

Communication

# MHC Genotyping by SSCP and Amplicon-Based NGS Approach in Chamois

Sunčica Stipoljev <sup>1</sup>, Elena Bužan <sup>2,3,\*</sup>, Barbora Rolečková <sup>4</sup>, Laura Iacolina <sup>1,2</sup> and Nikica Šprem <sup>1</sup>

<sup>1</sup> Faculty of Agriculture, University of Zagreb, 10000 Zagreb, Croatia; sstipoljev@agr.hr (S.S.); lauraiacolina@gmail.com (L.I.); nsprem@agr.hr (N.Š.)

<sup>2</sup> Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, 6000 Koper, Slovenia

<sup>3</sup> Environmental Protection College, 3320 Velenje, Slovenia

<sup>4</sup> Institute of Vertebrate Biology of the Czech Academy of Sciences, 60365 Brno, Czech Republic; roleckova@ivb.cz

\* Correspondence: elena.buzan@upr.si; Tel.: +386-5-611-75-70; Fax: +386-5-611-75-71

Received: 10 August 2020; Accepted: 16 September 2020; Published: 18 September 2020

**Simple Summary:** The major histocompatibility complex is a family of genes of central importance in vertebrate adaptive immunity. Genes of the major histocompatibility complex are among the most polymorphic genes ever described in vertebrates, with their number differing greatly among species. Due to their functional significance and exceptional diversity, they represent excellent markers in evolutionary ecology and conservation. Traditional methods were commonly used for their genotyping, but the introduction of next-generation sequencing facilitated more accurate and reproducible genotyping of such polymorphic gene families. Nevertheless, due to the high polymorphism of the major histocompatibility complex genes, genotyping is inherently difficult. Here, we compared the performance of traditional and next-generation sequencing on genotyping of the major histocompatibility complex genes in chamois. Although the major histocompatibility complex system in chamois is quite simple, we found genotyping discrepancies between the two methods in 25% of individuals. Our results show that next-generation sequencing has a higher detection capacity and thus allows for more accurate genotyping of highly polymorphic genes.

**Abstract:** Genes of the major histocompatibility complex (MHC) code for cell surface proteins essential for adaptive immunity. They show the most outstanding genetic diversity in vertebrates, which has been connected with various fitness traits and thus with the long-term persistence of populations. In this study, polymorphism of the MHC class II *DRB* locus was investigated in chamois with Single-Strand Conformation Polymorphism (SSCP)/Sanger genotyping and Ion Torrent S5 next-generation sequencing (NGS). From eight identified *DRB* variants in 28 individuals, five had already been described, and three were new, undescribed alleles. With conventional SSCP/Sanger sequencing, we were able to detect seven alleles, all of which were also detected with NGS. We found inconsistencies in the individual genotypes between the two methods, which were mainly caused by allelic dropout in the SSCP/Sanger method. Six out of 28 individuals were falsely classified as homozygous with SSCP/Sanger analysis. Overall, 25% of the individuals were identified as genotyping discrepancies between the two methods. Our results show that NGS technologies are better performing in sequencing highly variable regions such as the MHC, and they also have a higher detection capacity, thus allowing a more accurate description of the genetic composition, which is crucial for evolutionary and population genetic studies.

**Keywords:** major histocompatibility complex; next-generation sequencing; Ion Torrent; *Rupicapra rupicapra*

## 1. Introduction

In the conservation and management of species, genetic monitoring is essential to ensure that appropriate measures are taken to maintain the viability and adaptive potential of the population, particularly in the case of small and declining populations [1]. Although presumably neutral markers are often used to quantify genetic diversity, neutral variation is not enough to understand all mechanisms that shape the genetic variation of populations [2]. Studying the molecular polymorphism of adaptive genes (i.e., genes that directly influence fitness) can help to understand how adaptive genetic variation is generated and maintained within populations [3].

Among the most polymorphic genes in vertebrates are the major histocompatibility complex (MHC) genes, which are responsible for the adaptive immune response [4]. Polymorphism of the MHC genes has been shown to be related to the individual's fitness and thus to the long-term persistence of populations, which makes these genes important markers for various fitness traits, including factors important for population viability, such as resistance to parasites, survival, and reproductive success [4]. The most pronounced polymorphism of the MHC molecules occurs in the amino acid residues that code for the antigen binding groove [5]. As a result of the function of MHC molecules, pathogen-driven balancing selection has generally been assumed as a major evolutionary force for maintaining MHC polymorphism [6], but it is also produced by gene duplication, resulting in extensive copy number variation [7]. Thus, the genetic diversity of MHC is not only characterized by extreme allelic polymorphism and high nucleotide diversity, but also by the number of duplicated genes.

The introduction of high-throughput sequencing (HTS) technologies, also known as next-generation sequencing (NGS), has allowed for a large-scale assessment of genetic variation at reasonable times and costs [8]. In addition, NGS methods improved our ability to genotype highly polymorphic multigene families such as the MHC [9–11].

Great effort has been devoted to MHC genotyping in evolutionary and population studies [12,13]. However, the task is still quite demanding and challenging due to the complex genomic organization and the high sequence variation of MHC loci, as well as the difficulties to separate true alleles from artefacts [14,15]. MHC genes, often in multiple copies, vary widely between and within species, making the identification of all alleles carried by an individual and the reconstruction of its multilocus genotype very challenging [16]. Consequently, the problems that cause significant difficulties in MHC genotyping are (i) the frequent gene duplications and variation between haplotypes in the number of loci within and between species, (ii) the difficult design of locus-specific primers, (iii) varying degrees of concerted evolution, and (iv) the presence of pseudogenes, which cause additional difficulties in identifying functional variants.

For MHC genotyping, different techniques such as Restriction Fragment Length Polymorphism, Single-Strand Conformation Polymorphism (SSCP), Denaturing Gradient Gel Electrophoresis, and Reference Strand-mediated Conformational Analysis in combination with cloning were used [17].

The SSCP method is based on the observation that single-stranded DNA fragments with different DNA sequences will assume sequence-specific conformation when electrophoresed under non-denaturing conditions [18]. Since a change in a single base is sufficient to cause changes in the tertiary structure of single-stranded DNA fragments, SSCP is able to detect single-base substitutions. The SSCP method has been proven robust, with a high sensitivity to detect DNA sequence variation for MHC genes [19,20], but it can encounter problems when the amplification of some alleles is less efficient or when certain alleles are difficult to distinguish [17]. It especially becomes problematic and unreliable if multiple co-amplifying copies are present in the sample [17].

Despite significant advances in HTS technologies, the genotyping of MHC systems and the ability to discriminate between true alleles and artefacts is more challenging as the number of co-amplifying genes increases with this method. Of the available HTS platforms, Ion Torrent and Illumina are among the most appropriate choices for MHC genotyping due to the ultra-high coverage and read lengths, offering the potential to overcome this limitation [16].

Here, we compared the performance of two methods, SSCP and Ion Torrent S5 sequencing, for genotyping highly polymorphic exon 2 of the MHC class II *DRB* gene in chamois, which encodes

functionally important residues of the antigen-binding groove and therefore can be taken as a measure of functional diversity of *DRB* alleles [21].

## 2. Materials and Methods

DNA from 28 chamois muscle samples was extracted with peqGOLD Tissue DNA Mini Kit (VWR International, Leuven, Belgium) according to the manufacturer's protocol. All animals were legally harvested according to the Slovenian Hunting Law (Official Gazette of the Republic of Slovenia Ur. l. 16/04) and Croatian Hunting Law (Official Gazette of the Republic of Croatia 99/18) regular hunting allocations. Handling of animals and sampling were done according to the Ethical and Welfare Standards presented in the (Official Gazette of the Republic of Croatia 102/2017), Regulation on the Protection of Animals Used for Scientific Purposes (Official Gazette of the Republic of Croatia 55/13) and the Bioethical Committee for the Protection and Welfare of Animals of the University of Zagreb Faculty of Agriculture.

### 2.1. MHC Genotyping by SSCP/Sanger Sequencing

The Qiagen multiplex PCR kit was used to amplify the second exon of MHC class II *DRB* gene using primers HL030: 5'-ATCCTCTCTCTGCAGCACATTTCC-3' and HL032: 5'-TCGCCGCTGCACAGTGAACTCTC-3' [22]. PCR was performed following the protocol described elsewhere [22,23] and verified by electrophoresis on ethidium bromide stained agarose gel. Successful PCR products were cloned according to the protocol by Bryja et al. [24] and, following an additional PCR step [22], they were evaluated by single-stranded conformational polymorphism analysis and sequenced by capillary electrophoresis in an ABI 3130 analyzer. Sequencing and allele identification were performed according to the protocol described in Čížková et al. [23]. The results were validated with GeneMapper v.5.0 (Thermo Fisher Scientific, Waltham, MA, USA) software. The total length of the 236 bp allele sequences was assembled using CodonCode Aligner software (version 1.6.3; CodonCode Corporation, Dedham, MA, USA) and aligned with ClustalW 4.0, implemented in MEGA 7 [25].

### 2.2. MHC Genotyping by An Amplicon-Based NGS Approach

Amplicon sequencing was performed in three runs using an Ion Torrent S5 system (Thermo Fisher Scientific). Amplification was carried out by using the same two primers HL030 and HL032. In order to identify individuals, forward primers were designed to include the following motifs: (i) Ion Torrent A adapter, (ii) unique IonXpress barcode, (iii) barcode "GAT" linker, and (iv) forward primer HL030. PCR amplification was performed in triplicates in 25 µL reaction mixtures containing a 5 µL DNA template, 5X reaction buffer, 5 µL Q solution (Qiagen, Hilden, Germany), 3 µM MgCl<sub>2</sub>, 0.25 mM dNTPs, 0.15 µM of each primer, and 0.08 U HotStarTaq (Qiagen). The PCR program included an initial denaturation step at 95 °C for 2 min, 40 denaturation cycles at 95 °C for 20 s, primer annealing at 60 °C for 30 s, and primer extension at 72 °C for 45 s, and a final elongation step at 72 °C for 10 min. PCR products from triplicates were pooled and purified with Agencourt AMPure XP beads (Agencourt Bioscience Corporation, Beverly, MA, USA). The concentrations of the pooled and purified amplicons were estimated with the Qubit 3.0 Fluorometer using the Qubit dsDNA High Sensitivity Assay Kit (Thermo Fisher Scientific). Then, amplicons were normalized to 5 ng, pooled and purified again with Agencourt AMPure XP beads. Size and quality of the pooled amplicons were verified using the Agilent DNA High Sensitivity Kit on the 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA). The final library was normalized to 100 pM and sequenced with the Ion Torrent S5 on a 314 chip (Thermo Fisher Scientific).

Allele calling was conducted through the pipeline based on the Amplicon Sequence Assignment (AmpliSAS) web tool designed for high-throughput genotyping of duplicated polymorphic gene families, such as the MHC [26]. Initial quality and length filtering of raw data was performed with AmpliCLEAN by removing reads with a Phred quality score below 30 and all reads shorter than 279 bp and longer than 289 bp. AmpliSAS clusters true variants with their potential artefacts based on the platform-specific error rates. We used AmpliSAS default parameters for Ion Torrent sequencing

technology: a substitution error rate of 0.5% and an indel error rate of 1%. Exact length was required for the dominant sequence within a cluster. From previous work on this species [22,27–29], we did not expect more than two *DRB* variants per individual, so we kept the "minimum dominant frequency" clustering threshold at default 25%. Variants with frequency lower than 1% within an amplicon were discarded.

The maximum number of reads per amplicon that can be processed by AmpliSAS is 5000; hence, the amplicons with more than 5000 reads were randomly sub-sampled.

The datasets generated and analyzed during the current study were deposited in the NCBI Sequence Read Archive (SRA, accession number PRJNA662939).

True *DRB* variants of the second exon were aligned and translated into protein sequences to check whether there was evidence for pseudogenes, such as the presence of premature stop codons. Three undescribed alleles discovered in this study (two of them validated by both methods) were named *Ruru-DRB\*41*, *Ruru-DRB\*42*, and *Ruru-DRB\*43* and deposited in GenBank with accession numbers MT813042–MT813044.

The evolutionary relationships between the alleles were analyzed by a median-joining network, as implemented in the software PopART [30]. The parameter  $\epsilon$  was set to zero (default) to obtain a sparse spanning network. We compared frequencies of alleles obtained with the two genotyping approaches and estimated discrepancies in each individual's genotypes between the two methods.

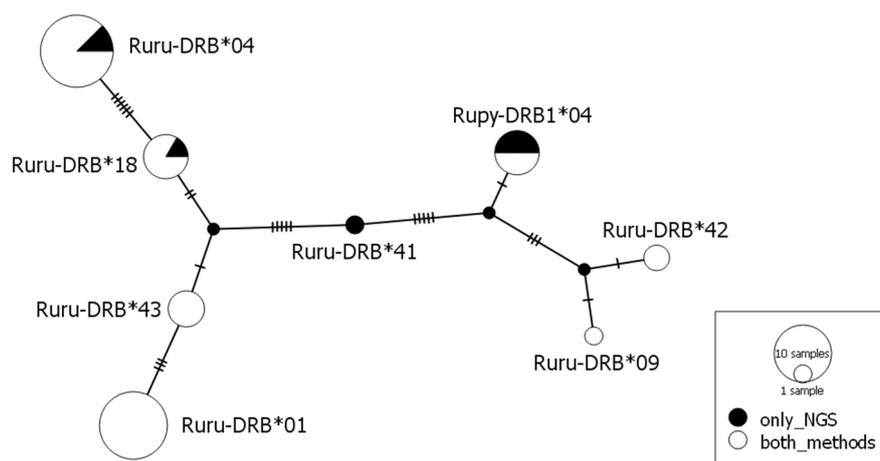
### 3. Results

The sequence of 236 bp (without primers) of the *DRB* exon 2 gene was obtained for all 28 chamois. After initial filtering of the raw Ion Torrent data, the amplicon coverage ranged from 656 to the maximum of 5000 reads allowed by AmpliSAS, with an average of  $3191 \pm 1996$  (SD) reads. The coverage in our study was sufficient to separate true alleles from artefacts and obtain reliable genotypes.

When clusters within an amplicon were ordered by descending per amplicon frequency (PAF), we observed a significant drop in frequency at the value of 3%, which probably represents the boundary between true alleles and artefacts. For one individual, we found chimeric variant at a frequency of 13%. Accordingly, we set the PAF threshold at 14% and removed the chimeric and low-frequency variants below this threshold (Figure S1 and Table S1).

Clustering sequencing errors with true variants increased the read depths of true variants and thereby allowed us to more easily distinguish alleles from low-frequency artefacts. On average, 54% of amplicon depths were assigned to alleles, and after clustering, this proportion increased to 83% (Figure S2). Ion Torrent tends to produce high rates of homopolymer indels, depending on the properties of the analyzed sequences. However, this had very little effect on the results of genotyping of chamois. The estimates of per amplicon variant frequencies after clustering were mainly affected by 1 bp substitution (Figure S3), which accounted, on average, for 21% of the coverage increase of true variants within the amplicon (Table S2).

No stop codons were found in the nucleotide sequence of any allele identified by SSCP/Sanger sequencing and amplicon-based NGS, indicating that all alleles encoded for functional proteins. The allele identities were based on 17 (7%) variable nucleotide positions with the lowest number of mutations (two) between the alleles *Ruru-DRB\*09* and *\*42*. The maximum number of 14 mutations was found between allele *Ruru-DRB\*01* and *\*42* (Figures 1 and 2).



**Figure 1.** Median-joining network ( $\epsilon = 0$ ) of *DRB* alleles found in chamois. Alleles are represented by pie charts whose size is proportional to the number of individuals. The colors indicate the number of individuals that have particular alleles obtained with two genotyping methods. Number of mutations separating nodes is represented by slashes crossed with the network branches. Small black circles indicate hypothetical alleles, as predicted by the model.

	9	10	54	72	88	89	91	99	118	145	147	148	156	187	190	211	235
<i>Ruru-DRB*01</i>	C	A	T	C	T	G	T	G	A	C	A	C	C	G	G	T	G
<i>Ruru-DRB*04</i>	A	G	.	.	.	.	.	.	T	.	G	A	T	.	A	A	T
<i>Ruru-DRB*09</i>	.	G	C	T	A	C	C	A	.	.	G	A	T	A	.	A	.
<i>Ruru-DRB*18</i>	A	G	.	.	.	.	.	.	.	.	.	.	.	A	C	A	T
<i>Ruru-DRB*41</i>	.	.	.	T	A	C	C	A	.	.	.	.	.	A	C	A	T
<i>Ruru-DRB*42</i>	A	G	C	T	A	C	C	A	.	.	G	A	T	A	.	A	T
<i>Ruru-DRB*43</i>	.	.	.	.	.	.	.	.	.	.	.	.	.	A	C	A	.
<i>Rupy-DRB1*04</i>	A	G	.	T	A	C	C	A	.	G	G	A	.	A	C	A	.

**Figure 2.** Alignment of the nucleotide sequences, only variable residues of the exon 2 of *DRB* alleles of chamois are shown. The codes before \* represent the species abbreviation and gene name, and the numbers after \* indicate allele numbers.

A total of eight alleles were identified in 28 individuals. Five of them had already been described [31,32], and three were new, not previously described alleles. We discovered seven alleles by SSCP/Sanger genotyping, and all of them were also detected by NGS sequencing. Further, NGS enabled detecting one more allele and extended the genotypes of six of the 28 individuals compared to SSCP/Sanger genotyping results (Table S3).

The maximum number of alleles per individual detected by both methods was two, which was consistent with the expectation of a single *DRB* locus from previous studies. A considerable variation in frequency between alleles was observed. For both the methods, the lowest frequencies were for *Ruru-DRB\*41* (0% and 2%, respectively) and *Ruru-DRB\*09* (2%), and the highest were for *Ruru-DRB\*01* (36% and 32%, respectively) (Table 1).

**Table 1.** *DRB* exon 2 allele frequencies estimated with Single-Strand Conformation Polymorphism (SSCP)/Sanger sequencing and amplicon-based next-generation sequencing (NGS).

Allele	SSCP/Sanger		NGS	
	N° Observations	Allele Frequency	N° Observations	Allele Frequency
<i>Ruru-DRB*01</i>	20	0.357	18	0.321
<i>Ruru-DRB*04</i>	18	0.321	17	0.304
<i>Ruru-DRB*09</i>	1	0.018	1	0.018
<i>Ruru-DRB*18</i>	5	0.089	6	0.107
<i>Ruru-DRB*41</i>	0	0.000	1	0.018

<i>Ruru-DRB*42</i>	2	0.036	2	0.036
<i>Ruru-DRB*43</i>	6	0.107	4	0.071
<i>Rupy-DRB1*04</i>	4	0.071	7	0.125

The number of called alleles per individual was consistent between the two genotyping methods in 75% of the cases. Discrepancies in the individual genotypes between the two methods were mainly caused by allelic dropout using the SSCP/Sanger analysis. Six out of 28 individuals (21%) were falsely classified as homozygous with SSCP/Sanger analysis due to dropout events. In addition to the allelic dropout, we had a false positive allele call for one individual with SSCP/Sanger analysis, which was due to a weak obtained sequence. In total, 25% of the individuals were identified as genotyping discrepancies. A summary of the *DRB* alleles found in each individual using both genotyping approaches is given in Table S3.

#### 4. Discussion

Our results show that NGS technologies are better performing in the sequencing of highly variable regions such as the MHC and could offer real advantages compared to the SSCP/Sanger methodology. However, knowledge of the complexity of the MHC system in the study species and, accordingly, the read depth that would be satisfactory to obtain such genotypes (in our case at least 600 reads per amplicon) is really important to exclude sequencing errors [15].

Recently, there have been many publications pointing to errors that can occur during the NGS sequencing process [14,16]. It is important to recognize that long-term limitations of genotyping by sequencing can also occur because, similar to traditional SSCP/Sanger genotyping, approaches of NGS sequencing are still based on amplification of the target DNA sequence by PCR. The variety of artefacts that occur during the PCR step can far exceed the number of variants corresponding to true alleles, and the total number of sequence reads derived from artefacts can be higher than those from true alleles [33]. PCR may also be responsible for the allelic dropout caused by differences in amplification efficiency between alleles, especially in the SSCP/Sanger methodology [34]. Allelic dropouts and false alleles can cause important genotyping errors and thus significantly alter downstream analyses based on heterozygosity, allele diversity, and genotype composition, which has a direct implication in population genetics.

The stochastic variation of allele frequencies will inevitably worsen in NGS methodologies if the number of reads is reduced. Ion Torrent sequencing is currently one of the methods of choice for MHC genotyping but, for accurate genotyping, it is important to cover a sufficient number of reads per amplicon and to use strict approaches for the quality control of generated data and allele calling [12]. Using Illumina sequencing, Biedrzycka et al. [15] tested four genotyping approaches for characterizing MHC diversity in a passerine bird with the maximum number of co-amplifying alleles exceeding 40 per individual. The four different genotyping methods tested in their study were developed based on different principles: (i) set a threshold based on the observed frequencies of variants that could be explained as a sequencing error [35]; (ii) use two amplicon replicates to separate true alleles from artefacts [33]; (iii) set a threshold between true alleles and artefacts based on the drop in sequencing depth between them [11]; and (iv) use the error rate of a particular HTS platform to cluster true alleles with their potential artefacts [26]. Despite these differences, the authors found generally high agreement across genotyping methods and demonstrated that HTS enables reliable genotyping of highly complex MHC systems if sufficient coverage per amplicon is reached (in their case, at least 5000 reads per amplicon).

NGS technologies are likely to replace cloning and Sanger sequencing for MHC genotyping. However, traditional genotyping methods may remain an alternative that should be considered for genotyping a limited number of samples, as NGS is still expensive, and many research groups do not have direct access to these technologies. Our study has shown that the results of the Ion Torrent and traditional cloning/Sanger sequencing are comparable on a qualitative level, but with NGS, we were able to detect a higher number of "putative alleles".

We must emphasize that this discrepancy between NGS and traditional methods is probably species-specific and depends on the complexity of the MHC II *DRB* system under investigation. The chamois MHC II *DRB* system is quite simple with only one functional copy of the gene, but the difference between the methods can be much bigger in more complex systems [15].

In our opinion, the higher probability of allele detection with NGS is based on the higher sequencing depth, which can compensate for the limitations of SSCP/Sanger sequencing, since the difficult amplification of some alleles i.e., differences in allele's amplification efficiency, is a real limitation of this traditional method.

The correct assessment of an individual's MHC constitution is the most fundamental prerequisite for understanding the functional significance of MHC allele diversity in evolutionary ecology, pathogen resistance, and conservation. The comparison of the methodological approaches is necessary to enable researchers to evaluate in detail and understand the data generated by NGS, which will improve confidence in the approach and in subsequent analyses and applications.

## 5. Conclusions

The major histocompatibility complex is polymorphic and polygenic in most species and therefore inherently difficult to genotype. We have shown that NGS technologies are better performing in sequencing highly variable regions such as the MHC compared to the SSCP/Sanger methodology, but strict pipelines for allele calling are required to distinguish true alleles from artefacts. Since SSCP/Sanger genotyping may underestimate the actual MHC variability, this method will most likely have very limited applicability. However, SSCP/Sanger analysis may remain an alternative for genotyping a limited number of samples and may be applicable to species with a less complex MHC system, such as chamois.

**Supplementary Materials:** The following are available online at [www.mdpi.com/2076-2615/10/9/1694/s1](http://www.mdpi.com/2076-2615/10/9/1694/s1), Figure S1: Variants ordered by descending per amplicon frequency after AmpliSAS clustering, Figure S2: Percentages of amplicon reads assigned to alleles before and after AmpliSAS clustering, Figure S3: Representation of amplicon depths from 28 individuals, and the number of reads of true sequences and artefacts within each amplicon, Table S1: AmpliSAS genotyping results of 28 chamois individuals, Table S2: Amplicon depths from 28 individuals, and the number of reads of true sequences and artefacts within each amplicon, Table S3: Genotyping results of 28 chamois individuals obtained by SSCP/Sanger sequencing and amplicon-based NGS.

**Author Contributions:** Conceptualization, E.B. and N.Š.; formal analysis, S.S., E.B., B.R., L.I.; writing—original draft preparation, S.S. and E.B.; writing—review and editing, S.S., E.B., B.R., L.I. and N.Š. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by (i) the Croatian Science Foundation, project IP 2016-06-5751, “DNA as evidence of distribution and vitality of endangered Balkan chamois”; (ii) the Slovenian Research Agency (programme group P6-0064); (iii) the STARBIOS2 European Union's Horizon 2020 Research and Innovation Program under grant agreement (No. 709517); (iv) the RESBIOS European Union's Horizon 2020 Research and Innovation Program (No. 872146); (v) COST action G-Bike Genomic Biodiversity Knowledge for Resilient Ecosystems and (vi) the Open Access Publication Fund of the University of Zagreb Faculty of Agriculture.

**Acknowledgments:** We would like to thank Sandra Potušek for her help with the laboratory work. We thank Ana Galov and Željko Pavlinec for advice and thoughtful comments on data analysis. We are grateful to two anonymous reviewers for their valuable comments that greatly improved the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Leroy, G.; Carroll, E.L.; Bruford, M.W.; DeWoody, J.A.; Strand, A.; Waits, L.; Wang, J. Next-generation metrics for monitoring genetic erosion within populations of conservation concern. *Evol. Appl.* **2018**, *11*, 1066–1083, doi:10.1111/eva.12564.
2. Eizaguirre, C.; Baltazar-Soares, M. Evolutionary conservation-evaluating the adaptive potential of species. *Evol. Appl.* **2014**, *7*, 963–967, doi:10.1111/eva.12227.

3. Funk, W.C.; McKay, J.K.; Hohenlohe, P.A.; Allendorf, F.W. Harnessing genomics for delineating conservation units. *Trends Ecol. Evol.* **2012**, *27*, 489–496, doi:10.1016/j.tree.2012.05.012.
4. Sommer, S. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front. Zool.* **2005**, *2*, 16, doi:10.1186/1742-9994-2-16.
5. Moreno-Santillán, D.D.; Lacey, E.A.; Gendron, D.; Ortega, J. Genetic variation at exon 2 of the MHC class II DQB locus in blue whale (*Balaenoptera musculus*) from the Gulf of California. *PLoS One* **2016**, *11*, 1–15, doi:10.1371/journal.pone.0141296.
6. Spurgin, L.G.; Richardson, D.S. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc. R. Soc. B Biol. Sci.* **2010**, *277*, 979–988, doi:10.1098/rspb.2009.2084.
7. O'Connor, E.A.; Strandh, M.; Hasselquist, D.; Nilsson, J.Å.; Westerdahl, H. The evolution of highly variable immunity genes across a passerine bird radiation. *Mol. Ecol.* **2016**, *25*, 977–989, doi:10.1111/mec.13530.
8. Reuter, J.A.; Spacek, D. V.; Snyder, M.P. High-Throughput Sequencing Technologies. *Mol. Cell* **2015**, *58*, 586–597, doi:10.1016/j.molcel.2015.05.004.
9. Babik, W.; Taberlet, P.; Ejsmond, M.J.; Radwan, J. New generation sequencers as a tool for genotyping of highly polymorphic multilocus MHC system. *Mol. Ecol. Resour.* **2009**, *9*, 713–719, doi:10.1111/j.1755-0998.2009.02622.x.
10. Zagalska-Neubauer, M.; Babik, W.; Stuglik, M.; Gustafsson, L.; Cichoń, M.; Radwan, J. 454 sequencing reveals extreme complexity of the class II Major Histocompatibility Complex in the collared flycatcher. *BMC Evol. Biol.* **2010**, *10*, 395, doi:10.1186/1471-2148-10-395.
11. Lighten, J.; van Oosterhout, C.; Paterson, I.G.; McMullan, M.; Bentzen, P. Ultra-deep Illumina sequencing accurately identifies MHC class IIb alleles and provides evidence for copy number variation in the guppy (*Poecilia reticulata*). *Mol. Ecol. Resour.* **2014**, *14*, 753–767, doi:10.1111/1755-0998.12225.
12. Grogan, K.E.; McGinnis, G.J.; Sauther, M.L.; Cuozzo, F.P.; Drea, C.M. Next-generation genotyping of hypervariable loci in many individuals of a non-model species: Technical and theoretical implications. *BMC Genomics* **2016**, *17*, 204, doi:10.1186/s12864-016-2503-y.
13. Huang, K.; Zhang, P.; Dunn, D.W.; Wang, T.; Mi, R.; Li, B. Assigning alleles to different loci in amplifications of duplicated loci. *Mol. Ecol. Resour.* **2019**, *19*, 1240–1253, doi:10.1111/1755-0998.13036.
14. Lighten, J.; van Oosterhout, C.; Bentzen, P. Critical review of NGS analyses for de novo genotyping multigene families. *Mol. Ecol.* **2014**, *23*, 3957–3972, doi:10.1111/mec.12843.
15. Biedrzycka, A.; Sebastian, A.; Migalska, M.; Westerdahl, H.; Radwan, J. Testing genotyping strategies for ultra-deep sequencing of a co-amplifying gene family: MHC class I in a passerine bird. *Mol. Ecol. Resour.* **2017**, *17*, 642–655, doi:10.1111/1755-0998.12612.
16. Rekdal, S.L.; Anmarkrud, J.A.; Johnsen, A.; Lifjeld, J.T. Genotyping strategy matters when analyzing hypervariable major histocompatibility complex-Experience from a passerine bird. *Ecol. Evol.* **2018**, *8*, 1680–1692, doi:10.1002/ece3.3757.
17. Babik, W. Methods for MHC genotyping in non-model vertebrates. *Mol. Ecol. Resour.* **2010**, *10*, 237–251, doi:10.1111/j.1755-0998.2009.02788.x.
18. Sunnucks, P.; Wilson, A.C.C.; Beheregaray, L.B.; Zenger, K.; French, J.; Taylor, A.C. SSCP is not so difficult: The application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Mol. Ecol.* **2000**, *9*, 1699–1710, doi:10.1046/j.1365-294X.2000.01084.x.
19. Garrigan, D.; Hedrick, P.W. Class I MHC polymorphism and evolution in endangered California Chinook and other Pacific salmon. *Immunogenetics* **2001**, *53*, 483–489, doi:10.1007/s002510100352.
20. Noakes, M.A.; Reimer, T.; Phillips, R.B. Genotypic characterization of an MHC class II locus in lake trout (*Salvelinus namaycush*) from Lake Superior by single-stranded conformational polymorphism analysis and reference strand-mediated conformational analysis. *Mar. Biotechnol.* **2003**, *5*, 270–278, doi:10.1007/s10126-002-0079-9.
21. Hughes, A.L.; Yeager, M. Natural selection at major histocompatibility complex loci of vertebrates. *Annu. Rev. Genet.* **1998**, *32*, 415–435, doi:10.1146/annurev.genet.32.1.415.
22. Schaschl, H.; Goodman, S.J.; Suchentrunk, F. Sequence analysis of the MHC class II DRB alleles in Alpine chamois (*Rupicapra r. rupicapra*). *Dev. Comp. Immunol.* **2004**, *28*, 265–277, doi:10.1016/j.dci.2003.08.003.
23. Číková, D.; de Bellocq, J.G.; Baird, S.J.E.; Piálek, J.; Bryja, J. Genetic structure and contrasting selection pattern at two major histocompatibility complex genes in wild house mouse populations. *Heredity (Edinb)*. **2011**, *106*, 727–740, doi:10.1038/hdy.2010.112.

24. Bryja, J.; Galan, M.; Charbonnel, N.; Cosson, J.F. Analysis of major histocompatibility complex class II gene in water voles using capillary electrophoresis-single stranded conformation polymorphism. *Mol. Ecol. Notes* **2005**, *5*, 173–176, doi:10.1111/j.1471-8286.2004.00855.x.
25. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874, doi:10.1093/molbev/msw054.
26. Sebastian, A.; Herdegen, M.; Migalska, M.; Radwan, J. Amplis: A web server for multilocus genotyping using next-generation amplicon sequencing data. *Mol. Ecol. Resour.* **2016**, *16*, 498–510, doi:10.1111/1755-0998.12453.
27. Fuselli, S.; Baptista, R.P.; Panziera, A.; Magi, A.; Guglielmi, S.; Tonin, R.; Benazzo, A.; Bauzer, L.G.; Mazzoni, C.J.; Bertorelle, G. A new hybrid approach for MHC genotyping: high-throughput NGS and long read MinION nanopore sequencing, with application to the non-model vertebrate Alpine chamois (*Rupicapra rupicapra*). *Heredity (Edinb)*. **2018**, *121*, 293–303, doi:10.1038/s41437-018-0070-5.
28. Zemanová, B.; Hájková, P.; Hájek, B.; Martínková, N.; Mikulíček, P.; Zima, J.; Bryja, J. Extremely low genetic variation in endangered Tatra chamois and evidence for hybridization with an introduced Alpine population. *Conserv. Genet.* **2015**, *16*, 729–741, doi:10.1007/s10592-015-0696-2.
29. Mona, S.; Crestanello, B.; Bankhead-Dronnet, S.; Pecchioli, E.; Ingrosso, S.; D’Amelio, S.; Rossi, L.; Meneguz, P.G.; Bertorelle, G. Disentangling the effects of recombination, selection, and demography on the genetic variation at a major histocompatibility complex class II gene in the alpine chamois. *Mol. Ecol.* **2008**, *17*, 4053–4067, doi:10.1111/j.1365-294X.2008.03892.x.
30. Leigh, J.W.; Bryant, D. POPART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **2015**, *6*, 1110–1116, doi:10.1111/2041-210X.12410.
31. Schaschl, H.; Wandeler, P.; Suchentrunk, F.; Obexer-Ruff, G.; Goodman, S.J. Selection and recombination drive the evolution of MHC class II *DRB* diversity in ungulates. *Heredity (Edinb)*. **2006**, *97*, 427–437, doi:10.1038/sj.hdy.6800892.
32. Alvarez-Busto, J.; García-Etxebarria, K.; Herrero, J.; Garin, I.; Jugo, B.M. Diversity and evolution of the *Mhc-DRB1* gene in the two endemic Iberian subspecies of Pyrenean chamois, *Rupicapra pyrenaica*. *Heredity (Edinb)*. **2007**, *99*, 406–413, doi:10.1038/sj.hdy.6801016.
33. Sommer, S.; Courtiol, A.; Mazzoni, C.J. MHC genotyping of non-model organisms using next-generation sequencing: A new methodology to deal with artefacts and allelic dropout. *BMC Genomics* **2013**, *14*, 542, doi:10.1186/1471-2164-14-542.
34. Montero, B.K.; Refaly, E.; Ramanamanjato, J.B.; Randriatafika, F.; Rakotondranary, S.J.; Wilhelm, K.; Ganzhorn, J.U.; Sommer, S. Challenges of next-generation sequencing in conservation management: Insights from long-term monitoring of corridor effects on the genetic diversity of mouse lemurs in a fragmented landscape. *Evol. Appl.* **2019**, *12*, 425–442, doi:10.1111/eva.12723.
35. Radwan, J.; Zagalska-Neubauer, M.; Cichoń, M.; Senddecka, J.; Kulma, K.; Gustafsson, L.; Babik, W. MHC diversity, malaria and lifetime reproductive success in collared flycatchers. *Mol. Ecol.* **2012**, *21*, 2469–2479, doi:10.1111/j.1365-294X.2012.05547.x.

