Check for updates

Functional immune diversity in reindeer reveals a high Arctic population at risk

Morten Lukacs[1], Ingebjørg H. Nymo[1], Knut Madslien[1], Jørn Våge[1], Vebjørn Veiberg[2], Christer Moe Rolandsen[2], Cathrine Arnason Bøe[1], Arvind Y. M. Sundaram[3] and Unni Grimholt[1]*

[1]Norwegian Veterinary Institute, Aas, Norway, [2]Department of Terrestrial Ecology, Norwegian Institute for Nature Research, Trondheim, Norway, [3]Department of Medical Genetics, Oslo University Hospital, Oslo, Norway

Climate changes the geographic range of both species as well as pathogens, causing a potential increase in the vulnerability of populations or species with limited genetic diversity. With advances in high throughput sequencing (HTS) technologies, we can now define functional expressed genetic diversity of wild species at a larger scale and identify populations at risk. Previous studies have used genomic DNA to define major histocompatibility complex (MHC) class II diversity in reindeer. Varying numbers of expressed genes found in many ungulates strongly argues for using cDNA in MHC typing strategies to ensure that diversity estimates relate to functional genes. We have used available reindeer genomes to identify candidate genes and established an HTS approach to define expressed MHC class I and class II diversity. To capture a broad diversity we included samples from wild reindeer from Southern Norway, semi-domesticated reindeer from Northern Norway and reindeer from the high Artic archipelago Svalbard. Our data show a medium MHC diversity in semi-domesticated and wild Norwegian mainland reindeer, and low MHC diversity reindeer in Svalbard reindeer. The low immune diversity in Svalbard reindeer provides a potential risk if the pathogenic pressure changes in response to altered environmental conditions due to climate change, or increased human-related activity.

KEYWORDS

climate change, MHC, reindeer, immune diversity, Svalbard, Norway

Introduction

Immune diversity affects how species handle new pathogens and major histocompatibility complex (MHC) diversity is a major factor in this area. The link between MHC and resistance towards pathogens has been documented in diverse species such as chickens, humans and Atlantic salmon (Plachy et al., 1992; Vallejo et al., 1998; Langefors et al., 2001; Grimholt et al., 2003; Kjøglum et al., 2008; Goulder and Walker, 2012; Chappell

et al., 2015). MHC genes are key elements of the adaptive immune system and affect how individuals handle infections (Paterson et al., 1998). MHC genes are amongst the most polymorphic genes known to date, and as such provide important information useful for conservation of endangered species. MHC molecules are divided into two classes, i.e., MHC class I (MHCI) and class II (MHCII) where classical genes are highly polymorphic, with tissue-specific expression patterns and bind peptides. MHCI molecules consist of an alpha chain non-covalently linked to beta2-microglobulin. They handle presentation of endogenously derived peptides, for instance from viruses to CD8 positive alpha beta T-cells and initiate an immune response against the pathogen. The MHCII molecule, composed of an alpha (MHCIIA) and a beta (MHCIIB) chain, bind and present exogenously derived peptides to CD4 positive alpha beta T-cells and initiate an immune response towards for instance bacteria. MHCI molecules are expressed on most cells while MHCII molecules are restricted to antigen presenting cells. Non-classical MHC genes are less polymorphic, bind various ligands and do not interact with alpha beta T-cells.

Cattle (*Bos Taurus*), addax (*Addax nasomaculatus*), as well as other mammals, all have their MHCI and MHCII genes encoded within one chromosomal region (Kulski et al., 2002; Traherne, 2008; Shiina et al., 2009; Behl et al., 2012; Li et al., 2020). Most species have unique numbers of MHCI and MHCII genes as exemplified by cattle with three DRB genes, two DQB genes and haplotypes ranging from 3 to 4 polymorphic MHCI genes (Behl et al., 2012; Schwartz et al., 2022). Comparatively, the addax has seven MHCI genes with unknown polymorphism and expression patterns, three expressed DRB genes and one expressed DQB gene (Li et al., 2020). Humans on the other hand, have nine DRB genes present in various combinations in different haplotypes (Andersson et al., 1987; Doxiadis et al., 2012). Of these, four are significantly expressed, whereof one is the major locus.[1] Assessing MHC diversity in less studied species using single exon domains and genomic DNA could amplify fragments from paralogue classical, non-classical and pseudogenes, where diversity estimates would not relate to a single polymorphic gene.

Advances in sequencing technology and availability of sequenced genomes have opened for assessing MHC diversity using genomes as a foundation for studying expressed sequences from single genes. This can improve the precision in our understanding of how genetic variability may be linked to fitness and long-term survival of a species, or how this relates to climate change or new pathogens (Teixeira and Huber, 2021).

To investigate this approach using genomes as a foundation and amplification of functional expressed sequences, we use reindeer as a model species. Reindeer is one of the species where climate has had an impact on the historical European population diversity and habitat range (Sommer et al., 2014). There are currently 12 extant subspecies of reindeer broadly divided into

---

1 https://www.ebi.ac.uk/ipd/imgt/hla/about/statistics/

forest, tundra, and Arctic subspecies (Wilson and Reeder, 2005). At a global level, the abundance of reindeer subspecies found throughout the Arctic today have declined, and reindeer has been listed as vulnerable in the IUCN Red List of Threatened Species (UCN, 2021). In Canada, the Peary caribou (*Rangifer tarandus pearyi*) has suffered drastic declines caused by severe weather, hunting and predation and is listed as endangered (Festa-Bianchet et al., 2011).

In Norway, wild reindeer are located both on the mainland as well as on the archipelago of Svalbard. Svalbard reindeer (*Rangifer tarandus platyrhynchus*) is a separate subspecies. It constitutes the most isolated of all reindeer populations (UCN, 2021) and their earliest known presence on Svalbard dates back more than 5,000 years ago (van der Knaap, 1989). Svalbard reindeer have been hunted since the 17th century, but harvest increased radically in the late 19th century. At the time of protection in 1925, only four meta-populations remained (Lønø, 1959) and the whole population was possibly as low as 1,000 animals. Since then, the population has increased with a population estimate of 22,435 animals in the period 2013–2016 (Le Moullec et al., 2019). Most likely due to the severe bottleneck, the Svalbard reindeer has low genetic diversity compared to other populations and subspecies of reindeer (Flagstad and Røed, 2003; Yannic et al., 2014).

On mainland Norway, there are today approximately 20,000–25,000 wild Eurasian tundra reindeer (*Rangifer tarandus tarandus*). The meta-population is divided into 24 management areas in mid- and southern parts of Norway. This division is partly a result of human infrastructure and disturbance (Gundersen et al., 2022). Population size and composition are managed by hunting and is organized in close cooperation between landowner organizations and state agencies (Strand et al., 2012; Mysterud et al., 2021).

In Norway, pastoralism of reindeer began by the Sámi people as early as the 16th century. Reindeer herding is based on natural pastures and moving animals between seasonal pastures (Tryland et al., 2016; Røed et al., 2018). About 75% of the approximately 220,000 semi-domesticated Eurasian tundra reindeer (*Rangifer tarandus tarandus*) on mainland Norway reside in the northernmost county Troms and Finnmark (Landbruksdirektoratet, 2021).

Climate change may have a direct impact on the health and diseases of wild animals and some infectious diseases in reindeer are classified as climate sensitive (Omazic et al., 2019). On the Norwegian mainland, the tick *Ixodes ricinus* is entering areas more used by semi-domesticated reindeer. The tick is the intermediate host for several vector-borne diseases such as anaplasmosis or babesiosis (Jore et al., 2014; Mysterud et al., 2017; Tryland et al., 2019). The distribution of mosquitoes and biting midges has also been found to correlate with changes in climatic conditions (Elbers et al., 2015), which may carry pathogens which reindeer are susceptible to such as West Nile Virus, Schmallenberg virus and bluetongue virus (Palmer et al., 2004; Laaksonen et al., 2009; Sánchez Romano et al., 2019). Such geographical changes in pathogen distribution could pose a future challenge to Norwegian reindeer.

Previous studies on MHC diversity in reindeer investigated MHCII diversity using genomic DNA (Mikko et al., 1999; Kennedy et al., 2011; Taylor et al., 2012; Gagnon et al., 2020) and there have been no investigations of expressed MHCI or MHCII diversity. With the availability of reindeer genome sequences, we aim to define classical versus non-classical MHC genes and use these data to assess expressed MHC diversity in samples from semi-domesticated reindeer from Northern Norway, wild reindeer from Southern Norway and reindeer from the high Arctic Svalbard archipelago.

Materials and methods

Materials

Mesenteric lymph nodes situated close to the ileocecal junction were collected in October 2021 from two different semi-domesticated reindeer herds in Northern Norway (Supplementary data 1) during slaughter in a commercial slaughterhouse totaling samples from 62 animals. The lymph nodes and kidney samples were stored in RNAlater (Merck, Darmstadt, Germany) according to manufacturer's protocol. These animals are hereafter defined as population number 1.

During March 2021, we collected blood from 18 wild free-ranging reindeer from two sub-populations in the Langfjella mountain range in Southern Norway. Of these, 12 were from Hardangervidda and six from Setesdal Ryfylke area (Supplementary data 1). Blood was collected from the jugular vein into EDTA-tubes during capture in connection with GPS collaring. The animal handling protocol was approved by the Norwegian Food Safety Authority (permit no. 20/227896). These animals are hereafter defined as population number 2.

During October 2021, kidney samples from nine Svalbard reindeer were collected from harvested animals originating from two neighboring valleys (Semmeldalen and Colesdalen) on Nordenskiöld Land, Svalbard (Supplementary data 1). The harvest of animals was approved by the Governor of Svalbard (permit no. 16/01632–40). These animals constitute population number 3.

Genome mining

To identify amplicons and design primers we performed tBlastN search of available reindeer genomes, i.e., GCA_019903745.1 (*Rangifer tarandus caribou*, denoted caribou), GCA_014898785.1 (*Rangifer tarandus granti*, denoted granti) and GCA_902712895.1 (*Rangifer tarandus tarandus*, denoted tarandus) using MHCI and MHCII gene sequences from human and cow (Supplementary data 2). None of the reindeer genomes are annotated so genes were predicted using FGenesh (Solovyev et al., 2006) in addition to tblastN with bovine and later reindeer genes were used to manually correct gene predictions. Geographical origin of the caribou and granti genome

animals is uncertain, but the tarandus genome originates from a Norwegian Hardangervidda animal (pers. comm. Montserrat Torres Oliva, Institute of Clinical Molecular Biology, Kiel University).

Preparing the sequencing library

RNA was isolated from blood using Tempus™ Spin RNA Isolation Kit (Thermo Fisher Scientific, Waltham, United States) according to manufacturer's protocol and the RNA was eluted in 90 μl RNase-free water. RNA concentrations were measured on TapeStation capillary electrophoresis (Agilent, Santa Clara, CA, United States) and Qubit™ fluorometer (Invitrogen, Carlsbad, CA, United States).

Kidney and lymph node samples preserved in RNAlater were dissolved using Tissue lyzer (Qiagen, Hilden, Germany) with MagNA Pure LC RNA Isolation Tissue Lysis Buffer (Roche Molecular Systems, Mannheim, Germany). RNA was isolated using MagNA Pure 96 Cellular RNA Large volume kit (Roche Molecular Systems) according to the manufacturer's recommendation. cDNA was synthesized using up to 1,000 ng total RNA and the QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's recommendation resulting in 20 μl cDNA stored at −20°C for later use.

We used 10 ng of cDNA from two animals (V0300 and V0301) in 20 μl PCR reactions for the specific amplicons A1, A5, DQB and DRB genes with KAPA HiFi HotStart ReadyMix (Roche Molecular Systems), with the gene specific primers shown in Supplementary data 3. Products were verified on a 1% agarose gel and sequenced using the PCR primers with standard Sanger sequencing, to verify that primers amplified correct targets (data not shown). The resultant PCR products were further cleaned using 1.5× PCR volume of Agencourt AMPure XP PCR purification kit (Beckman Coulter, Brea, CA, United States) according to manufacturer's recommendation and dissolved in 30 μl TE buffer.

Libraries containing each of the four PCR products were pooled in proportions to ensure similar coverage and subjected to 10 cycles of PCR using the second set of primers (Supplementary data 3), thus adding one unique Illumina index pair for each animal. One μl forward and reverse index primer were used for subsequent PCR reaction with KAPA HiFi HotStart ReadyMix. The following program was used for amplification: 94°C for 2 min; 10× (94°C for 30 s, 58°C for 30 s, 72°C for 60 s); 72°C for 10 min. Amplicons were purified using 1.0× PCR volume of Agencourt AMPure XP PCR purification kit according to manufacturer's recommendation and dissolved in 30 μl TE. Samples were finally pooled together and concentrated with 1.0 x PCR volume of AMPure XP beads.

Based on data from Tapestation and Qubit, the PCR pools were mixed totaling 2 μg DNA and sequencing was performed on Illumina MiSeq platform (Illumina, San Diego, CA, United States) using the v3 chemistry to achieve 300 base pairs paired end reads.

Bioinformatics filtering and population genetics analyses

Sequenced data was analyzed as previously described (Sundaram et al., 2020). Briefly, raw reads for each animal were obtained based on the Illumina index introduced during the second PCR reaction. BBDuk v34.56 (part of BBTools) was used to remove/trim bad quality reads and sequencing adapter sequences and the cleaned reads were further demultiplexed based on the primers used during the first PCR reaction using multiplexer v1.9 (de Muinck et al., 2017).

Paired reads for each MHC subgroup were combined using FLASH v1.2.11 (Magoč and Salzberg, 2011) with default settings and the resulting full-length amplified reads were collapsed to identify all the unique reads and sorted based on the number of times it was present in the data using fastx_collapser [part of FASTX Toolkit v0.0.13 (FASTX-Toolkit, 2010)].

Standard indices of genetic diversity, including number of alleles, observed and expected heterozygosity, allelic richness and deviation from Hardy Weinberg equilibrium, were computed using RStudio (RStudioTeam, 2022) and the package PopGenReport version 3.0.7 (Adamack and Gruber, 2014). Allelic richness (El Mousadik and Petit, 1996) for each combination of population and locus is calculated. To account for differences in sample sizes and genotyping success, rarefaction is used where the sample size for each combination of population and locus is set equal to the smallest number of alleles seen in a sample across all combinations of population and locus.

MHC amplicon sequencing

Performing successful amplicon sequencing relies on decent amplification of the genes in question. In our case, the final sequenced pool is a mixture of 89 animals each containing four different cDNA sequences amplified with dual index primers to allow us to sort sequences into animals based on index and then into each individual gene based on primer sequence. All sequences per primer pair for each animal were collapsed with a requirement of 100% sequence identity providing a number of consensus sequences listed with number of matching reads. Sequence data have been submitted to NCBI SRA under the BioProject accession number PRJNA832088.

MHC gene sequences identified in this study have been submitted to GenBank with accession numbers ON411528-ON411580. Expressed MHC sequences and allelic variants for each animal are shown in Supplementary data 5.

Phylogenetic analysis

Amino acid sequence alignments were performed in ClustalX (Larkin et al., 2007) and polished using Jalview 2 (Waterhouse et al., 2009). The evolutionary histories were inferred using the Maximum Likelihood method and JTT matrix-based model (Figure 1; Jones et al., 1992) or Tamura 3-parameter model (Figures 2, 3; Tamura, 1992). The trees with the highest log likelihood are shown. Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using individual models, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites. The rate variation model allowed for some sites to be evolutionarily invariable. Bootstrap values in percentage from 1,000 trials are shown next to the branches. The trees are drawn to scale, with branch lengths measured in the number of substitutions per site. All positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option). Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018).
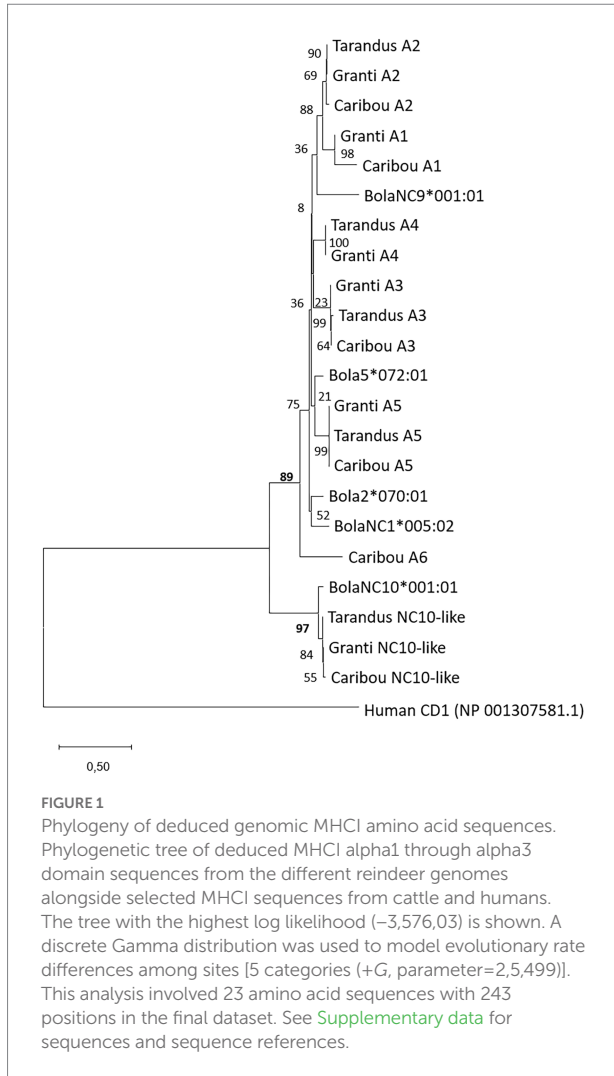
MHC nomenclature

To accommodate naming MHC genes in numerous species the MHC nomenclature committee has decided on naming MHC genes with the first two letters of the Latin name of the genus as well as species (Maccari et al., 2018; Ballingall et al., 2018a). Curated MHC sequences with species-specific nomenclature for various non-human species are gathered in the Immuno Polymorphism Database (IPD) MHC (IPD-MHC, 2021). Our chosen nomenclature is in accordance with these rules. For example, the sequences *Rata-DRB1*01:01*, *Rata-DRB1*01:02* and *Rata-DRB1*02:01* describe sequences from *Rangifer tarandus* locus DRB1. The sequence Rata-DRB1*01 differs from *Rata-DRB1*02* in more than three amino acids while *Rata-DRB1*01:01* and *Rata-DRB1*01:02* differ in three or less amino acids.

For bovines, IPD-MHC MHCI nomenclature is divided into numbered loci for genes assumed to be classical, and a NC-heading for genes assumed to be non-classical, i.e., non-polymorphic. We follow this nomenclature, but chose to include a single letter A in addition to a number based on location on the chromosome for our reindeer MHCI genes. Adding another two digits as for instance in *Rata-A1*14:01:01* versus *Rata-A1*14:01:02* implies changes in the nucleotide sequence that do not provide any changes in the translated sequence.

Results

Genome mining and choice of amplicons

Based on data from other mammalian species, the MHCIIA genes are not considered highly polymorphic and were thus not included in our study. *Rangifer tarandus tarandus*, *Rangifer tarandus granti* and *Rangifer tarandus caribou* genomes were all searched for MHCI and MHCII DRB and DQB genes to choose
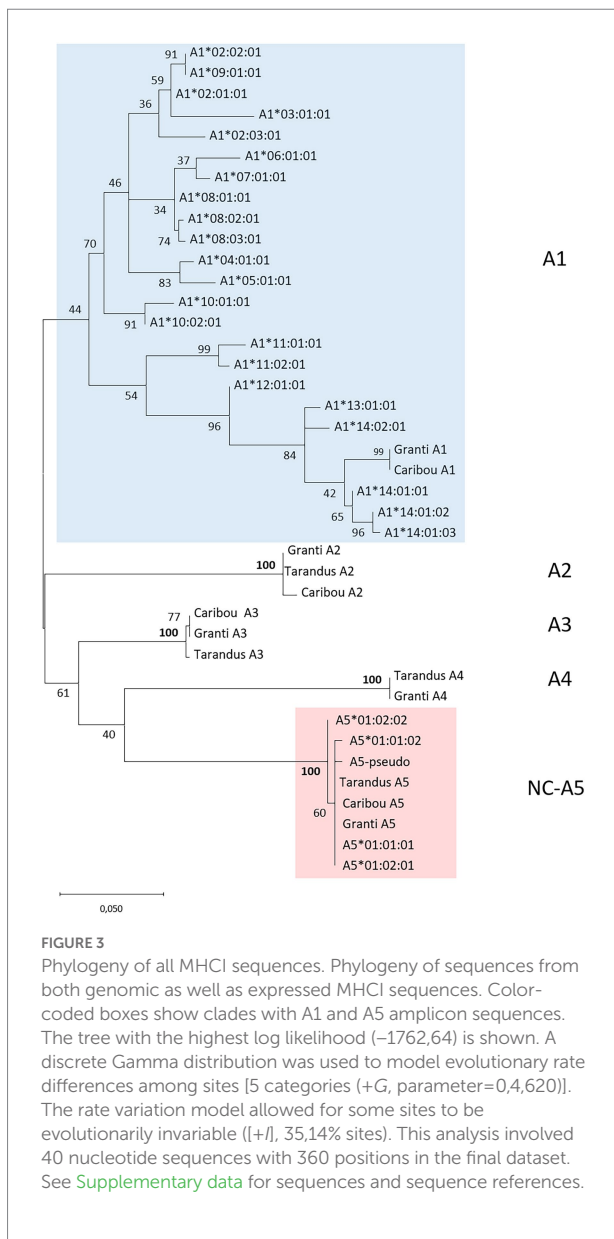
FIGURE 1
Phylogeny of deduced genomic MHCI amino acid sequences. Phylogenetic tree of deduced MHCI alpha1 through alpha3 domain sequences from the different reindeer genomes alongside selected MHCI sequences from cattle and humans. The tree with the highest log likelihood (−3,576,03) is shown. A discrete Gamma distribution was used to model evolutionary rate differences among sites [5 categories (+G, parameter=2,5,499)]. This analysis involved 23 amino acid sequences with 243 positions in the final dataset. See Supplementary data for sequences and sequence references.

FIGURE 2
Phylogeny of DRB1 and DQB sequences. Phylogenetic tree of nucleotide sequences of amplicon region from the different reindeer samples alongside selected MHCII sequences from the three reindeer genomes. Red, blue, orange and green colored boxes show amplicon clades. The tree with the highest log likelihood (−3,061,43) is shown. A discrete Gamma distribution was used to model evolutionary rate differences among sites [5 categories (+G, parameter=0,8,885)]. This analysis involved 41 nucleotide sequences with 433 positions in the final dataset. See Supplementary data for sequences and sequence references.

as amplicons. Hereafter we refer to data from these three genomes as "tarandus," "granti" and "caribou." Deduced genomic amino acid sequences are gathered in Supplementary data 4. We used genes and MHC organization of cattle as a reference as they represent an ungulate with a well-assembled genome and well-defined MHC genes (Vincent et al., 1996; Yamamoto et al., 2020; Schwartz et al., 2022).

Assembling genomic MHC regions can be difficult, in particular when the sequence identity between loci is high or when the regions contain multiple repeated sequences. Since all studied mammals so far has been found to contain one MHC region (Kulski et al., 2002), this was also expected for reindeer. However, none of the reindeer genomes had a single MHC region similar to the 20 mega bases (Mb) region found in cattle (Takeshima and Aida, 2006). All genomes displayed three MHCII DRB loci, one to three MHCII DQB loci and five to seven MHCI loci (Figure 4; Supplementary data 4.1). Granti MHC genes resided on six scaffolds, tarandus MHC genes resided on four scaffolds and caribou MHC genes resided on two scaffolds. Only
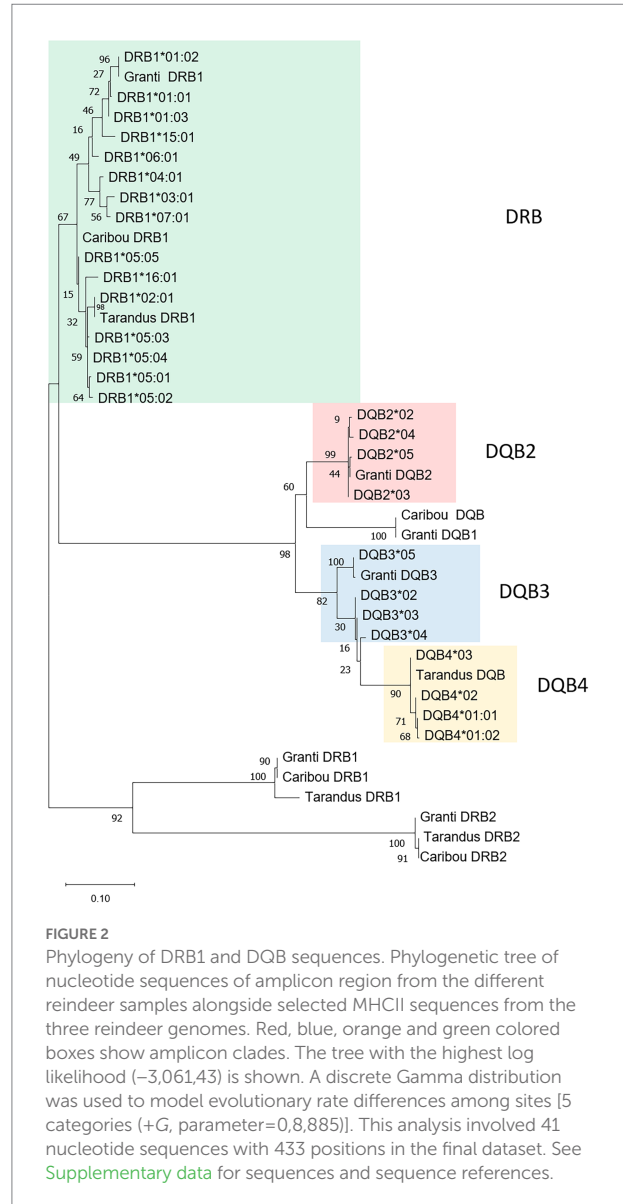
the caribou genome had a scaffold linking MHCI and MHCII genes, while only the tarandus genome had a scaffold linking all MHCII genes.

All three genomes displayed three DRB loci where the DRB2 locus seems like a pseudogene in all genomes (Figure 4; Supplementary data 4.1). Previous studies of MHCII DRB polymorphism in reindeer (Kennedy et al., 2011; Taylor et al., 2012; Gagnon et al., 2020) have all used genomic DNA and amplification of DRB beta 1 sequences with the LA31 and LA32 primer pair designed for bovines (Sigurdardóttir et al., 1991). These studies amplify sequences originating from the gene defined as DRB1, which seem highly polymorphic (Mikko et al., 1999; Kennedy et al., 2011; Taylor et al., 2012; Gagnon et al., 2020). We therefore chose this gene for amplicon sequencing. Location of primers are shown in Supplementary data 4.2.

FIGURE 3
Phylogeny of all MHCI sequences. Phylogeny of sequences from both genomic as well as expressed MHCI sequences. Color-coded boxes show clades with A1 and A5 amplicon sequences. The tree with the highest log likelihood (−1762,64) is shown. A discrete Gamma distribution was used to model evolutionary rate differences among sites [5 categories (+*G*, parameter=0,4,620)]. The rate variation model allowed for some sites to be evolutionarily invariable ([+*I*], 35,14% sites). This analysis involved 40 nucleotide sequences with 360 positions in the final dataset. See Supplementary data for sequences and sequence references.

Although Kennedy et al. (2011) have sequenced DQB variants from reindeer; we could not find these sequences in the NCBI database. As there is no expressed support in NCBI for DQB or MHCI genes, gene and primer choice rely on sequences identified in genomes. The caribou and tarandus genomes only display single DQB genes while the granti genome has three DQB genes (Figure 4; Supplementary data 4). We chose to design primers that should amplify sequences from all DQB genes.

For the MHCI genes, it was more complicated to identify relevant amplicons due to incomplete gene predictions or possibly existence of varying number of pseudogenes in the three genomes analyzed. Genomic organization and sequence phylogenies from the three reindeer genomes supported the presence of 5–7 MHCI genes (Figures 1, 4;

Supplementary data 4). Comparatively, genomes from cattle subspecies contain a varying number of MHCI genes ranging from 8 to10 genes (Schwartz et al., 2022). The sequence defined as A1 in the granti genome matches an A1 gene sequence in caribou, but does not have a match in tarandus. Gene sequences here defined as A2 and A3 both have internal stop codons in all three genomes, indicating pseudogenes. The gene sequence defined as A6 is also a pseudogene present in granti and caribou, but is not present in tarandus. An MHC-like sequence residing in between the A2 and A3 genes in all three genomes, has non-classical characteristics showing high sequence identity with the non-classical (NC) cattle NC10 gene sequence (Figures 1, 4). The A5 gene sequence is present in all three genomes, while the A4 gene sequence is present in the granti and tarandus genomes only. Without *a priori* knowledge about which genes are expressed and polymorphic, we chose to test primers for the three potentially classical polymorphic genes A1, A4 and A5 (Supplementary data 4.2). Amplification with primers for the A4 gene were not successful and were not included further.

MHC class II DRB1

Of the three DRB loci found in all reindeer genomes, the *DRB2* locus is a pseudogene with internal stop codons, while the polymorphic content of the *DRB3* gene is currently uncharacterized. As the *DRB1* locus has previously been defined as polymorphic (Mikko et al., 1999; Kennedy et al., 2011; Taylor et al., 2012; Gagnon et al., 2020), we chose to study this locus. No animal displayed more than two *DRB1* sequences with high support suggesting the primers were restricted to one locus. In our 89 animals, we identified 15 *DRB1* alleles, all clustering with the reindeer gene sequences defined as *DRB1* (Mikko et al., 1999; Figure 2; Supplementary data 5.1). For our new alleles, we continued the *DRB1* nomenclature initiated by Mikko and co-workers (Mikko et al., 1999).

Two of the defined alleles were found in one animal only, i.e., *DRB*01:03* and *DRB*05:04* and may as such represent PCR or sequencing errors. However, at least for *DRB*01:03* there are two other DRB sequences from reindeer in the NCBI database with an identical sequence, suggesting it is a true allele (AAB66611.1 and QCU79926.1).

The semi-domesticated population with 62 animals displayed 12 *DRB1* alleles (Table 1; Supplementary data 5.1). The 18 wild reindeer from Langfjella had nine alleles, where all but *DRB1*05:04* are shared with the semi-domesticated population. The nine individuals from Svalbard only displayed two DRB1 alleles, both unique to this population.

The main sequence diversity between alleles resides in the beta 1 domain with only a single amino acid difference in the leader sequence and in the beta 2 domain included in our amplicon (Supplementary data 5.1). Amino acid sequence identity in the beta 1 domain ranges from 80 to 99% with 22 variable positions (data not shown).
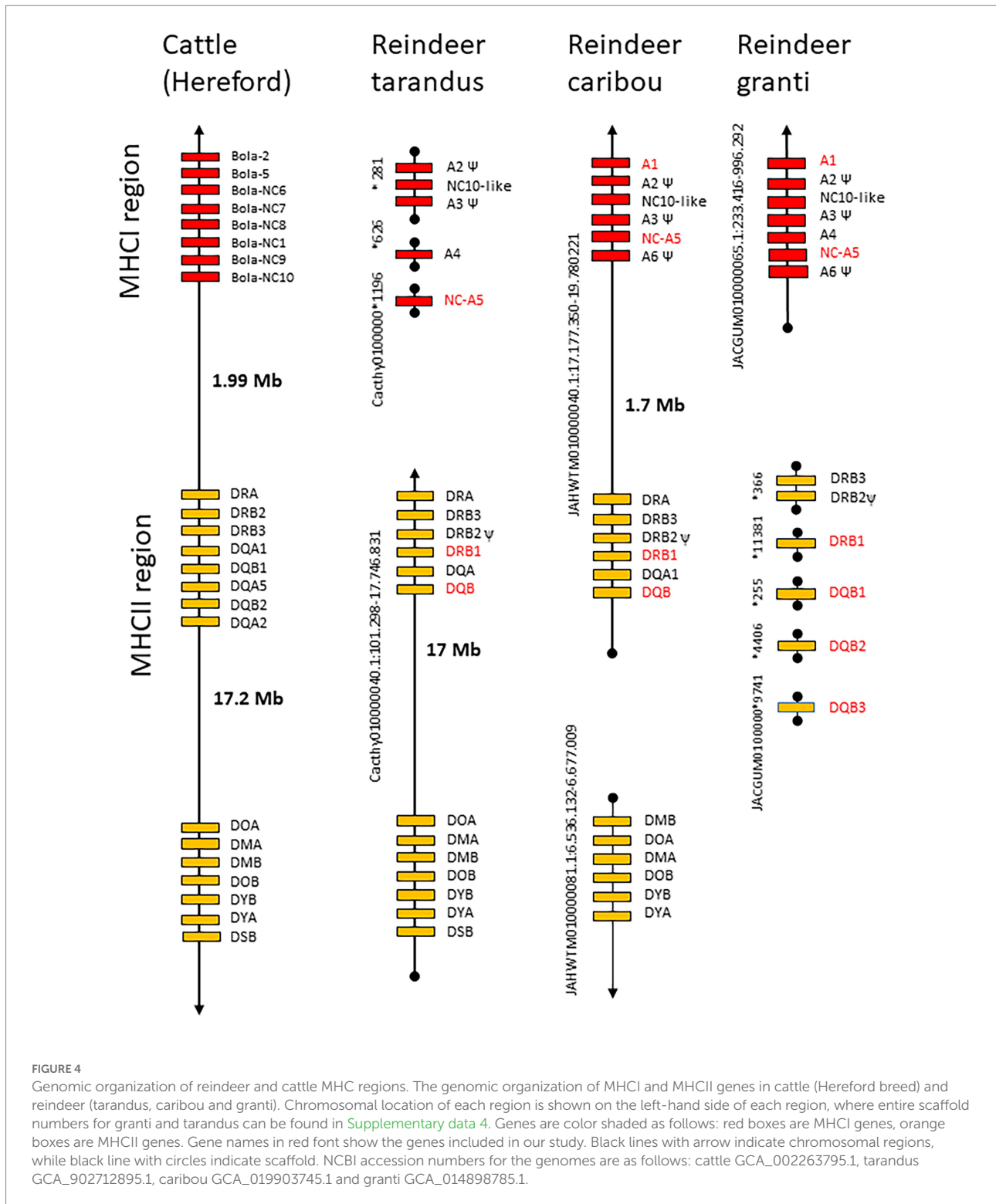
FIGURE 4
Genomic organization of reindeer and cattle MHC regions. The genomic organization of MHCI and MHCII genes in cattle (Hereford breed) and reindeer (tarandus, caribou and granti). Chromosomal location of each region is shown on the left-hand side of each region, where entire scaffold numbers for granti and tarandus can be found in Supplementary data 4. Genes are color shaded as follows: red boxes are MHCI genes, orange boxes are MHCII genes. Gene names in red font show the genes included in our study. Black lines with arrow indicate chromosomal regions, while black line with circles indicate scaffold. NCBI accession numbers for the genomes are as follows: cattle GCA_002263795.1, tarandus GCA_902712895.1, caribou GCA_019903745.1 and granti GCA_014898785.1.

The alleles *DRB1*05:01*, *DRB1*05:03*, *DRB1*15:01*, *DRB1*16:01* were unique for semi-domesticated reindeer, while *DRB1*05:04* was unique to wild reindeer from Langfjella (Supplementary data 5.1). *DRB1*02:01* and *DRB1*04:01* were the most frequent alleles in both semi-domesticated and wild reindeer. The semi-domesticated reindeer had 17 animals homozygous for

given *DRB1* alleles while the wild reindeer had five homozygous animals. *DRB1*04:01* was also the most frequent allele amongst the 22 *DRB1* homozygous mainland reindeer.

Eight of the individuals from Svalbard were homozygous for *DRB1*06:01*, while one individual was heterozygous for *DRB1*06:01* and *DRB1*01:03*. For the other two populations,

there were more homozygous animals than expected with *DRB1*04:01:01* homozygous animals being most pronounced (Figure 5; Table 2). The majority of these homozygous *DRB1*04:01:01* alleles originate from population 1 (Supplementary data 7).

MHC class II DQB results

In the three analyzed genomes, there was only one DQB gene in tarandus and caribou, but three genes in granti. As all mainland reindeer included in the present study were tarandus we expected one DQB gene, but our primers were designed to amplify sequences from all three genes.

TABLE 1 MHC allele diversity and allele richness per reindeer population.

	N	*DRB1* alleles	*A1*	*A5* alleles	Mean Ar	Total Ar
Population 1	62	12 (6.7)	18 (10.09)	4 (2.60)	6.46	19.39
Population 2	18	9 (6.8)	13 (8.96)	4 (3.53)	6.42	19.27
Population 3	9	2 (1.76)	2 (1.76)	1 (1)	1.5	4.51

N defines the number of animals per each of the semi-domesticated (Population 1, *Rangifer tarandus tarandus*), wild mainland (Population 2, *Rangifer tarandus tarandus*) and wild Svalbard (Population 3, *Rangifer tarandus platyrhynchus*) reindeer populations, all sampled in 2021. Number of MHC alleles with allelic richness (Ar) in parenthesis are shown for each of the *DRB1* (MHCII), *A1* (MHCI) and *A5* (MHCI) genes. Mean Ar is calculated across the three loci while total Ar relates to the sum of all three loci for each of the three populations.

We found 11 different DQB sequence variants in the 89 reindeer. No animal displayed more than two strongly supported sequences, which translates to one expressed DQB gene in accordance with the single DQB locus found in the caribou and tarandus genomes (Figure 4; Supplementary data 5.2). One sequence, DQB3*04, was found in one animal only and may represent a PCR or sequencing artefact. It differs from the remaining DQB3 sequences in two amino acids and needs further verification to be defined as a genuine expressed MHC variant.

Based on sequence phylogenies, the DQB sequences cluster into three clades (Figure 2). Four sequences cluster with the granti-DQB2 sequence here defined as DQB2*01 and 4 with the tarandus-DQB here defined as DQB4*01 sequence, both clusters supported by bootstrap values above 90%. Four additional sequences branch out as a separate clade linked to the granti-DQB3 here defined as DQB4*01 sequence. Here, the sequences defined as DQB3*02–04 are strongly supported while the DQB3*05 sequence is on a separate branch alongside granti-DQB3 with less bootstrap support. This contradicts our expectation of all mainland DQB sequences clustering with the single genomic tarandus DQB sequence. Amino acid sequence identity in the beta 1 domain ranged from 73 to 99% spread over 28 amino acid positions (Supplementary data 5.2; data not shown). Although we cannot link expressed sequence to one given locus, we refer to them as alleles originating from the hypothetical DQB2, DQB3 and DQB4 loci. Without being able to define these sequences as true alleles, we did not include DQB sequences in our population analysis.
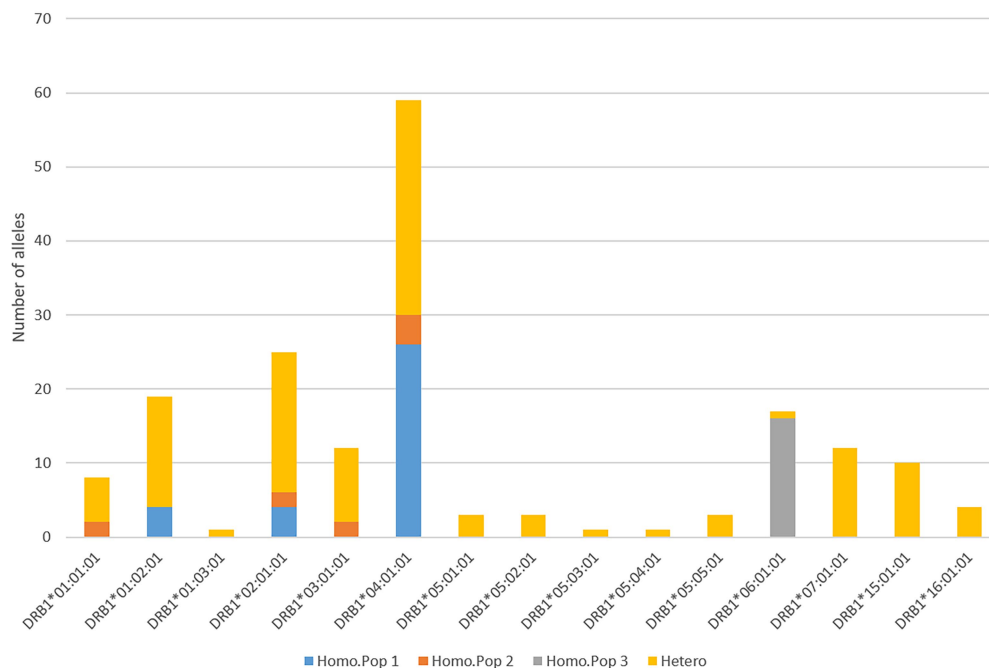

FIGURE 5

Homozygosity at DRB1 locus. Diagram shows the number homozygous *DRB1* alleles per each of the three populations where blue defines homozygous alleles from population 1 (Homo.Pop 1), red from population 2 1 (Homo.Pop 2), and grey from population 3 1 (Homo.Pop 3), while remaining alleles in heterozygous (hetero) pairs are show using orange.

TABLE 2 Expected versus observed MHC heterozygosity.

	He	Ho	% difference
A1	0.93	0.39	57.6
A5	0.53	0.44	17.4
DRB1	0.83	0.66	20.5

Population genetic estimates [expected (He) and observed (Ho) heterozygosity] across all three populations of reindeer (*Rangifer tarandus tarandus* and *Rangifer tarandus platyrhynchus*) sampled in 2021. The percentage differences between He and Ho for each locus are also presented.

In the semi-domesticated reindeer, there were nine different DQB alleles, where three sequences are unique and six are shared with wild reindeer (Supplementary data 5.2). Only one allele was unique to wild reindeer, i.e., DQB3*03. In both semi-domesticated and wild reindeer, DQB2*02 was by far the most frequent allele. All eight Svalbard reindeer were homozygous for the unique allele DQB3*02. We did not successfully amplify a DQB sequence from one of the Svalbard reindeer.

MHC class II DRB-DQB haplotypes

Haplotypes could provide a clue as to the number of DQB genes amplified in our dataset. Based on what is known from other species, as well as the tarandus and caribou genomes (Figure 4), the DRB and DQB genes are closely linked and should segregate as specific haplotypes of allelic gene variants. There are indeed a considerable number of discernible combinations of DRB1 and DQB alleles in our material. The most frequent combination is *DRB1*04:01-DQB2*02* present in 43 animals of which 15 animals are homozygous (Supplementary data 5.2). Other frequent haplotypes are *DRB1*01:02-DQB2*02*, *DRB1*02:01-DQB4*03*, *DRB1*06:01*-DQB3*02, *DRB1*07:01-DQB2*02* and *DRB1*15:01*-DQB4*01. Being confident that we only amplify sequences from the *DRB1* locus, this would imply that the *DRB1* locus co-segregates with alleles from each of the three loci we tentatively denoted DQB2, DQB3 and DQB4.

MHC class I

Our primers designed for the *A1* and *A5* loci were successful in our initial testing hence these genes were included for amplicon sequencing. Primers for the *A4* gene provided no PCR product and were not included further, while the *A2* and *A3* were both pseudogenes with internal stop codons in all analyzed genomes. The A5 primers showed a locus that was virtually non-polymorphic with two different alleles defined as *A5*01:01* and *A5*01:02* varying in one amino acid only (Figure 3; Supplementary data 5.3). Sequences from six mainland reindeer displayed a pseudogene in addition to a functional *A5*01:01* or *A5*01:02* allele. The reindeer from Svalbard were all homozygote for *A5*01:02*. Three reindeer from Langfjella also displayed one nucleotide difference in the

A5*01:02 sequence that did not affect the translated amino acid sequence (e.g., V0297) and these alleles were designated A5*01:02:02. This A5 gene is then defined as a non-classical gene widely expressed in all tissues and animals and we renamed this locus to *NC-A5*.

The other MHCI gene we designed primers for, here defined as *A1*, provided PCR products in all animals. No reindeer displayed more than two sequences with high matching numbers of reads suggesting that all sequences originate from one locus. We found 22 sequence variants in the 89 reindeer (Supplementary data 5.3). Sequence phylogenies support the notion of these sequences being allelic variants (Figure 3) and we chose our nomenclature accordingly. We conclude that the tarandus genome assembly lacks the *A1* gene. Amino acid sequence identity between alleles in the alpha 1 domain ranges from 76 to 99% distributed over 21 amino acid positions (Supplementary data 5.3). Two sequences are found in one animal only and may represent a PCR or sequencing artefact. *A1*11:01* differs from *A1*11:02* in three amino acids where two are located in positions also polymorphic in other alleles. The other allele found in only one animal is *A1*14:01:03*, which differs from *A1*14:01:01* and *A1*14:01:02* in two nucleotides not affecting the amino acid sequence.

The semi-domesticated reindeer displayed 18 alleles where six alleles were unique, as they were not found in wild reindeer from Langfjella or Svalbard (Supplementary data 5.3). The Langfjella reindeer displayed 13 alleles where 11 were shared with the semi-domesticated population while two were unique. Allele frequencies varied between these two mainland populations where *A1*06:01* and *A1*14:01:02* were the most frequent in the semi-domesticated reindeer, while *A1*07:01* was the most frequent allele in Langfjella reindeer. Svalbard reindeer displayed two unique alleles, *A1*10:02* and *A1*11:01* where eight animals were homozygous for *A1*10:02*.

Despite being more polymorphic than the *DRB1* gene, the *A1* gene was homozygous in 54 out of 89 animals. This excess homozygosity is also visible in the analysis relating to expected versus observed homozygosity per individual locus (Table 2). Looking at individual alleles, homozygous *A1*06:01:01*, *A1*07:01:01*, *A1*13:01:01* and *A1*10:02:01* alleles are most pronounced (Figure 6).

Overall population and MHC assessment

Analysis of allele richness (Ar) adjusted for sample size showed the two mainland populations having comparable diversity while the Svalbard population had far less diversity (Table 1). The wild and semi-domesticated animals had a mean allelic richness of 6.46 and 6.42, while the comparable number for the Svalbard animals was only 1.50.

Based on the available genome information, the MHCI loci *A1* and *A5* are located approximately 1.7 megabases apart from
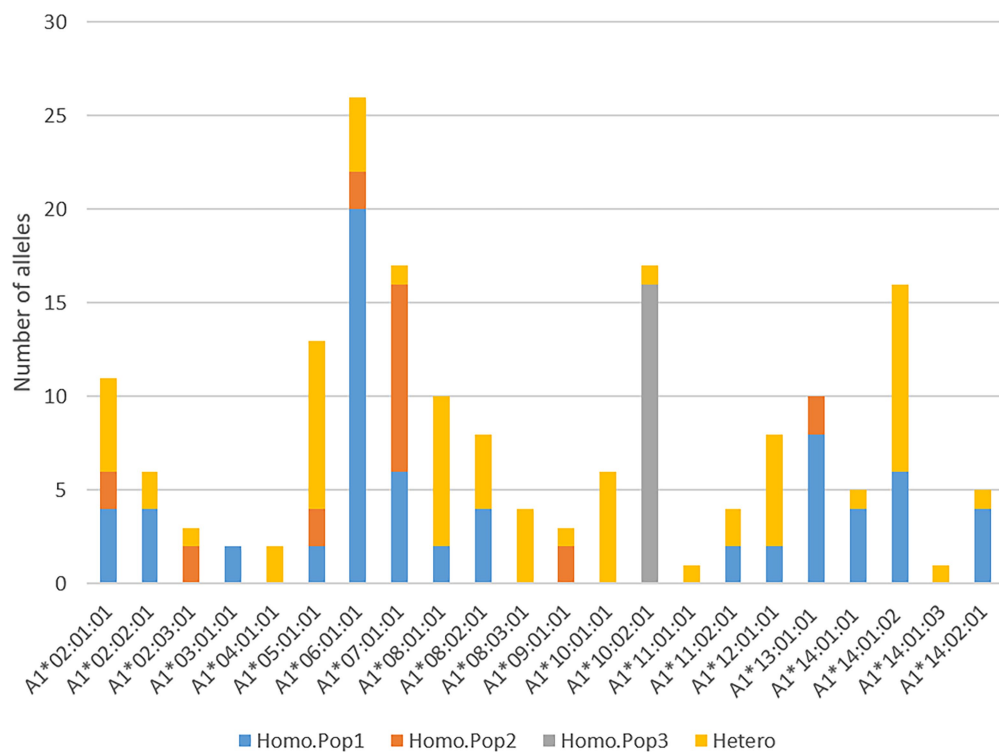
FIGURE 6
Homozygosity at A1 locus. Diagram shows the number of homozygous A1 alleles per each of the three populations where blue defines homozygous alleles from population 1 (Homo.Pop 1), red from population 2 (Homo.Pop 2), and grey from population 3 (Homo.Pop 3), while remaining alleles in heterozygous (Hetero) pairs in all populations are shown using orange.

the *DRB1* locus (Figure 4). We therefore investigated the existence of discernible MHCI and MHCII haplotypes, i.e., specific combinations of MHCI and MHCII alleles (Supplementary data 7). Although four haplotypes were detected, three of them contained one *DRB1* allele combined with three different A1 alleles. We thus conclude that there were no stable *DRB1*-DQB-*A1-A5* haplotypes in our mainland reindeer material. Sequence diversity is too low in the Svalbard animals to make any conclusions on haplotype stability.

Discussion

In the field of conservation genetics, genomic DNA and amplification of single exon domains is widely used as a foundation to define MHC diversity (Gangoso et al., 2012; Gillingham et al., 2017; Slade et al., 2017; Awadi et al., 2018; Qurkhuli et al., 2019). However, the precision in such diversity analyses may be low due to the variation in number of MHC genes as exemplified by human DRB. In this study, we identified the number of MHCI and MHCII genes in three available reindeer genomes and used this information to choose amplicons and design primers. This approach was then used to investigate

expressed, i.e., functional MHC diversity in Norwegian mainland wild and semi-domesticated reindeer as well as wild reindeer from the archipelago of Svalbard.

In all three genomes, we found three DRB genes where the *Rata-DRB1* gene is defined as highly polymorphic in previous studies (Mikko et al., 1999; Kennedy et al., 2011; Taylor et al., 2012; Gagnon et al., 2020). The *DRB2* gene is a pseudogene lacking parts of the beta 1 domain while the polymorphic content of the *DRB3* locus is unknown as our primers may not have amplified sequences from this locus. Cattle, another ungulate, has three DRB genes but only one gene is polymorphic and highly expressed (Aida et al., 1995). A study of the polymorphic content of the reindeer *DRB3* locus is necessary before we can decide whether reindeer, like cattle, only has one highly polymorphic expressed DRB locus.

Comparing the *DRB1* alleles identified in our study with those found in North-American caribou (Kennedy et al., 2011; Taylor et al., 2012; Gagnon et al., 2020), seven of the 15 *DRB1* alleles we identified are also present in North American animals (Supplementary data 6). When the glacial connection in the Beringia melted approximately 11,000 years ago (Elias et al., 1996; Jakobsson et al., 2017), sharing of genes between North American and Eurasian reindeer ended. This implies that the *DRB1* alleles

shared between North American and European reindeer are at least 11,000 years old. Complete conservation of allelic sequences over such a long time span suggests that these shared alleles are well adapted to pathogens encountered in boreal and Arctic areas.

For DQB, there were one to three genes in the three reindeer genomes analyzed (Figure 4). We were not able to assign the expressed DQB sequences to specific genes as our expressed DQB sequences cluster with each of the granti-DQB2, granti-DQB3 or the tarandus-DQB gene sequences. Haplotypic variation in number of expressed loci both within as well as between populations is quite common when it comes to MHC. For instance in cattle, there is haplotypic variation with one or two DQB genes (Fukunaga et al., 2020). In sheep, there seems to be two DQA and DQB genes (DQA1-DQB1 and DQA2-DQB2), but in some haplotypes the DQA1-DQB1 genes are replaced by DQA2-like and DQB2-like genes (Ballingall et al., 2018b). The fact that we only found two expressed DQB alleles in each animal supports the notion of a single expressed DQB gene in each animal, where the expressed gene is of a DQB2, DQB3 or DQB4 type. One could imagine DQB expression profiles to differ between tissues as seen in humans (Johansson et al., 2021), but we found no link between DQB sequence variant and tissues, i.e., blood, intestinal lymph nodes and kidney. A more thorough study is needed to clarify if the DQB alleles found in our study represent allelic versions of one locus, single genes with different evolutionary histories expressed in different haplotypes or three separate genes.

For MHCI, we found 5–7 genes in the three analyzed tarandus, granti and caribou genomes. The A2 and A3 genes are pseudogenes with internal stop codons in all genomes. We found little polymorphism in the non-classical A5 gene with only two amino acid alleles identified in the 89 animals investigated. To evaluate the reindeer MHCI genes further we compared against the MHCI genes defined in cattle (Schwartz et al., 2022). Based on sequence phylogeny, Rata-NC10-like is an orthologue of the non-classical cattle NC10 gene (Figure 1). This cattle NC10 gene is non-polymorphic, but expressed in all animals (Schwartz et al., 2022). We did not include this gene in our study, but expect it to be less polymorphic. So far, the function of this gene has not been studied in any species, but conservation in multiple ungulates suggests an essential function and makes it an interesting candidate for future studies. For the remaining genes, there is low bootstrap support between cattle and reindeer MHCI gene sequences suggesting there has been a lot of within species duplications and deletions.

The other MHCI gene included in our amplicon study was the A1 gene. This gene is highly polymorphic and expressed in all tissues included, i.e., blood (Langfjella), intestinal lymph nodes (semi-domesticated) and kidney (Svalbard). In cattle, there are haplotypes ranging from two to four polymorphic classical MHCI genes in addition to 5–6 non-classical genes while humans have a fixed number of three classical MHCI genes (Klein et al., 1990; Robbins et al., 1997; Schwartz et al., 2022). Based on our data,

reindeer have two pseudogenes (A2, A3), two non-classical genes (NC-A5 and NC10-like) and possibly two classical genes (A1, A4) where the polymorphic content of A4 needs to be defined in future studies.

There is no fixed number defining a healthy MHC diversity in any species, but it is generally assumed that the alleles present are adapted to the pathogens they encounter in their environment if they are left undisturbed by man-made or environmental factors. In our semi-domesticated, wild, and Svalbard reindeer, we identified 12, 9 and 2 alleles for DRB1, 9, 7 and 1 alleles for DQB and 18, 13 and 2 alleles for A1. Overall, in all study areas, MHCI was the most polymorphic gene with 22 different alleles while DRB1 and DQB had 15 and 11 different alleles, respectively.

When comparing the MHC diversity between our three study areas, sample composition and sample size need to be considered. Our sample size varied from nine Svalbard reindeer to 62 semi-domesticated reindeer. At relatively low sample sizes, the number of detected alleles is expected to increase with an increasing number of sampled individuals. Therefore, the number of alleles detected in the two areas with lower sample size, Langfjella and Svalbard, could have been higher if more individuals had been sampled. However, our allele richness analysis uses rarefaction to adjust for differences in samples size and this analysis also shows the Svalbard population to be far less diverse than the two mainland populations (Table 1). A study of allele richness using microsatellites in six Svalbard reindeer subpopulations displayed similar diversity estimates with mean Ar ranging from 2,07 to 2,58 (Peeters et al., 2020). Most likely the wild reindeer population has experience some introgression from semi-domesticated reindeer while the Svalbard population is a different isolated subspecies.

Even though the majority of samples analysed in our study originated from semi-domesticated reindeer, the geographical distribution of these samples was limited. Herds of semi-domesticated reindeer are widely distributed across the mid and northern parts of both Norway and Sweden. In many of the border areas, there are both an overlap in distribution range, and an expected genetic exchange between Norwegian and Swedish herds. The MHC diversity found in our two neighbouring herds included in our study is therefore not expected to provide the total immune diversity found within the Norwegian metapopulation of semi-domesticated reindeer.

The samples from wild reindeer were collected from two sub-populations with limited genetic exchange (Kvie et al., 2019). Historically, these sub-populations have been genetically influenced by semi-domesticated animals introduced to the region (Røed et al., 2014; Kvie et al., 2019). The genetic exchange both between wild reindeer sub-populations, as well as between wild reindeer sub-populations and herds of semi-domestic reindeer, has been found to vary substantially (Kvie et al., 2019). Future studies should therefore aim to investigate both semi-domesticated as well as wild reindeer from wider geographical regions to conclude on the overall immune diversity in Norwegian reindeer.

To further evaluate the overall MHC diversity in our mainland study populations, we compared against MHC diversity found in previous studies. Mikko and co-workers studied *DRB1* diversity in 20 reindeer from Northern Norway and 20 animals from Central Norway where they found 5 and 6 alleles, respectively, (Mikko et al., 1999). All these *DRB1* alleles were present in our study where we added six new alleles. A study of *DRB1* polymorphism by Gagnon and co-workers included two declining Canadian migratory caribou populations with 245 and 146 animals (Gagnon et al., 2020). Overall, these two populations displayed 18 and 20 *DRB1* alleles, where 18 alleles were shared between populations. Such low numbers correlate well with the 99 and 70% decline in population size for these populations (Gagnon et al., 2020). Taylor and co-workers investigated *DRB1* polymorphism in both historic as well as contemporary high Arctic caribou from Ellesmere Island in Canada (Taylor et al., 2012). They analyzed 11 and 24 animals in each group and found 10 and 15 alleles, respectively. Kennedy and co-workers sampled caribou from five populations spread in Alaska and found 25 *DRB1* alleles in 114 animals (Kennedy et al., 2011).

Although Svalbard reindeer is a different subspecies, their MHC diversity is very different. A previous study found only two *DRB1* alleles based on a sample of 20 Svalbard reindeer from 1985 (Mikko et al., 1999). These are the same two alleles we identified in samples harvested 36 years later. This extremely low *DRB1* diversity is further supported by only one DQB allele and only two A1 alleles identified in nine animals. A low MHC diversity is further supported by a low overall genetic diversity (Peeters et al., 2020). Despite Svalbard reindeer being close to extinction at the beginning of the last century (Lønø, 1959), protection measures, natural recolonization and human-mediated reintroductions have contributed to a gradual increase in the total population size (Le Moullec et al., 2019). Today, Svalbard reindeer are found in considerable numbers as semi-isolated populations all across this remotely located archipelago (Le Moullec et al., 2019).

We could not find any stable *DRB1*-DQB-*A1-A5* haplotypes in our material (Supplementary data 7). This most likely results from frequent recombination within the 1.7 Mb region separating the class I and class II region as defined by the caribou genome (Figure 4). With frequent recombination between the MHCI and MHCII regions, different selective forces could operate on MHCI and MHCII genes. This would imply that studies of both MHCI and MHCII are essential for understanding how selection affects MHC diversity in reindeer.

Increasing the number of animals from a bottlenecked population does not seem to have increased the overall genetic diversity or the immune diversity. There are examples of populations that thrive with low MHC and genetic diversity (Munguia-Vega et al., 2007; Robinson et al., 2022). A small Mexican porpoise population (*Phocoena sinus*, or vaquita) with limited overall genetic and MHC diversity does not seem affected by inbreeding or pathogens. Several species of teleosts have also lost their MHC class II, but seem to have replaced

this function by multiple MHC class I genes with potentially dual functions (Star et al., 2011). However, these examples are most likely the exception rather than the rule as most vertebrates studied so far have a broad MHC diversity.

Our data displayed an unexpected number of homozygous animals in particular for the A1 locus where 54 of 89 animals were homozygous (Table 1; Figure 6). Disregarding the homozygous animals originating from the Svalbard population, the *A1*06:01:01*, *A1*07:01:01* and *A1*13:01:01* alleles display by far more homozygous animals than expected. We do not have population genetic data to show kinship, but animal husbandry may have promoted homozygosity in the semi-domesticated population1. Population 2 only consists of wild reindeer although they may have been influenced by semi-domestic reindeer sharing the same habitat.

Another explanation for the increased homozygosity could be that MHC composition is influenced by selection, thus increasing homozygosity for specific alleles at the MHCI locus. Recently it has been shown that both human as well as chicken MHCI alleles can be defined as specialists or generalists (Chappell et al., 2015; Kaufman, 2018). Specialist alleles have strict peptide binding capacity while generalists can bind a broader specter of peptides. In humans, a specialist allele *HLA-B*57:01* confers resistance to AIDS while in chicken the generalist *BF2*2101* confers resistance towards Mareks disease. A consequence of presenting many peptides could be a broader T-cell response, needed for protection against certain pathogens while other pathogens require a specific T-cell response for fighting the infection. In humans, a recent study found that in pathogen-rich geographical regions, humans are more likely to carry highly promiscuous MHC class II alleles, thus enabling protection against many pathogens (Manczinger et al., 2019). The peptide binding ability of reindeer MHCI alleles is currently unknown, but increased homozygosity for certain MHC alleles might have evolved through adaptation to certain pathogens. As exemplified by the human adaptation to pathogen-rich regions with increasing numbers of generalist MHC alleles, a similar approach could be in place for populations such as reindeer showing limited MHC diversity.

Low immune diversity would not be problematic assuming the alleles are adapted to all pathogens present in the environment or can accommodate peptides from a wide variety of pathogens. However, if there is a change in pathogens that elude binding to existing MHC alleles, low immune diversity could be fatal. Svalbard reindeer seem to have such a limited immune diversity and could be particularly vulnerable if exposed to new pathogens.

Regarding conservation measures to increase this diversity the Svalbard reindeer is endemic to Svalbard and is found in isolated or semi-isolated population across the entire archipelago (Peeters et al., 2020). In accordance with the purpose of the Svalbard Environmental Protection Act,[2] the landscape, flora and fauna

---

2 https://www.regjeringen.no/en/dokumenter/svalbard-environmental-protection-act/id173945/

should be preserved virtually untouched and with minimal influence of human presence. This means that the more traditional management regimes involving population regulation through harvesting is not valid for Svalbard. The intended absence of human footprint in population dynamics also means that species distribution and population viability should not be affected by management actions.

Coal mining was previously the most important industry in Svalbard, but the mines and the associated activity has now to a large degree been closed down. Today, tourism has become the largest employer and is a rapidly growing sector in Svalbard (Saville, 2022). The increasing number of visitors both within and outside settlements, represents a potential risk related to transmission of contamination of infectious diseases. In this context, increased awareness of the vulnerability of endemic species is highly relevant information. Such information is not expected to cause altered wildlife management strategies, but will hopefully further increase the awareness of the need to regulate human presence and impact in this pristine and vulnerable ecosystem.

Conclusion

Previous studies have used genomic DNA to define *DRB1* diversity in reindeer. Here we use reindeer genomes to identify one expressed classical MHC class I gene (*A1*) and two non-classical MHC class I genes (*NC-A5* and *NC10-like*). For class II we verify that the previously amplified DRB sequences originate from the classical expressed locus while different DQB genes seem expressed in different haplotypes. We found no stable MHCI-MHCII haplotypes suggesting frequent recombination between these regions. Collectively our data show a medium MHC diversity in semi-domesticated and wild Norwegian mainland reindeer. On the other hand, animals from Svalbard have experienced a severe bottleneck resulting in a low MHC diversity, placing them at risk if the pathogenic pressure changes.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/, PRJNA832088.

Ethics statement

The animal study was reviewed and approved by Norwegian Food Safety Authority (permit no. 20/227896). The harvest of animals was approved by the Governor of Svalbard (permit no. 16/01632–40).

Author contributions

UG and ML designed the experiment. IN, JV, KM, VV, and CR were responsible for providing the samples. ML performed the sample and library prep. CB performed the sequencing. AS and ML performed the bioinformatics analysis. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2022.1058674/full#supplementary-material

References

Adamack, A. T., and Gruber, B. (2014). PopGenReport: simplifying basic population genetic analyses in R. *Methods Ecol. Evol.* 5, 384–387. doi: 10.1111/2041-210x.12158

Aida, Y., Niimi, M., Asahina, M., Okada, K., Nakai, Y., and Ogimoto, K. (1995). Identification of a new bovine MHC class II DRB allele by nucleotide sequencing and an analysis of phylogenetic relationships. *Biochem. Biophys. Res. Commun.* 209, 981–988. doi: 10.1006/bbrc.1995.1594

Andersson, G., Larhammar, D., Widmark, E., Servenius, B., Peterson, P. A., and Rask, L. (1987). Class II genes of the human major histocompatibility complex. Organization and evolutionary relationship of the DR beta genes. *J. Biol. Chem.* 262, 8748–8758. doi: 10.1016/S0021-9258(18)47480-7

Awadi, A., Ben Slimen, H., Smith, S., Knauer, F., Makni, M., and Suchentrunk, F. (2018). Positive selection and climatic effects on MHC class II gene diversity in hares (Lepus capensis) from a steep ecological gradient. *Sci. Rep.* 8:11514. doi: 10.1038/s41598-018-29657-3

Ballingall, K. T., Bontrop, R. E., Ellis, S. A., Grimholt, U., Hammond, J. A., Ho, C.-S., et al. (2018a). Comparative MHC nomenclature: report from the ISAG/IUIS-VIC committee 2018. *Immunogenetics* 70, 625–632. doi: 10.1007/s00251-018-1073-3

Ballingall, K. T., Lantier, I., Todd, H., Lantier, F., and Rocchi, M. (2018b). Structural and functional diversity arising from intra- and inter-haplotype combinations of duplicated DQA and B loci within the ovine MHC. *Immunogenetics* 70, 257–269. doi: 10.1007/s00251-017-1029-z

Behl, J. D., Verma, N. K., Tyagi, N., Mishra, P., Behl, R., and Joshi, B. K. (2012). The major histocompatibility complex in bovines: a review. *ISRN veterinary science* 2012:872710. doi: 10.5402/2012/872710

Chappell, P., Meziane el, K., Harrison, M., Magiera, L., Hermann, C., Mears, L., et al. (2015). Expression levels of MHC class I molecules are inversely correlated with promiscuity of peptide binding. *elife* 4:e05345. doi: 10.7554/eLife.05345

de Muinck, E. J., Trosvik, P., Gilfillan, G. D., Hov, J. R., and Sundaram, A. Y. M. (2017). A novel ultra high-throughput 16S rRNA gene amplicon sequencing library preparation method for the Illumina HiSeq platform. *Microbiome* 5:68. doi: 10.1186/s40168-017-0279-1

Doxiadis, G. G., Hoof, I., de Groot, N., and Bontrop, R. E. (2012). Evolution of HLA-DRB genes. *Mol. Biol. Evol.* 29, 3843–3853. doi: 10.1093/molbev/mss186

El Mousadik, A., and Petit, R. J. (1996). High level of genetic differentiation for allelic richness among populations of the argan tree [Argania spinosa (L.) Skeels] endemic to Morocco. *Theor. Appl. Genet.* 92, 832–839. doi: 10.1007/BF00221895

Elbers, A. R., Koenraadt, C. J., and Meiswinkel, R. (2015). Mosquitoes and Culicoides biting midges: vector range and the influence of climate change. *Rev. Sci. Tech.* 34, 123–137. doi: 10.20506/rst.34.1.2349

Elias, S. A., Short, S. K., Nelson, C. H., and Birks, H. H. (1996). Life and times of the Bering land bridge. *Nature* 382, 60–63. doi: 10.1038/382060a0

FASTX-Toolkit (2010). *The FASTX-toolkit is a collection of command line tools for short-reads FASTA/FASTQ files preprocessing.* [Online]. Available: http://hannonlab.cshl.edu/fastx_toolkit/

Festa-Bianchet, M., Ray, J. C., Boutin, S., Côté, S. D., and Gunn, A. (2011). Conservation of caribou (*Rangifer tarandus*) in Canada: an uncertain future1This review is part of the virtual symposium "flagship species – flagship problems" that deals with ecology, biodiversity and management issues, and climate impacts on species at risk and of Canadian importance, including the polar bear (*Ursus maritimus*), Atlantic cod (*Gadus morhua*), piping plover (*Charadrius melodus*), and caribou (*Rangifer tarandus*). *Can. J. Zool.* 89, 419–434. doi: 10.1139/z11-025

Flagstad, O., and Røed, K. H. (2003). Refugial origins of reindeer (*Rangifer tarandus* L.) inferred from mitochondrial DNA sequences. *Evolution* 57, 658–670. doi: 10.1111/j.0014-3820.2003.tb01557.x

Fukunaga, K., Yamashita, Y., and Yagisawa, T. (2020). Copy number variations in BOLA-DQA2, BOLA-DQB, and BOLA-DQA5 show the genomic architecture and haplotype frequency of major histocompatibility complex class II genes in Holstein cows. *Hla* 96, 601–609. doi: 10.1111/tan.14086

Gagnon, M., Yannic, G., Boyer, F., and Cote, S. D. (2020). Adult survival in migratory caribou is negatively associated with MHC functional diversity. *Heredity (Edinb)* 125, 290–303. doi: 10.1038/s41437-020-0347-3

Gangoso, L., Alcaide, M., Grande, J. M., Munoz, J., Talbot, S. L., Sonsthagen, S. A., et al. (2012). Colonizing the world in spite of reduced MHC variation. *J. Evol. Biol.* 25, 1438–1447. doi: 10.1111/j.1420-9101.2012.02529.x

Gillingham, M. A., Bechet, A., Courtiol, A., Rendon-Martos, M., Amat, J. A., Samraoui, B., et al. (2017). Very high MHC class IIB diversity without spatial differentiation in the mediterranean population of greater flamingos. *BMC Evol. Biol.* 17:56. doi: 10.1186/s12862-017-0905-3

Goulder, P. J., and Walker, B. D. (2012). HIV and HLA class I: an evolving relationship. *Immunity* 37, 426–440. doi: 10.1016/j.immuni.2012.09.005

Grimholt, U., Larsen, S., Nordmo, R., Midtlyng, P., Kjoeglum, S., Storset, A., et al. (2003). MHC polymorphism and disease resistance in Atlantic salmon (*Salmo salar*); facing pathogens with single expressed major histocompatibility class I and class II loci. *Immunogenetics* 55, 210–219. doi: 10.1007/s00251-003-0567-8

Gundersen, V., Myrvold, K. M., Kaltenborn, B. P., Strand, O., and Kofinas, G. (2022). A review of reindeer (*Rangifer tarandus tarandus*) disturbance research in northern Europe: towards a social-ecological framework? *Landsc. Res.* 1-17, 1–17. doi: 10.1080/01426397.2022.2078486

IPD-MHC. (2021). *IPD-MHC release 3.8.0.0 (2021-12) build 204* [Online]. Available: https://www.ebi.ac.uk/ipd/mhc/

Jakobsson, M., Pearce, C., Cronin, T. M., Backman, J., Anderson, L. G., Barrientos, N., et al. (2017). Post-glacial flooding of the Bering land bridge dated to 11 cal ka BP based on new geophysical and sediment records. *Clim. Past* 13, 991–1005. doi: 10.5194/cp-13-991-2017

Johansson, T., Yohannes, D. A., Koskela, S., Partanen, J., and Saavalainen, P. (2021). HLA RNA sequencing with unique molecular identifiers reveals high allele-specific variability in mRNA expression. *Front. Immunol.* 12:629059. doi: 10.3389/fimmu.2021.629059

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275

Jore, S., Vanwambeke, S. O., Viljugrein, H., Isaksen, K., Kristoffersen, A. B., Woldehiwet, Z., et al. (2014). Climate and environmental change drives Ixodes ricinus geographical expansion at the northern range margin. *Parasit. Vectors* 7:11. doi: 10.1186/1756-3305-7-11

Kaufman, J. (2018). Generalists and specialists: a new view of how MHC class I molecules fight infectious pathogens. *Trends Immunol.* 39, 367–379. doi: 10.1016/j.it.2018.01.001

Kennedy, L. J., Modrell, A., Groves, P., Wei, Z., Single, R. M., and Happ, G. M. (2011). Genetic diversity of the major histocompatibility complex class II in Alaskan caribou herds. *Int. J. Immunogenet.* 38, 109–119. doi: 10.1111/j.1744-313X.2010.00973.x

Kjøglum, S., Larsen, S., Bakke, H. G., and Grimholt, U. (2008). The effect of specific MHC class I and class II combinations on resistance to furunculosis in Atlantic salmon (*Salmo salar*). *Scand. J. Immunol.* 67, 160–168. doi: 10.1111/j.1365-3083.2007.02052.x

Klein, J., Gutknecht, J., and Fischer, N. (1990). The major histocompatibility complex and human evolution. *Trends Genet.* 6, 7–11. doi: 10.1016/0168-9525(90)90042-5

Kulski, J. K., Shiina, T., Anzai, T., Kohara, S., and Inoko, H. (2002). Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity and complexity from shark to man. *Immunol. Rev.* 190, 95–122. doi: 10.1034/j.1600-065x.2002.19008.x

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Kvie, K. S., Heggenes, J., Bårdsen, B.-J., and Røed, K. H. (2019). Recent large-scale landscape changes, genetic drift and reintroductions characterize the genetic structure of Norwegian wild reindeer. *Conserv. Genet.* 20, 1405–1419. doi: 10.1007/s10592-019-01225-w

Laaksonen, S., Solismaa, M., Kortet, R., Kuusela, J., and Oksanen, A. (2009). Vectors and transmission dynamics for Setaria tundra (Filarioidea; Onchocercidae), a parasite of reindeer in Finland. *Parasit. Vectors* 2:3. doi: 10.1186/1756-3305-2-3

Landbruksdirektoratet. (2021). *Ressursregnskap for reindriftsnæringen.* Available at: https://www.landbruksdirektoratet.no/nb/filarkiv/rapporter/Ressursregnskap%20for%20reindriftsn%C3%A6ringen%202020-21%20versjon%202.pdf/_/attachment/inline/016e952-6cb4-4bdd-893b-915b93d95c63:77d7b3e921a90632b1fff6070ab9c23b4049c194/Ressursregnskap%20for%20reindriftsn%C3%A6ringen%202020-21%20versjon%202.pdf

Langefors, A., Lohm, J., Grahn, M., Andersen, O., and von Schantz, T. (2001). Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. *Proceedings. Biological sciences* 268, 479–485. doi: 10.1098/rspb.2000.1378

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Le Moullec, M., Pedersen, Å. Ø., Stien, A., Rosvold, J., and Hansen, B. B. (2019). A century of conservation: the ongoing recovery of Svalbard reindeer. *J. Wildl. Manag.* 83, 1676–1686. doi: 10.1002/jwmg.21761

Li, C., Huang, R., Nie, F., Li, J., Zhu, W., Shi, X., et al. (2020). Organization of the Addax Major Histocompatibility Complex Provides Insights into Ruminant Evolution. *Front. Immunol.* 11:260. doi: 10.3389/fimmu.2020.00260

Lønø, O. (1959). Reinen på Svalbard. *Fauna* 12, 40–70.

Maccari, G., Robinson, J., Bontrop, R. E., Otting, N., de Groot, N. G., Ho, C.-S., et al. (2018). IPD-MHC: nomenclature requirements for the non-human major histocompatibility complex in the next-generation sequencing era. *Immunogenetics* 70, 619–623. doi: 10.1007/s00251-018-1072-4

Magoč, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507

Manczinger, M., Boross, G., Kemeny, L., Muller, V., Lenz, T. L., Papp, B., et al. (2019). Pathogen diversity drives the evolution of generalist MHC-II alleles in human populations. *PLoS Biol.* 17:e3000131. doi: 10.1371/journal.pbio.3000131

Mikko, S., Røed, K., Schmutz, S., and Andersson, L. (1999). Monomorphism and polymorphism at Mhc DRB loci in domestic and wild ruminants. *Immunol. Rev.* 167, 169–178. doi: 10.1111/j.1600-065x.1999.tb01390.x

Munguia-Vega, A., Esquer-Garrigos, Y., Rojas-Bracho, L., Vazquez-Juarez, R., Castro-Prieto, A., and Flores-Ramirez, S. (2007). Genetic drift vs. natural selection in a long-term small isolated population: major histocompatibility complex class II variation in the Gulf of California endemic porpoise (*Phocoena sinus*). *Mol. Ecol.* 16, 4051–4065. doi: 10.1111/j.1365-294X.2007.03319.x

Mysterud, A., Jore, S., Østerås, O., and Viljugrein, H. (2017). Emergence of tick-borne diseases at northern latitudes in Europe: a comparative approach. *Sci. Rep.* 7:16316. doi: 10.1038/s41598-017-15742-6

Mysterud, A., Viljugrein, H., Lund, J. H. L. A., Lund, S. E., Rolandsen, C. M., and Strand, O. (2021). The relationship between quotas and harvest in the alpine reindeer population on Hardangervidda, Norway. *European J. Wildlife Res.* 67:100. doi: 10.1007/s10344-021-01542-x

Omazic, A., Bylund, H., Boqvist, S., Högberg, A., Björkman, C., Tryland, M., et al. (2019). Identifying climate-sensitive infectious diseases in animals and humans in northern regions. *Acta Vet. Scand.* 61:53. doi: 10.1186/s13028-019-0490-0

Palmer, M. V., Stoffregen, W. C., Rogers, D. G., Hamir, A. N., Richt, J. A., Pedersen, D. D., et al. (2004). West Nile Virus Infection in Reindeer (*Rangifer Tarandus*). *J. Vet. Diagn. Investig.* 16, 219–222. doi: 10.1177/104063870401600307

Paterson, S., Wilson, K., and Pemberton, J. M. (1998). Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population. *Proc. Natl. Acad. Sci. U. S. A.* 95, 3714–3719. doi: 10.1073/pnas.95.7.3714

Peeters, B., Le Moullec, M., Raeymaekers, J. A. M., Marquez, J. F., Røed, K. H., Pedersen, Å. Ø., et al. (2020). Sea ice loss increases genetic isolation in a high Arctic ungulate metapopulation. *Glob. Chang. Biol.* 26, 2028–2041. doi: 10.1111/gcb.14965

Plachy, J., Pink, J. R., and Hála, K. (1992). Biology of the chicken MHC (B complex). *Crit. Rev. Immunol.* 12, 47–79.

Qurkhuli, T., Schwensow, N., Brandel, S. D., Tschapka, M., and Sommer, S. (2019). Can extreme MHC class I diversity be a feature of a wide geographic range? The example of Seba's short-tailed bat (Carollia perspicillata). *Immunogenetics* 71, 575–587. doi: 10.1007/s00251-019-01128-7

Robbins, F., Hurley, C. K., Tang, T., Yao, H., Lin, Y. S., Wade, J., et al. (1997). Diversity associated with the second expressed HLA-DRB locus in the human population. *Immunogenetics* 46, 104–110. doi: 10.1007/s002510050248

Robinson, J. A., Kyriazis, C. C., Nigenda-Morales, S. F., Beichman, A. C., Rojas-Bracho, L., Robertson, K. M., et al. (2022). The critically endangered vaquita is not doomed to extinction by inbreeding depression. *Science* 376, 635–639. doi: 10.1126/science.abm1742

Røed, K. H., Bjørklund, I., and Olsen, B. J. (2018). From wild to domestic reindeer – genetic evidence of a non-native origin of reindeer pastoralism in northern Fennoscandia. *J. Archaeol. Sci. Rep.* 19, 279–286. doi: 10.1016/j.jasrep.2018.02.048

Røed, K. H., Bjørnstad, G., Flagstad, Ø., Haanes, H., Hufthammer, A. K., Jordhøy, P., et al. (2014). Ancient DNA reveals prehistoric habitat fragmentation and recent domestic introgression into native wild reindeer. *Conserv. Genet.* 15, 1137–1149. doi: 10.1007/s10592-014-0606-z

RStudioTeam. (2022). "RStudio: Integrated development environment for R". (Boston, MA, USA: RStudio, PBC).

Sánchez Romano, J., Grund, L., Obiegala, A., Nymo, I. H., Ancin-Murguzur, F. J., Li, H., et al. (2019). A multi-pathogen screening of captive reindeer (*Rangifer tarandus*) in Germany based on serological and molecular assays. *Front Vet Sci* 6:461. doi: 10.3389/fvets.2019.00461

Saville, S. M. (2022). Valuing time: tourism transitions in Svalbard. *Polar Record* 58:0055. doi: 10.1017/S0032247422000055

Schwartz, J. C., Maccari, G., Heimeier, D., and Hammond, J. A. (2022). Highly-contiguous bovine genomes underpin accurate functional analyses and updated nomenclature of MHC class I. *Hla* 99, 167–182. doi: 10.1111/tan.14494

Shiina, T., Hosomichi, K., Inoko, H., and Kulski, J. K. (2009). The HLA genomic loci map: expression, interaction, diversity and disease. *J. Hum. Genet.* 54, 15–39. doi: 10.1038/jhg.2008.5

Sigurdardóttir, S., Borsch, C., Gustafsson, K., and Andersson, L. (1991). Cloning and sequence analysis of 14 DRB alleles of the bovine major histocompatibility complex by using the polymerase chain reaction. *Anim. Genet.* 22, 199–209. doi: 10.1111/j.1365-2052.1991.tb00670.x

Slade, J. W., Sarquis-Adamson, Y., Gloor, G. B., Lachance, M. A., and MacDougall-Shackleton, E. A. (2017). Population differences at MHC do not explain enhanced resistance of song sparrows to local parasites. *J. Hered.* 108, esw082–esw134. doi: 10.1093/jhered/esw082

Solovyev, V., Kosarev, P., Seledsov, I., and Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* 7, S10.1–S1012. doi: 10.1186/gb-2006-7-s1-s10

Sommer, R. S., Kalbe, J., Ekström, J., Benecke, N., and Liljegren, R. (2014). Range dynamics of the reindeer in Europe during the last 25,000 years. *J. Biogeogr.* 41, 298–306. doi: 10.1111/jbi.12193

Star, B., Nederbragt, A. J., Jentoft, S., Grimholt, U., Malmstrom, M., Gregers, T. F., et al. (2011). The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477, 207–210. doi: 10.1038/nature10342

Strand, O., Nilsen, E. B., Solberg, E. J., and Linnell, J. C. D. (2012). Can management regulate the population size of wild reindeer (*Rangifer tarandus*) through harvest? *Can. J. Zool.* 90, 163–171. doi: 10.1139/z11-123

Sundaram, A. Y. M., Garseth, Å. H., Maccari, G., and Grimholt, U. (2020). An Illumina approach to MHC typing of Atlantic salmon. *Immunogenetics* 72, 89–100. doi: 10.1007/s00251-019-01143-8

Takeshima, S.-N., and Aida, Y. (2006). Structure, function and disease susceptibility of the bovine major histocompatibility complex. *Anim. Sci. J.* 77, 138–150. doi: 10.1111/j.1740-0929.2006.00332.x

Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Mol. Biol. Evol.* 9, 678–687. doi: 10.1093/oxfordjournals.molbev.a040752

Taylor, S. S., Jenkins, D. A., and Arcese, P. (2012). Loss of MHC and neutral variation in Peary caribou: genetic drift is not mitigated by balancing selection or exacerbated by MHC allele distributions. *PLoS One* 7:e36748. doi: 10.1371/journal.pone.0036748

Teixeira, J. C., and Huber, C. D. (2021). The inflated significance of neutral genetic diversity in conservation genetics. *Proc. Natl. Acad. Sci.* 118:e2015096118. doi: 10.1073/pnas.2015096118

Traherne, J. A. (2008). Human MHC architecture and evolution: implications for disease association studies. *Int. J. Immunogenet.* 35, 179–192. doi: 10.1111/j.1744-313X.2008.00765.x

Tryland, M., Nymo, I. H., Sánchez Romano, J., Mørk, T., Klein, J., and Rockström, U. (2019). Infectious disease outbreak associated with supplementary feeding of semi-domesticated reindeer. *Front. Veterinary Science* 6:126. doi: 10.3389/fvets.2019.00126

Tryland, M., Stubsjøen, S. M., Ågren, E., Johansen, B., and Kielland, C. (2016). Herding conditions related to infectious keratoconjunctivitis in semi-domesticated reindeer: a questionnaire-based survey among reindeer herders. *Acta Vet. Scand.* 58:22. doi: 10.1186/s13028-016-0203-x

UCN. (2021). *The IUCN red list of threatened species. Version 2021-3* [online]. Available: https://www.iucnredlist.org [Accessed March 7, 2022].

Vallejo, R. L., Bacon, L. D., Liu, H. C., Witter, R. L., Groenen, M. A., Hillel, J., et al. (1998). Genetic mapping of quantitative trait loci affecting susceptibility to Marek's disease virus induced tumors in F2 intercross chickens. *Genetics* 148, 349–360. doi: 10.1093/genetics/148.1.349

van der Knaap, W. O. (1989). Past vegetation and reindeer on Edgeoya (Spitsbergen) between c. 7900 and c. 3800 BP, studied by means of peat layers and reindeer Faecal pellets. *J. Biogeogr.* 16, 379–394. doi: 10.2307/2845229

Vincent, R., Louis, P., Gongora, C., Papa, I., Clot, J., and Eliaou, J. F. (1996). Quantitative analysis of the expression of the HLA-DRB genes at the transcriptional level by competitive polymerase chain reaction. *J. Immunol.* 156, 603–610.

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033

Wilson, D. E., and Reeder, D. M. (2005). *Mammal species of the world: A taxonomic and geographic reference*. Baltimore, Md [London]: Johns Hopkins University Press.

Yamamoto, F., Suzuki, S., Mizutani, A., Shigenari, A., Ito, S., Kametani, Y., et al. (2020). Capturing differential allele-level expression and genotypes of all classical HLA loci and haplotypes by a new capture RNA-Seq method. *Front. Immunol.* 11:941. doi: 10.3389/fimmu.2020.00941

Yannic, G., Pellissier, L., Ortego, J., Lecomte, N., Couturier, S., Cuyler, C., et al. (2014). Genetic diversity in caribou linked to past and future climate change. *Nat. Clim. Chang.* 4, 132–137. doi: 10.1038/nclimate2074