidly regenerate *MHC* diversity (Spurgin et al. 2011), and both gray and harbor seal genomes reveal signatures of historical bottlenecks across neutral markers (Cammen et al. 2018).

In addition to the evidence of shared selection, we also observed differences between gray and harbor seals in our proxies of *MHC-I* gene content variation and allelic diversity. Hammond et al. (2012) previously reported finding higher gene copy numbers among gray seals than harbor seals in their small sample size from the United Kingdom. Though we could not assign our variants to genes, at exon 2, we similarly observed greater supertype richness among gray seals, calculated as the average number of supertypes per individual. This finding contrasts with exon 2 allelic richness, which was higher among harbor seals. The apparent incongruity between allelic and supertype richness can be explained by the number of harbor seals that exhibited more than 2 exon 2 variants from a single supertype (Table 2), and hints at functional redundancy among harbor seal variants.

We also observed private supertypes, or supertypes that were observed in only one species. Three exon 2 supertypes and 2 exon 3 supertypes were observed only among gray seals, while 2 exon 2 supertypes were unique to harbor seals. Both supertype richness and private supertypes may be important in explaining observed differences in disease resistance between species. MHC-related fitness can be attributed to having a more diverse arsenal of antigen-recognition proteins or the right type of antigen-recognition protein for a specific pathogen challenge (Doherty and Zinkernagel 1975). Private supertypes observed only in gray seals could play critical roles in recognizing and mounting appropriate immune responses to PDV and/or IAV, the viruses that have caused large-scale mortality events of harbor seals on both east and west coasts of the North Atlantic (Härkönen et al. 2006; Anthony et al. 2012; Bodewes et al. 2015; Siembieda et al. 2017; Puryear et al. 2021) but which are observed circulating in seemingly healthy individuals from the sympatric gray seal populations (Pomeroy et al. 2005; Puryear et al. 2016).

Our analysis identified one new supertype, *ex2ST*15*, that was not observed in any of the gray and harbor seal *MHC-I* sequences previously generated by Hammond et al. (2012)

from individuals sampled in the United Kingdom. It is possible that this supertype is also present in the Northeast Atlantic, but was not found in Hammond's small sample set of 8 harbor seals and 4 gray seals; however, this seems unlikely given that a variant from this lineage (often multiple variants) was found in every individual we sequenced from the Northwest Atlantic. The fact that *ex2ST*15* was identified at least once in every individual, similar to *ex2ST*01*, *ex2ST*02*, and *ex3ST*02*, which are associated with the nonclassical *MHC-I* lineage and diversity sublineage 1, suggests that *ex2ST*15* may represent a new Northwest Atlantic *MHC-I* diversity sublineage. More than 2 variants from this supertype were often observed among our sequenced individuals, suggesting that it represents multiple loci. Though further population scale sequencing in the Northeast Atlantic is necessary for a robust comparison, the cross-basin difference suggested here is interesting to note in the context of the different scale of disease outbreaks observed between the Northeast and Northwest Atlantic, which may suggest differences in seal disease resistance, as well as the recent observation of a Northwest Atlantic variant of PDV (Puryear et al. 2021), which may affect selection on the MHC. Disease outbreaks in the Northwest Atlantic have typically been smaller in the number of individuals affected than outbreaks in the Northeast Atlantic (Härkönen et al. 2006; Siembieda et al. 2017). This difference may be the singular or combined result of multiple factors, including variation in pathogen virulence, habitat factors that affect exposure rates, host density, and host resistance. To address the latter, this novel Northwest Atlantic *MHC-I* diversity sublineage warrants further study to assess the role it may play in contributing to differential host resistance across the North Atlantic.

The putative Northwest Atlantic *MHC-I* diversity sublineage, *ex2ST*15*, was the only supertype that exhibited notably higher functional redundancy, a measure of the proportion of unique variants within a supertype that share identical PBR sequences. This higher functional redundancy, which was observed among gray seals but not harbor seals, suggests that this supertype's functionality has been preserved over time in the Northwest Atlantic as new variants were generated by mutations that do not affect the functionality of the PBR. Overall, its ubiquitous presence among individuals, the high number of variants, and high functional redundancy all suggest that *ex2ST*15* may be experiencing particularly strong balancing selection among gray seals.

Lighten et al. (2017) have previously proposed (and demonstrated in guppies, *Poecilia reticulate,* and *P. obscura*) that higher functional redundancy enables genetic differentiation of *MHC* alleles within supertypes. We tested that hypothesis for *ex2ST*15* by examining the frequency of variants associated with this supertype across sampling sites, but found no difference in the dominant variants across sites for either gray or harbor seals, with high and low functional redundancy, respectively (Supplementary Figure S10). This analysis did reveal that the 2 variants of this supertype that are shared between gray and harbor seals were observed only in harbor seals sampled in the Northeast United States and not among Canadian harbor seals; these 2 variants were also the dominant types observed at all gray seal sampling sites.

Our study included gray seals sampled at 3 breeding colonies, 2 in Canada, and 1 in the United States, and harbor seals sampled at 2 locations, 1 in Canada and 1 in the United States. A previous study that used restriction site-associated DNA (RAD) sequencing to characterize genome-wide variation across the 3 gray seal breeding colonies detected little to no evidence of population structure, suggesting panmixia as a result of recent and ongoing movement of individuals between Canada and the United States (Cammen et al. 2018). Similarly, a genotyping-by-sequencing study of harbor seals sampled worldwide found no significant substructure within the Northwest Atlantic, though fine-scale genetic substructure was detected within other North Atlantic localities (Liu et al. 2022).

In our overall assessment of genetic differentiation at the *MHC-I*, we found little difference in the frequency of sequence variants or supertypes among the sampling sites for either species. Estimates of genetic differentiation were generally higher in comparisons between Canada and the United States than within Canada for gray seals, consistent with an isolation by distance pattern (Table 3). We observed that estimates of genetic differentiation calculated using population-level allelic frequencies were consistently higher than those calculated using population-level supertype frequencies. Similar observations in guppies have been attributed to the forces of balancing selection acting on supertypes in such a way that homogenizes supertype frequencies across locations that share similar selection patterns (Lighten et al. 2017). Limited sampling of live adult gray seals from Sable Island, Canada, and Massachusetts, United States has revealed similar levels of IAV seroprevalence (Puryear et al. 2016); however, a comprehensive regional survey of pinniped pathogen presence, the putative selective force of interest, could shed additional light on factors that are affecting selection on the MHC. Hammond et al. (2012) reported preliminary evidence of differences in *MHC-I* diversity between sampling sites that differed in mortality rates during a past European PDV epidemic. In the Northwest Atlantic, reported pinniped disease outbreaks have generally been localized to the Northeast United States, but the presence of a coordinated marine mammal stranding network in the United States, and the lack thereof in Canada, may lead to geographic biases in disease reporting.

As initially called for by Hammond et al. (2012), our population scale analysis of *MHC-I* diversity in gray and harbor seals has advanced our understanding of pinniped immunogenetics in ways that are complementary to the deep dive provided previously by full gene sequencing. The 2 studies represent clear trade-offs between sample size and sequencing depth, which provide complementary insights into the forces of natural selection and resulting gene evolution. With continued advances in long-read sequencing technology (Viļuma et al. 2017; Cheng et al. 2022; He et al. 2021), such trade-offs may eventually become obsolete, making it possible for future studies to explore full-sequence *MHC-I* allele diversity at a population scale. Additionally, future studies that are able to implement a case:control study design, for example by comparing *MHC-I* diversity between survivors and nonsurvivors of a disease outbreak, may offer additional insights into the role of the observed genetic diversity in disease resistance. In other marine mammal systems, such approaches have shown promise for investigating the role of both the MHC and other genes in toxin and disease resistance (Cammen et al. 2015; Batley et al. 2021). Despite all of the evidence gathered here and previously of the role of the MHC in disease resistance, we must acknowledge that it is probably 1 piece of a complex puzzle (Acevedo-Whitehouse

and Cunningham 2006). Future studies that leverage recent advances in whole-genome sequencing approaches for nonmodel species will illuminate further insights into the genetic underpinnings of pinniped disease resistance.

Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

Funding

This work was supported by the University of Maine Office of the Vice President for Research. Funding for initial sample collection was provided by the Department of Fisheries and Oceans Canada; Natural Sciences and Engineering Research Council of Canada; NOAA Northeast Fisheries Science Center; Northeast Fisheries Observer Program; and the National Institutes of Health Centers of Excellence for Influenza Research and Surveillance (grant number HHSN272201400008C).

Acknowledgments

The collection of samples in US waters and import of seal samples from Canada were conducted under NOAA Permit No. 17670-03 issued to the NOAA Northeast Fisheries Science Center. We thank Kelley Thomas and Jeffrey Hall for consulting on the next-generation sequencing approach used in this study, and Pauline Kamath, Nishad Jayasundara, and 2 anonymous reviewers for feedback on a prior version of this manuscript.

Data Availability

All DNA sequences involved in this study are available at Genbank accessions OM648960–OM649179 and NCBI BioProject PRJNA805543.

References

Acevedo-Whitehouse K, Cunningham AA. 2006. Is MHC enough for understanding wildlife immunogenetics? *Trends Ecol Evol*. 21:433–438.

Aldridge BM, Bowen L, Smith BR, Antonelis GA, Gulland F, Stott JL. 2006. Paucity of class I MHC gene heterogeneity between individuals in the endangered Hawaiian monk seal population. *Immunogenetics*. 58:203–215.

Anthony SJ, St Leger JA, Pugliares K, Ip HS, Chan JM, Carpenter ZW, Navarrete-Macias I, Sanchez-Leon M, Saliki JT, Pedersen J, et al. 2012. Emergence of fatal avian influenza in New England harbor seals. *mBio*. 3:e00166-12.

Arden B, Klein J. 1982. Biochemical comparison of major histocompatibility complex molecules from different subspecies of *Mus musculus*: evidence for trans-specific evolution of alleles. *Proc Natl Acad Sci USA*. 79:2342–2346.

Ballingall KT, Miltiadou D, Chai ZW, McLean K, Rocchi M, Yaga R, McKeever DJ. 2008. Genetic and proteomic analysis of the MHC class I repertoire from four ovine haplotypes. *Immunogenetics*. 60:177–184.

Batley KC, Sandoval-Castillo J, Kemper CM, Zanardo N, Tomo I, Beheregaray LB, Möller LM. 2021. Whole genomes reveal multiple candidate genes and pathways involved in the immune response of dolphins to a highly infectious virus. *Mol Ecol*. 30:6434–6448.

Bodewes R, Bestebroer TM, van der Vries E, Verhagen JH, Herfst S, Koopmans MP, Fouchier RA, Pfankuche VM, Wohlsein P, Siebert U, et al. 2015. Avian Influenza A (H10N7) virus-associated mass deaths among harbor seals. *Emerg Infect Dis*. 21:720–722.

Boyson JE, Shufflebotham C, Cadavid LF, Urvater JA, Knapp LA, Hughes AL, Watkins DI. 1996. The MHC class I genes of the rhesus monkey. Different evolutionary histories of MHC class I and II genes in primates. *J Immunol*. 156:4656–4665.

Cammen K, Hoffman JI, Knapp LA, Harwood J, Amos W. 2011. Geographic variation of the major histocompatibility complex in Eastern Atlantic grey seals (*Halichoerus grypus*). *Mol Ecol*. 20:740–752.

Cammen KM, Rasher DB, Steneck RS. 2019. Predator recovery, shifting baselines, and the adaptive management challenges they create. *Ecosphere*. 10:e02579.

Cammen KM, Schultz TF, Don Bowen W, Hammill MO, Puryear WB, Runstadler J, Wenzel FW, Wood SA, Kinnison M. 2018. Genomic signatures of population bottleneck and recovery in Northwest Atlantic pinnipeds. *Ecol Evol*. 8:6599–6614.

Cammen KM, Wilcox LA, Rosel PE, Wells RS, Read AJ. 2015. From genome-wide to candidate gene: an investigation of variation at the major histocompatibility complex in common bottlenose dolphins exposed to harmful algal blooms. *Immunogenetics*. 67:125–133.

Carroll EL, Hall A, Olsen MT, Onoufriou AB, Gaggiotti OE, Russell DJ. 2020. Perturbation drives changing metapopulation dynamics in a top marine predator. *Proc Biol Sci*. 287:20200318.

Cheng Y, Grueber C, Hogg CJ, Belov K. 2022. Improved high-throughput MHC typing for non-model species using long-read sequencing. *Mol Ecol Resour*. 22:862–876.

den Heyer CE, Bowen WD, Dale J, Gosselin J-F, Hammill MO, Johnston DW, Lang SLC, Murray KT, Stenson GB, Wood SA. 2020. Contrasting trends in gray seal (*Halichoerus grypus*) pup production throughout the increasing northwest Atlantic metapopulation. *Mar Mamm Sci*. 37:611–630.

Doherty PC, Zinkernagel RM. 1975. Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*. 256:50–52.

Duignan PJ, Sadove S, Saliki JT, Geraci JR. 1993. Phocine distemper in harbor seals (*Phoca vitulina*) from Long Island, New York. *J Wildl Dis*. 29:465–469.

Earle JAP, Melia MM, Doherty NV, Nielsen O, Cosby SL. 2011. Phocine distemper virus in seals, east coast, United States, 2006. *Emerg Infect Dis*. 17:215–220.

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 26:2460–2461.

Gulland FMD, Hall AJ. 2007. Is marine mammal health deteriorating? Trends in the global reporting of marine mammal disease. *EcoHealth*. 4:135–150.

Haldane JBS. 1949. Disease and evolution. *Suppl Ricerca Sci*. 19:68–76.

Hall AJ, Jepson PD, Goodman SJ, Härkönen T. 2006. Phocine distemper virus in the North and European Seas – Data and models, nature and nurture. *Biol Conserv.* 131:221–229.

Hammond JA, Guethlein LA, Norman PJ, Parham P. 2012. Natural selection on marine carnivores elaborated a diverse family of classical MHC class I genes exhibiting haplotypic gene content variation and allelic polymorphism. *Immunogenetics*. 64:915–933.

Harding KC, Härkönen T, Caswell H. 2002. The 2002 European seal plague: epidemiology and population consequences. *Ecol Lett*. 5:727–732.

Härkönen T, Dietz R, Reijnders P, Teilmann J, Harding K, Hall A, Brasseur S, Siebert U, Goodman SJ, Jepson PD, et al. 2006. The 1988 and 2002 phocine distemper virus epidemics in European harbour seals. *Dis Aquat Organ*. 68:115–130.

He K, Minias P, Dunn PO. 2021. Long-read genome assemblies reveal extraordinary variation in the number and structure of MHC loci in birds. *Genome Biol Evol*. 13:evaa270.

Heide-Jørgensen MP, Härkönen T. 1992. Epizootiology of the seal disease in the eastern North Sea. *J Appl Ecol*. 29:99–107.

Hoelzel AR, Stephens JC, O'Brien SJ. 1999. Molecular genetic diversity and evolution at the MHC DQB locus in four species of pinnipeds. *Mol Biol Evol*. 16:611–618.

Hughes AL, Nei M. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci USA*. 86:958–962.

Jo WK, Osterhaus AD, Ludlow M. 2018. Transmission of morbilliviruses within and among marine mammal species. *Curr Opin Virol*. 28:133–141.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24:1403–1405.

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*. 8:275–282.

Jost L. 2008. G(ST) and its relatives do not measure differentiation. *Mol Ecol*. 17:4015–4026.

Katona SK, Rough V, Richardson DT. 1993. *A field guide to whales, dolphins, and seals from Cape Cod to Newfoundland*. Washington (DC): Smithsonian Institution Press.

Kim K, Dobson AP, Gulland FMD, Harvell CD. 2005 Diseases and the conservation of marine biodiversity. In: Norse EA, Crowder LB, editors, *Marine conservation biology: the science of maintaining the sea's biodiversity*. Washington (DC): Island Press. p. 149–166.

Klein J, Sato A, Nikolaidis N. 2007. MHC, TSP, and the origin of species: from immunogenetics to evolutionary genetics. *Annu Rev Genet*. 41:281–304.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 35:1547–1549.

Lau Q, Chow N, Gray R, Gongora J, Higgins DP. 2015. Diversity of MHC DQB and DRB genes in the endangered Australian sea lion (*Neophoca cinerea*). *J Hered*. 106:395–402.

Lighten J, Papadopulos AS, Mohammed RS, Ward BJ, Paterson IG, Baillie L, Bradbury IR, Hendry AP, Bentzen P, Van Oosterhout C. 2017. Evolutionary genetics of immunological supertypes reveals two faces of the Red Queen. *Nat Commun*. 8:1–11.

Liu X, Rønhøj Schjøtt S, Granquist SM, Rosing-Asvid A, Dietz R, Teilmann J, Galatius A, Cammen K, O'Corry-Crowe G, Harding K, et al. 2022. Origin and expansion of the world's most widespread pinniped: range-wide population genomics of the harbour seal (*Phoca vitulina*). *Mol Ecol*. 31:1682–1699.

Lively CM, Apanius V. 1995. Genetic diversity in host-parasite interactions. In: Grenfell BT, Dobson AP, editors, *Ecology of Infectious Diseases in Natural Populations*. Cambridge (MA): Cambridge University Press. p. 421–449.

Markussen NH, Have P. 1992. Phocine distemper virus infection in harp seals *(Phoca groenlandica)*. *Mar Mamm Sci*. 8:19–26.

Marsh SG, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, Fernández-Viña M, Geraghty DE, Holdsworth R, Hurley CK, et al. 2010. Nomenclature for factors of the HLA system, 2010. *Tissue Antigens*. 75:291–455.

McCarthy AJ, Shaw MA, Jepson PD, Brasseur SM, Reijnders PJ, Goodman SJ. 2011. Variation in European harbour seal immune response genes and susceptibility to phocine distemper virus (PDV). *Infect Genet Evol*. 11:1616–1623.

Neefjes J, Jongsma ML, Paul P, Bakke O. 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol*. 11:823–836.

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol*. 3:418–426.

NOAA Fisheries. 2021. 2018–2020 Pinniped Unusual Mortality Event Along the Northeast Coast [Internet]; [cited January 5 2022]. Available from: https://www.fisheries.noaa.gov/new-england-mid-atlantic/marine-life-distress/2018-2020-pinniped-unusual-mortality-event-along.

Osborne AJ, Pearson J, Negro SS, Chilvers BL, Kennedy MA, Gemmell NJ. 2015. Heterozygote advantage at MHC DRB may influence response to infectious disease epizootics. *Mol Ecol*. 24:1419–1432.

Papkou A, Guzella T, Yang W, Koepper S, Pees B, Schalkowski R, Barg MC, Rosenstiel PC, Teotónio H, Schulenburg H. 2019. The genomic basis of Red Queen dynamics during rapid reciprocal host-pathogen coevolution. *Proc Natl Acad Sci USA*. 116:923–928.

Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR, Quail M, Smith F, Walker D, Libberton B, et al. 2010. Antagonistic coevolution accelerates molecular evolution. *Nature*. 464:275–278.

Piertney SB, Oliver MK. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity (Edinb)*. 96:7–21.

Pomeroy PP, Hammond JA, Hall AJ, Lonergan M, Duck CD, Smith VJ, Thompson H. 2005. Morbillivirus neutralising antibodies in Scottish grey seals *Halichoerus grypus*: assessing the effects of the 1988 and 2002 PDV epizootics. *Mar Ecol Prog Ser*. 287:241–250.

Puryear WB, Keogh M, Hill N, Moxley J, Josephson E, Davis KR, Bandoro C, Lidgard D, Bogomolni A, Levin M, et al. 2016. Prevalence of influenza A virus in live-captured North Atlantic gray seals: a possible wild reservoir. *Emerg Microbes Infect*. 5:e81.

Puryear W, Sawatzki K, Bogomolni A, Hill N, Foss A, Stokholm I, Olsen MT, Nielsen O, Waltzek T, Goldstein T, et al. 2021. Longitudinal analysis of pinnipeds in the northwest Atlantic provides insights on endemic circulation of phocine distemper virus. *Proc Biol Sci*. 288:20211841.

Rada C, Lorenzi R, Powis SJ, van den Bogaerde J, Parham P, Howard JC. 1990. Concerted evolution of class I genes in the major histocompatibility complex of murine rodents. *Proc Natl Acad Sci USA*. 87:2167–2171.

Sandberg M, Eriksson L, Jonsson J, Sjöström M, Wold S. 1998. New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. *J Med Chem*. 41:2481–2491.

Sanderson CE, Alexander KA. 2020. Unchartered waters: climate change likely to intensify infectious disease outbreaks causing mass mortality events in marine mammals. *Glob Chang Biol*. 26:4284–4301.

Sanderson CE, Jobbins SE, Alexander KA. 2013. With Allee effects, life for the social carnivore is complicated. *Popul Ecol*. 56:417–425.

Saper MA, Bjorkman PJ, Wiley DC. 1991. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. *J Mol Biol*. 219:277–319.

Siembieda JL, Hall AJ, Gulland FMD, Rowles T, Garron M, Matassa K, Rotstein DS, Gonzalez S, Northeast Regional Marine Mammal Stranding Network, Johnson CK. 2017. Epidemiology of a phocine distemper virus outbreak along the North Atlantic coast of the United States. *Aquat Mamm*. 43:254–263.

Sigourney DB, Murray KT, Gilbert JR, Ver Hoef JM, Josephson E, DiGiovanni RA Jr. 2021. Application of a Bayesian hierarchical model to estimate trends in Atlantic harbor seal *(Phoca vitulina vitulina)* abundance in Maine, U.S.A., 1993–2018. *Mar Mamm Sci*. 38:500–516.

Simeone CA, Gulland FM, Norris T, Rowles TK. 2015. A systematic review of changes in marine mammal health in North America, 1972-2012: the need for a novel integrated approach. *PLoS One*. 10:e0142105.

Slade RW. 1992. Limited MHC polymorphism in the southern elephant seal: implications for MHC evolution and marine mammal population biology. *Proc Biol Sci*. 249:163–171.

Sommer S, Courtiol A, Mazzoni CJ. 2013. MHC genotyping of non-model organisms using next-generation sequencing: a new methodology to deal with artefacts and allelic dropout. *BMC Genomics*. 14:542.

Spurgin LG, van Oosterhout C, Illera JC, Bridgett S, Gharbi K, Emerson BC, Richardson DS. 2011. Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations. *Mol Ecol*. 20:5213–5225.

Tallmadge RL, Campbell JA, Miller DC, Antczak DF. 2010. Analysis of MHC class I genes across horse MHC haplotypes. *Immunogenetics*. 62:159–172.

Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol*. 10:512–526.

Tanaka-Matsuda M, Ando A, Rogel-Gaillard C, Chardon P, Uenishi H. 2009. Difference in number of loci of swine leukocyte antigen classical class I genes among haplotypes. *Genomics*. 93:261–273.

Těšický M, Vinkler M. 2015. Trans-species polymorphism in immune genes: general pattern or MHC-restricted phenomenon? *J Immunol Res*. 2015:838035.

Thompson JN. 1998. Rapid evolution as an ecological process. *Trends Ecol Evol*. 13:329–332.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 40:e115.

van Valen L. 1973. A new evolutionary law. *Evol Theory*. 1:1–30.

Viļuma A, Mikko S, Hahn D, Skow L, Andersson G, Bergström TF. 2017. Genomic structure of the horse major histocompatibility complex class II region resolved using PacBio long-read sequencing technology. *Sci Rep*. 7:45518.

Weber DS, Stewart BS, Schienman J, Lehman N. 2004. Major histocompatibility complex variation at three class II loci in the northern elephant seal. *Mol Ecol*. 13:711–718.

Wen C, Wu L, Qin Y, Van Nostrand JD, Ning D, Sun B, Xue K, Liu F, Deng Y, Liang Y, et al. 2017. Evaluation of the reproducibility of amplicon sequencing with Illumina MiSeq platform. *PLoS One*. 12:e0176716.

Wood SA, Murray KT, Josephson E, Gilbert J. 2020. Rates of increase in gray seal (*Halichoerus grypus atlantica*) pupping at recolonized sites in the United States, 1988–2019. *J Mamm*. 101:121–128.

Woolhouse ME, Webster JP, Domingo E, Charlesworth B, Levin BR. 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat Genet*. 32:569–577.

Yu G. 2020. Using ggtree to visualize data on tree-like structures. *Curr Protoc Bioinformatics*. 69:e96.

Zohari S, Neimanis A, Härkönen T, Moraeus C, Valarcher JF. 2014. Avian influenza A (H10N7) virus involvement in mass mortality of harbour seals (*Phoca vitulina*) in Sweden, March through October 2014. *Euro Surveill*. 19:20967.