MHCI variability and fitness in blue tits (*Cyanistes caeruleus*)

Mia Erika Rochman Fagerlund Johansen

Master Thesis Ecology and Evolution 60 credits

The Natural History Museum

Department of Biosciences

Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

January 2023

Photo by Bjørn Aksel Bjerke

MHC and fitness in blue tits (*Cyanistes caeruleus*)

Master Thesis Ecology and Evolution

Mia Erika Rochman Fagerlund Johansen

The Natural History Museum

Department of Biosciences

Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

January 2023

Acknowledgements

This master thesis was carried out at SERG (Sex and Evolution Research Group) at Natural history museum under the Department of Biosciences, University of Oslo, under the supervision of my main supervisor Arild Johnsen and co-supervisors Tore Slagsvold, Jarl Andreas Anmarkrud and Silje Larsen Rekdal. There are a few people I would like to thank.

Arild, thank you for your support and guidance during the entirety of my thesis, and for being my teacher in in all aspects of this field. Thank you for always answering the door when I knock, for answering all my questions seriously, and for lending me your time, ear and knowledge. Tore, thank you for being my teacher and guide, and helping me become independent in my fieldwork. Jarl Andreas, thank you for your unconditional support in the lab, for answering all my questions, and for lending me your time and knowledge when I was at a loss. Silje, thank you for your enormous support and uplifting words. You guided me through the toughest parts of my thesis and assured me all would be well. The time you have all allocated to my thesis is invaluable.

I also want to thank Christian Kierulf Aas, who spent many hours with me in the field, guiding me through sampling, and building my confidence to sample on my own.

Thank you to Lars Erik Johannessen, Tore Slagsvold, Bo Terning Hansen and Jan T. Lifjeld who collected samples and participated in the 2001 study. A shout out to everyone in the NHM DNA lab, for helping me whenever I was in need. To Bjørn Aksel Bjerke, thank you for the beautiful picture.

To my boyfriend Jørgen, thank you for keeping me company through the late nights, early mornings, for supporting me through the stressful parts, and for all our discussions about anything biologically related. You have given me unconditional support during the entirety of my thesis.

To my parents, thank you for always believing in me and supporting me academically throughout my life. You have always cheered me on, even when I doubted myself.

Finally, thank you to all close friends and extended family for checking up on me, believing in me and cheering me on. You know who you are.

Abstract

The Major Histocompatibility Complex (MHC) is a group of closely linked loci that play an important role in the immune system within all jawed vertebrates, and it has been shown to yield a higher fitness with an intermediate diversity. Several studies present how females choose their mates based on the MHC variability, hereby providing their offspring with the optimal variability. MHC class I (MHCI) mediates immunity against intracellular pathogens such as *Haemoproteus*, *Plasmodium* and *Leucocytozoon*. Here I investigate the MHCI composition in a population of blue tits (*Cyanistes caeruleus*) sampled in 2001 and 2021, and test whether social pairing departs from what is expected from random pairings. I also investigate the connection between the MHCI morphology, age and reproductive success. Additionally, I investigate the correlation between the MHCI and prevalence of the blood parasites *Haemoproteus*, *Plasmodium* and *Leucocytozoon*, and the correlation between these blood parasites, morphology, age and reproductive success. The younger females tended to be in social pairs with fewer positively selected site (PSS) alleles than expected from random pairings, whereas the older females tended to be in social pairs with more PSS alleles than expected. A higher number of MHCI alleles was linked to longer wings. Interestingly, there was an association between the number of MHCI alleles and *Haemoproteus* infection within the 2001 blue tits, with individuals with a higher number of alleles being more likely to be infected. Finally, a close to significant positive association was also found between the MHCI PSS allele IWYWYVGR and *Haemoproteus* infection. My results are most consistent with selection for a higher or lower number of alleles rather than an intermediate optimum, but additional studies are needed to further investigate these results.

Table of contents

1.0 Introduction

The Major histocompatibility complex (MHC) is a large genetic region of closely linked loci that play an important role in the immune system of all jawed vertebrates (Kaufman, 2018; Klein & Figueroa, 1986). The purpose of the MHC is to present peptide fragments from pathogens to the appropriate T-cells in order to eliminate or neutralize them (Janeway et al., 2001).

There are two classes of antigen presenting MHC genes; MHC class I (hereafter referred to as MHCI) and MHC class II (hereafter referred to as MHCII). MHCI mediates immunity against intracellular pathogens and is expressed on all nucleated cells. It presents antigens derived from pathogens in the cytoplasm to cytotoxic $CD8^+$ T-cells (Janeway et al., 2001; Klein & Figueroa, 1986; Neefjes et al., 2011; Schubert et al., 2021). MHCII mediates immunity against extracellular pathogens such as bacteria or parasites, and it is expressed by the professional antigen-presenting cells (B-cells, macrophages, dendritic cells, etc.) to $CD4^+$ T-cells (Janeway et al., 2001; Klein & Figueroa, 1986; Neefjes et al., 2011; Schubert et al., 2021).

The MHC is polygenic and highly polymorphic. This means that each organism has several loci encoding for proteins with a similar function, as well as having numerous alleles at each locus (Aeschlimann et al., 2003; Janeway et al., 2001), thus making it challenging for pathogens to escape immune responses. Multiple copies of the MHC genes can increase the organisms' resistance to pathogens, though having too many copies will increase the risk of autoimmune diseases (Lenz et al., 2015). In situations where too many MHC alleles are present, there will be too few circulating T-cells due to negative selection in the thymus (Milinski, 2006; Nowak et al., 1992). This means that the T-cells that bind to self-antigens bound to the MHC-molecules too well, will be removed, avoiding an immune response to one's own tissue. An intermediate number of MHC alleles is thus thought to be the better evolutionary approach (Aeschlimann et al., 2003).

Female mate choice can occur before, during and after copulation (Ziegler et al., 2005), where for example cryptic female choice plays an important role. The female can base her choice upon several traits of her potential mate to optimize the genetic constitution of her future offspring. MHC-based mate choice has been observed in a wide range of species (Chaix et al., 2008; Kamiya et al., 2014; Leclaire et al., 2017; Tregenza & Wedell, 2000; Winternitz et al., 2017), although less frequently in birds. For MHCI-based mate choice in birds, Rekdal et al. (2019), Strandh et al. (2011) and Strandh et al. (2012) provide good examples. This type of mate choice can be based on direct benefits, compatible genes or good genes. Direct benefits could be in the form of parental investment from a mate with a good immune system, however there is little evidence to support this hypothesis (Kamiya et al., 2014). According to the compatible genes

hypothesis, in pursuance of a mate to provide the optimal MHC variability for her offspring, the female must be aware of her own genetic composition as well as the genetic composition of her potential mate (Milinski, 2006).

Hamilton and Zuk (1982) suggested that the female is able to evaluate her potential mates' genetic composition based on the expression of secondary sexual traits. More specifically, they suggested that there is a positive correlation between the expression of male secondary sexual characters and immunity against parasites, such that only males with a good immune system were able to produce high-quality ornaments (Hamilton & Zuk, 1982). Since such characters are often costly, they signal that the male is a strong individual with good genes. In other words, the good genes hypothesis states that the males' phenotypical traits are honest indicators of his health and fitness (Mays & Hill, 2004). Thus, there might be a correlation between morphological traits and resistance to parasites. A study performed by Eizaguirre et al. (2009) supports the good genes hypothesis, as they found that there was a correlation between MHC, size, and ability to fight a common parasite (*Gyrodactylus* sp.) in male three-spined sticklebacks (*Gasterosteus aculeatus*). This supports the stated hypothesis that there is a correlation between phenotypic traits and MHC variability. A study performed by Dunn et al. (2012) also supports this theory. They found that female common yellowthroats (*Geothlypis trichas*) preferred extra-pair paternity (EPP) with a larger black facial mask, as this is linked to a greater MHC variation. Additionally, the study by Richardson et al. (2005) found that females were more likely to obtain EPP with a higher MHC diversity, if their social partner had a low variability, which is consistent with the good genes hypothesis. Genetic dissimilarity can also offer a favorable genetic composition, and a fitness benefit, to the offspring in the form of genetic compatibility (Mays & Hill, 2004), thus giving the female indicators as to which male she should choose. The bluethroat (*Luscinia svecica*) in Rekdal et al. (2019) provides a good example of genetic compatibility as their study found how mate choice could reduce the population variance in individual MHC diversity and imply a stabilizing effect on the trait.

MHCI mediates immunity against intracellular pathogens, such as *Plasmodium* (avian malaria parasite), *Leucocytozoon* and *Haemoproteus* (malaria-like parasites) (Valkiūnas & Iezhova, 2017). These parasites are among the most common avian blood parasite genera (Hellgren et al., 2004), and they are recognized as important evolutionary factors in birds (Atkinson et al., 2009). These parasites have been linked to various types of diseases, including Marek's disease in chickens and necrosis in passerine birds (Donovan et al., 2008; Zekarias et al., 2002). Anemia, anorexia, diarrhea, weight loss, depression, respiratory and neurologic signs in turkeys, ducks and chickens, but especially in non-domestic birds, have also been reported

(Wettere, 2022). Additionally, *Haemoproteus* infection has been documented to cause a decrease in host fitness, nestling mortalities, fledging success and a delayed recovery in infected birds (Wettere, 2022).

The blue tit (*Cyanistes caeruleus*) is a small passerine bird that lives in woodlands, parks and gardens. It is a socially monogamous species, with a relatively low frequency of extra-pair paternity (Johannessen et al., 2005; Lifjeld & Krokene, 2000). The MHCI has previously been described in a number of passerine birds, although within the blue tit it has been described as less diverse compared to other passerines. Wutzler et al. (2012) studied the MHCI variation in a natural blue tit population and found 19 MHCI sequences and 10 amplified MHCI loci. Additionally, they identified 50 unique functional sequences.

In the blue tit, blood parasites are known to negatively affect reproductive success and other measurements of fitness (Knowles et al., 2010; Tomás et al., 2007). Puente et al. (2010) found that infection by *Haemoproteus* in wild birds reduced their survival rate, and thus reducing their fitness. However, the study by Podmokła et al. (2014) showed that infected birds increased their reproductive investment resulting in no visible differences in clutch size. Aguilar et al. (2016) tested the relationship between MHCI in the blue tit and the two blood parasites *Haemoproteus* and *Leucocytozoon*. They found that three MHCI alleles (UA104, UA108 and UA117) were associated with the infection rate of *Leucocytozoon.* However, they found no correlation between MHCI and *Haemoproteus*. These results further raise questions concerning the correlation between specific MHCI alleles and blood parasites.

This study aims to address questions concerning MHCI variation, mate choice and the correlation between MHCI and fitness-related characters. I will genotype blue tit MHCI comprehensively using high-throughput sequencing. I will then describe the MHCI variation in a Norwegian blue tit population through the following 4 main questions; (1) is the MHCI composition of social pairs different from what would be expected from random pairings? (2) Are there any associations between MHCI variation, morphological traits, age and reproductive success? (3) Are there any associations between individual MHCI variation and blood parasite prevalence? And finally (4) are there any associations between the blood parasite prevalence and morphological traits, age and reproductive success? There are three distinct ways in which I will analyze the MHCI variation; as the number of alleles, as presence of single alleles and as a distance to the optimum, in an attempt to separate the effect of good genes and compatible genes.

2.0 Materials and methods

2.1 Study area and study species

The blue tits were studied during 2001 and 2021 in mixed deciduous and coniferous woodlands around Dæli near Oslo, Norway (59°56'N, 10°32'E) (Appendix 6 Supplementary Figure 3 and 4). The species is an excellent model organism because it is common, uses nest boxes for breeding so that breeding success can easily be observed, and tolerates catching and handling. The study area is about 1.6 km^2 provided with about 500 wooden nest boxes attached to trees about 1.5 m above the ground and 40-50 meters apart (Slagsvold et al., 2013). Blue tits are small (about 10-12 g), secondary cavity nesting birds and defend territories. Only the female builds the nest, incubate the eggs and brood the young. In our study area, most blue tits are resident with egg laying from around 20th of April to 20th of May. The peak for the start of egg laying is at the end of April in most years. The population of blue tits has been studied since 1995, with annual numbers of nests between 80 and 140 (T. Slagsvold, pers.com.).

2.2 Fieldwork and data collection

The fieldwork of 2001 was conducted for the purpose of another project (Johannessen et al., 2005). The collected samples were conserved in the DNA-bank at the Natural History Museum in Oslo. The fieldwork of 2021 began in middle of March to identify nest box owners and the presence of nest building materials. From about 20th of April, the nest boxes were inspected every 2-3 days to record date of egg laying, final clutch size, hatching date, and the number of nestlings that hatched and fledged.

Mist nets were placed in close proximity to the nest boxes first and foremost in order to capture the male blue tits, though some female blue tits were also captured using this method. We brought a previously caught male in a cage from a different study area to further capture the focal male birds, and played the song of the blue tit on a loud speaker. The birds captured were placed in a small bag in order to calm them down and reduce stress, while preparing for sampling. Firstly, the sex of the bird was determined by the color of its plumage, and by checking for presence of a brood patch which only the female has. The age of the bird was determined according to Svensson (1992) to be either two years old (in its second calendar year), hereafter referred to as SY, or 3+ years old (in its third or more calendar year), hereafter referred to as ASY, by the color of its wing. If the bird had rings from previous capturing, their identification number was written down. If they did not already have rings, they received an aluminum ring with an identification number, and bands with a color combination indicating their sex, age and year of marking. Thereafter they were placed in a small plastic bag and weighed to the nearest 0.5g using a 50g Pesola spring balance. Wing

length (with flattened chord) was measured to the nearest 0.5 mm with a stopped metal ruler at the wing bend (Svensson, 1992). The tarsus was measured with a sliding Vernier caliper to the nearest 0.1 mm. Finally, a blood sample was taken from the branchial vein located on the elbow joint on the wing. A small needle was used to prick the vein so a blood droplet could form. The droplet was collected with a capillary glass tube and deposited in a 2mL tube containing 96% ethanol.

Two days after presumed hatching time (based on date of last egg laid and a length of 13 days for the incubation period), the nestlings were counted. The largest nestling of the brood, which was considered to be the first hatchling, was weighed with a Pesola 10g spring balance in order to calculate the hatching date of the first nestling from a growth curve for nestling blue tits (Gibb, 1950). When the nestlings were approximately 15 days old, they were weighed again with a Pesola 50g spring balance and banded with an aluminum band. In addition, a blood sample was taken from the brachial vein as described above for the parents. For nest boxes where neither of the chicks survived, dead chicks and eggs were samples for tissue to be used in further analysis. The remaining females were caught in their nest boxes by covering their exits and grabbing them by hand. They were also ringed, weighed, measured and blood sampled. A total of 95 birds were sampled from 2001, and the additional 162 birds from 2021. These samples also included 74 nestlings from 2001 and 106 nestlings from 2021 which were not included in this study.

2.3 Laboratory work

2.3.1 DNA extraction

The DNA from the 2001 samples was extracted by Lars Erik Johannessen following the method of Johannessen et al. (2005). DNA from the 2021 samples was extracted from blood samples using the E.Z.N.A® Blood and Tissue DNA kit (OMEGA bio-tek), with the following changes: a) Approximately 50µL of blood tissue preserved in ethanol were transferred to a 1,5mL tube. The remaining ethanol was removed using a vacuum centrifuge. b) To increase DNA yield, incubation of heated elution buffer was increased to 10-15 minutes in the final elution step. Complete list of samples provided in Appendix 7 (Supplementary Table 9, 10 and 11).

2.3.2 MHC – polymerase chain reaction (PCR)

MHC class 1 exon 3, hereafter referred to as MHCI, was amplified using the gene target primer pair HNalla/HN46 (O'Connor et al., 2016). The 2001 samples were amplified using primers consisting of the gene target motif and a 12 base pair (bp) index given in Fadrosh et al. (2014). The 2021 samples were amplified with the full primer design (Illumina adapters, index, heterogenic spacers and gene target motif) as

described in Fadrosh et al. (2014). All samples were amplified in duplicates in independent PCR reactions (See Appendix 1 Supplementary Table 1 for PCR setup and thermal profile). A complete list of the primers used in this project is provided in Appendix 2 (Supplementary Table 3). The following conditions were met in the PCR reactions; 1X Q5® high fidelity master mix (New England Biolabs (NEB)), 0.5µM forward and reverse primer, 1.5µL DNA template and Nuclease-free water to a total volume of 25µL.

Cleanup and library preparation of the 2001 samples were utilized following the NEBNext Library Prep kit for Ion Torrent (New England Biolabs) and size selected using the BluePippin (Sage Science). The final libraries were sequenced on two Ion530 Chips Thermofisher) using an Ion S5™ System (Thermofisher).

Amplicons were visualized on agarose gels and quantified using the ImageLab software v6.0 (Bio-Rab laboratory). Based on the quantification, uniform amounts of each amplicon were merged using the Biomeck 4000 liquid handling robot (Beckman coulter). These pools of amplicons are hereafter referred to as libraries. The library from the 2021 samples were cleaned using Illustra ExoStar (Merck) and concentrated using the DNA clean & Concentrator kit (Zymo research). As a final step to remove leftover primers, the library was cleaned using 1.2X AMpure beads. The final library was sequenced on an Illumina MiSeq instrument at the Norwegian Sequencing Centre.

2.3.3 Blood parasites – polymerase chain reaction (PCR)

Blood parasites from the protist genera *Haemoproteus*, *Plasmodium* and *Leucocytozoon*, were detected in adult birds following the methods in Ciloglu et al. (2019). PCR amplicons yielding inadequate genotyping scores were reran in additional PCRs with slight changes to the protocol (See Appendix 1 Supplementary Table 2 for PCR setup and thermal profile). Blood parasite amplicons were then run on 1.5% agarose gels to determine which parasites were present in the given individuals. The final gel was run on 90V for one hour. FastRuler Low Range (Thermo Fischer) was utilized as DNA ladder, which gives bands on 50, 200, 40, 850 and 1500 bp. Based on Ciloglu et al (2019) I expected *Haemoproteus* at 533 bp, *Plasmodium* at 378 bp and *Leucocytozoon* at 200 bp (Figure 1).

2.4 Analyzing the data – Bioinformatics

A pipeline script for the bioinformatics was prepared and provided by Silje Larsen Rekdal, who carried out most of the bioinformatics programming. The final dataset consisted exclusively of those individuals who had two amplicons with > 1500 reads.

The steps for the bioinformatics of the IonTorrent data and the MiSeq data were mainly similar, but they differed in one aspect. First, using FLASH (Magoč & Salzberg, 2011), the paired MiSeq reads were merged. This first step was not necessary for the IonTorrent dataset due to them not having paired reads. Thereafter they were quality filtered, where we removed the reads where > 5% of the bases had a phred score of < 20, using fastx toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). For the MiSeq dataset only 3% (454 597 reads) were discarded for the first library and 2% (313 006 reads) were discarded for the second library, whereas for the IonTorrent dataset 58% (6 151 572 reads) were discarded for the first library, and 55% (5 456 749 reads) were discarded for the second library. This difference could be due to homopolymer errors that is a known issue with Ion Torrent data (Loman et al., 2012; Quail et al., 2012), or due to the phred scores not being directly comparable between technologies from different platforms. After this first filtration, the average number of reads per amplicon for the IonTorrent data were very similar (Lib1: 37 482 reads per amplicon and Lib2: 37 494 reads per amplicon) whereas the MiSeq data differed a bit more (Lib1: 6 543 reads per amplicon and Lib2: 4 426 reads per amplicon). The heterogenetic spacer was removed in the MiSeq data, with standard unix-codes. jMHC (Stuglik et al., 2011) took care of the clustering of identical reads and assigning them to individuals. Variants with < 3 reads in any amplicon, and those with < 0.20% per-amplicon frequency were filtered out as these are more likely to be artefacts (Rekdal et al., 2018). The variants were thereafter trimmed to 261 +/-3 basepairs and aligned to previously published MHCI-sequences from the willow warbler (GenBank accessions KU169602-KU169603, KU169705-KU169706, KU169709-KU169713, KU169715-KU169719 (O'Connor et al., 2016)). Next, variants with a shift in the reading frame, that had stop codons or that lacked crucial cysteine residues (Cys7 and Cys70, Connor et al 2016) were filtered out. Due to the trimming, some variants were now identical and therefore clustered. Further, all remaining variants that had a frequency of more than 0.20% in an amplicon, were considered putative alleles for that amplicon. Using the uchime3 denovo algorithm (Edgar, 2016), chimera detection was conducted in every amplicon. The variants were scored as alleles if listed as a putative allele for both amplicons for one individual, and not listed as a chimera in both. In addition, the > 1 bp variants (variants that are more than one basepair different from their most similar, more frequent variant in the same amplicon) were scored as allele if present in another individual across the whole dataset. Lastly, the positively selected sites (PSS) alleles were established, in order to solely base the analysis on the functionally important sites (sites 1, 3, 5, 19, 21, 57, 60, 61; (Balakrishnan et al., 2010; Rekdal, 2020)). In and Rekdal (2020), these sites were positively selected for other passerines and the bluethroat respectively (See Appendix 3). The number of individuals with identified genotypes were 95 for the IonTorrent data, and 151 for the MiSeq data. However, the nestlings were not used for further analysis, leaving 22 adult

individuals from the IonTorrent data and 47 adult individuals from the MiSeq data. In the end there was generated a data frame with the number of unique alleles within each social pair, in addition to a data frame with the number of unique alleles within each possible pair (meaning all possible female-male combinations).

One important note: Although I could not detect amplification of negative controls in the agarose gels, I obtain a substantial amount of sequence reads from these controls in the IonTorrent data. Sequence reads from the negative controls have most likely been introduced by leftover primers in the library amplification. The data from this dataset was still included in the study. The allele scoring pipeline requires replicates and Mendelian inheritance patterns. It is unlikely that the lab introduced PCR errors have appeared independently in distinct amplicons. Hence, the noise introduced by these respective sequence reads is expected to be eliminated in the bioinformatics pipeline, and the chance they could introduce any qualitative bias to the data set is negligible.

2.5 Ethical note

The study complies with the current laws of Norway, and permission to ring the birds with metal identification rings was approved by the Norwegian Environment agency (2014/2620). Permission to ring with color rings, take blood samples and keep a bird in a cage was given by the Norwegian Food Safety Authority (2001: S-1456/01; 2021:2020/23426).

2.6 Statistical analyses

All statistical analyses were done in R, using RStudio (version 4.1.1). I tested for normal distribution with Shapiro-Wilk tests, and adjusted for multiple testing by calculating the false discovery rate with the Benjamini-Hochberg method (Benjamini & Hochberg, 1995). The MHCI variables for 2001 and 2021 were centered to justify pooling of the variables due to the difference in number of alleles between the sequencing platforms. Tarsus and wing length of both sexes were also centered due to the significant difference in length (See Appendix 4 for sex-separated tests). Tests performed on clutch size and egg laying date was centered across both years, and they were performed exclusively on females, as the male has no or little influence on these parameters. All tests on the number of fledged brood were performed exclusively on the 2021 material, due to missing information from the 2001 material (See Appendix 5 Supplementary Table 8a for full table of measurements and individuals). *Plasmodium* had no infected individuals from 2001, and *Leucocytozoon* had only infected individuals from 2001. All tests for these two blood parasites were therefore exclusively run on the 2021 material. It is also important to mention that

there was some missing age information from the adults of the 2001 material (n = 7), these were not included in age-based analyses (Appendix 5 Supplementary Table 8a).

I use three levels of MHCI variation in this study: 1) Number of PSS alleles, 2) distance from the optimal number of PSS alleles, square rooted in order to make the distribution normally distributed. The optimum was set to the population mean for number of alleles across all adults for both datasets (Aeschlimann et al., 2003; Rekdal et al., 2019). 3) The presence / absence of specific alleles.

Analyses were run in order to test whether there was an intermediate optimum in the number of alleles within the observed pairs, compared to the random pairs (Rekdal et al., 2019). I also tested for the absolute number of unique alleles in a pair, in order to test whether there can be any selection for mates that produce offspring with many, or few, MHCI alleles (i.e., if the optimal number of alleles are at a maximum/minimum rather than at an intermediate level). A file was generated for all possible pairs for each dataset (2001 and 2021). For each analysis a female from the given dataset was picked and paired with a random male. The male was thereafter put back into the pool. This process was repeated for all females in the sample. A mean was thereafter calculated. This process was repeated 10 000 times for each year (2001 and 2021) in each analysis. Separate analyses for the MiSeq dataset (2021) were ran where pairs were separated according to the females' age (SY and ASY), due to only having 3 SY males and 29 ASY males. This was due to the IonTorrent dataset (2001) missing some age information.

Tests were performed to investigate whether there were any associations between MHCI, morphological traits and reproductive success. Due to only the wing and tarsus measurements being normally distributed, non-parametric tests were performed for most variables. To test for correlations, Spearman correlation tests were performed for MHCI against morphological measurements, and reproductive success, while the number of fledged offspring within female individuals were tested with a Pearson correlation test.

In order to test for a relationship between the MHCI and blood parasite prevalence, Wilcoxon rank sum tests were performed. This test was supported by a statistically more correct logistic regression due to the blood parasite prevalence being the dependent variable, and the MHCI measures being the independent variable. Additionally, to test whether specific MHCI alleles were associated with the blood parasites, Fisher's exact tests were performed on the presence / absence of the MHCI alleles and each of the three parasite genera. A control for multiple testing was performed as described by Benjamini and Hochberg (1995). The tests are significant after correction if the p-value is lower than a threshold value, which is based on the rank of the test, number of tests in each group (which is 31 due to having 31 PSS alleles) and

the number of false positive tests allowed. I also tested whether blood parasite prevalence was associated with morphology, reproductive success and age, using unpaired t-tests (blood parasite versus wing and tarsus length, and number of fledged offspring from females), Wilcoxon rank sum tests (blood parasite versus mass, and egg laying date and clutch size) and Fisher's exact test (blood parasite prevalence and age correlation).

3.0 Results

3.1.1 MHC1 variation

Out of 46 MHCI nucleotide alleles from 2001 and 78 MHCI nucleotide alleles from 2021, they grouped into 14 PSS alleles for the 2001 material, and 33 PSS alleles for the 2021 material in the minimum of one individual. The average number of nucleotide alleles per individual was 12.36 from the IonTorrent (range 4 – 21), and 6.37 from the MiSeq data (range 2 – 13). The estimated number of loci was 11 and 7 for the 2001 and 2021 data, respectively. On average there was 4.59 PSS alleles per individual in the IonTorrent data from 2001 and 3.48 PSS alleles per individual in the MiSeq data from 2021.

3.1.2 Blood parasite prevalence

Blood parasite scoring in a selection of adult blue tits from 2001 and 2021 (is shown in Figure 1). There are no occurrences of *Plasmodium* in the 2001 data (Table 1). *Haemoproteus* and *Leucocytozoon* is distributed across both years, however all individuals from 2001 were infected by *Leucocytozoon* (Appendix 5 Supplementary Table 8b).

Table 1: Table of blood parasite infected individuals.

Year	n	*Haemoproteus*	*Plasmodium*	*Leucocytozoon*	Multiple infection
2001 Male SY	8	5 (62.5%)	0	8(100%)	5 (62.5%)
2001 Male ASY	3	1 (33.33%)	0	3 (100%)	1 (33.33%)
2001 Female SY	10	3 (30%)	0	10 (100%)	3 (30%)
2001 Female ASY	1	0	0	1 (100%)	0
2021 Male SY	3	0	0	3 (100%)	0
2021 Male ASY	29	10 (34.48%)	5 (17.24%)	24 (82.76%)	12 (41.38%)
2021 Female SY	8	0	1 (12.5%)	6 (75%)	0
2021 Female ASY	18	5 (27.78%)	8 (44.44%)	16 (88.89%)	11 (61.11%)
Total	80	24 (30%)	14 (17.5%)	71 (88.75)	32 (40%)

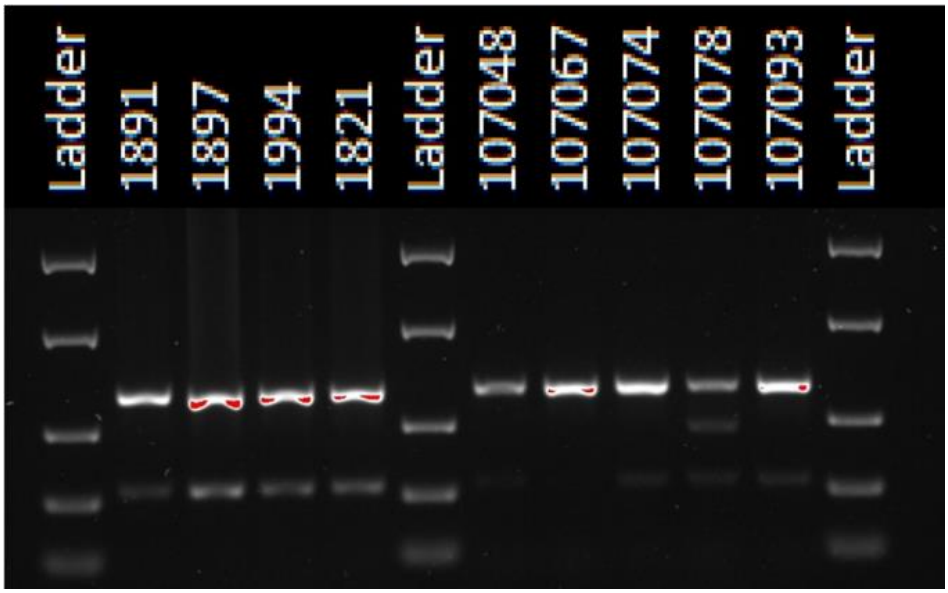


Figure 1: Blood parasite test on gel on a selection of adult individuals of blue tits from 2001 (1891, 1897, 1994 and 1821) and 2021 (107048, 107067, 107074, 107078 and 107093), testing for the presence of blood parasites Haemoproteus (533 base pairs (bp)), Plasmodium (378 bp) and Leucocytozoon (218 bp). The ladder has the following bands (bp): 1500, 850, 400, 200 and 50). All individuals from 2001 have Haemoproteus and Leucocytozoon. All individuals from 2021 have Haemoproteus, four individuals have Leucocytozoon (107047, 107074, 107078 and 107093) and one individual has Plasmodium (107078).

3.2 The MHCI composition of observed and random pairs

Overall, there was no significant departure from random social mating in the two study years, neither with respect to the combined number of alleles in pairs (exact test; p = 0.51; Figure 2) nor the distance to the optimum number of alleles in pairs (exact test; p = 0.50; Figure 3).

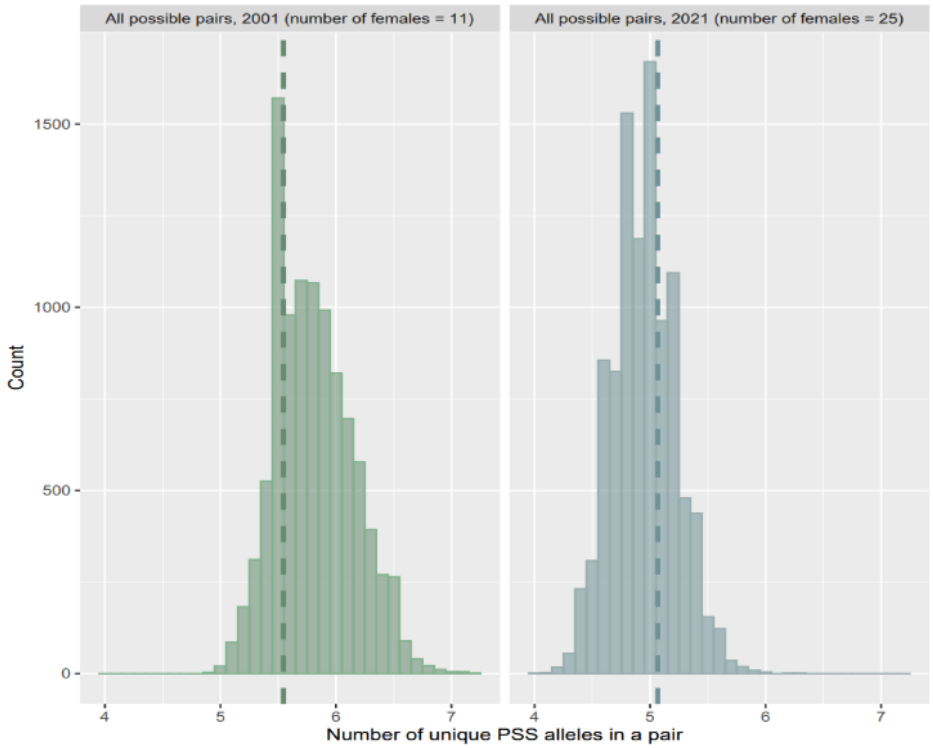


Figure 2: Histograms of 10 000 means of number of PSS alleles for random blue tit pairs for 2001 (left panel) and 2021 (right panel). The dashed line represents the mean of the observed pairs (2001; n pairs = 11. 2021; n pairs = 15).

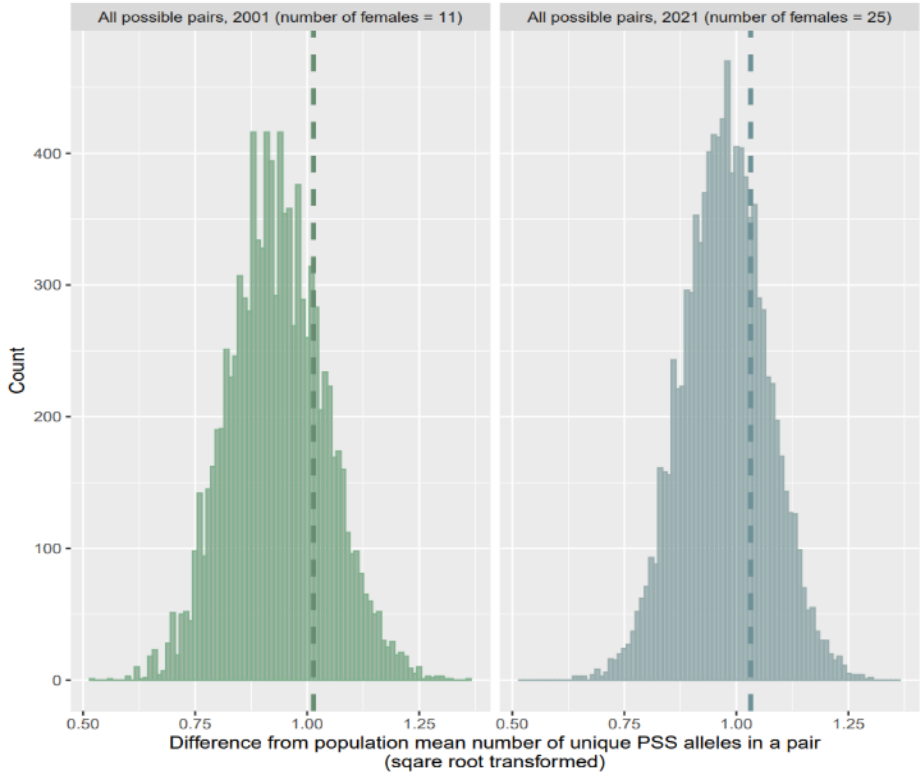


Figure 3: Histograms of 10 000 mean distances to the optimum for random blue tit pairs for 2001 (left panel) and 2021 (right panel). The dashed line represents the mean of the observed pairs (2001; n pairs = 11. 2021; n pairs = 15).

I also tested the differences between younger (SY) and older (ASY) females in a pairing pattern within the 2021 material. The observed SY females (Figure 4) were in pairs that had fewer PSS alleles than the random pairs (exact test: compared to all simulated pairs: p = 0.015), whereas the observed ASY females proved to be in pairs that had more PSS alleles than the random pairs (exact test: compared to all simulated pairs: p = 0.038). Testing for associations between females age and their partners number of MHCI alleles in the 2021 material gave a close to significant result. SY females tended to choose males with fewer alleles, and ASY females tended to choose males with more alleles (Wilcoxon rank sum test: W = 13.5, p = 0.07). There was no significant departure from random mating with respect to the distance to the mean number of unique PSS alleles in a pair (Figure 5), neither for SY females (exact test: compared to all simulated pairs: p = 0.29) nor for ASY females (exact test: compared to all simulated pairs: p = 0.087).

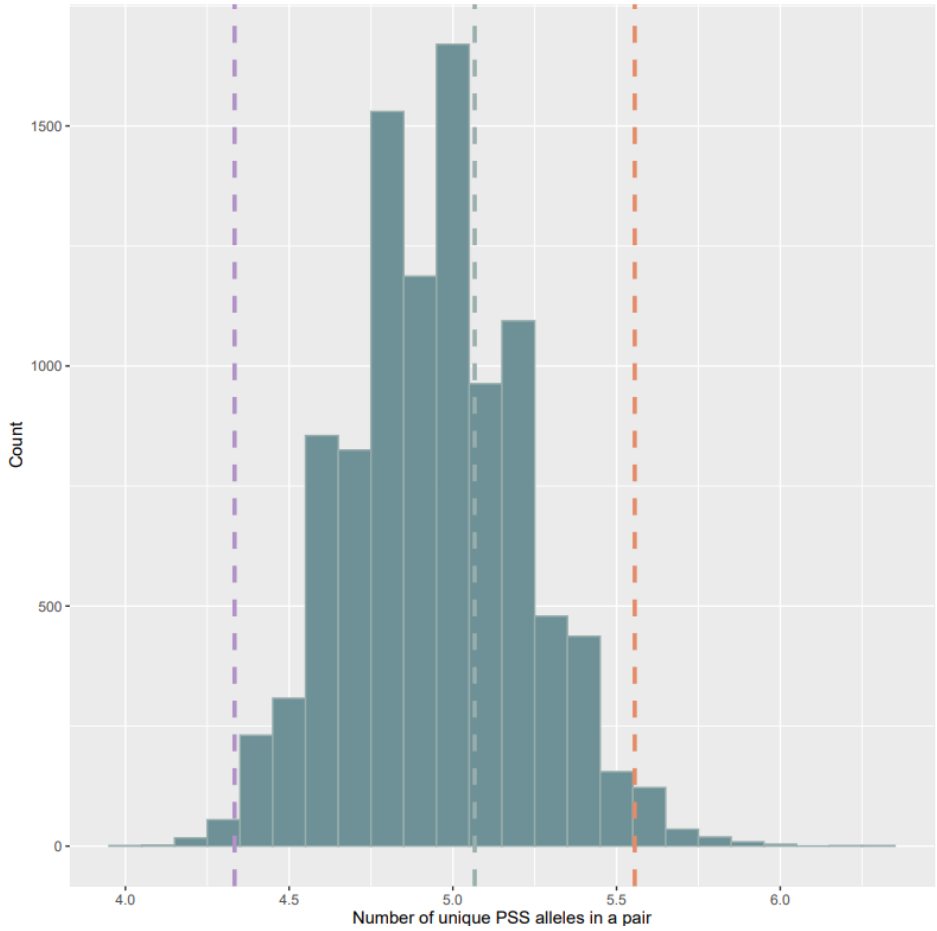


Figure 4: Histogram of 10 000 means of PSS allele counts of all possible blue tit pairs from 2021. Grey dashed line represents the mean of the observed pairs. Purple dashed line; SY = second year, a social pair where the female is in her second calendar year. Orange dashed line; ASY = after second year, a social pair where the female is in her third or more calendar year (n SY female = 8, n ASY female = 18).

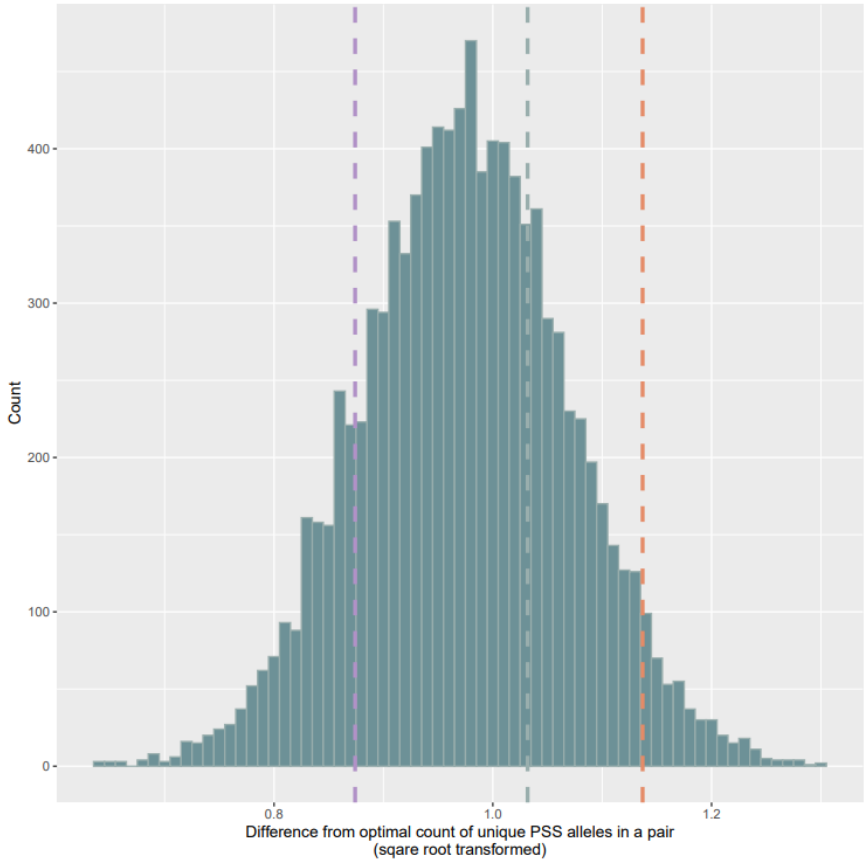


Figure 5: Histogram of 10 000 mean distance to the optimum of all possible blue tit pairs from 2021. Grey dashed line represents the mean of the observed pairs. Purple dashed line; SY = second year, a social pair where the female is in her second calendar year. Orange dashed line; ASY = after second year, a social pair where the female is in her third or more calendar year (n SY female = 8, n ASY female = 18).

3.3 Associations between MHCI variation, morphological traits, age and reproductive success

There was a significant positive association between the number of MHCI alleles and wing length (r_s = 0.31, p = 0.025), however when correcting for multiple testing it was no longer significant. This trend was not seen for separate sexes (Appendix 4 Supplementary Table 4). Neither of the remaining tests between MHCI variables and morphological measurements provided significant associations (Table 2).

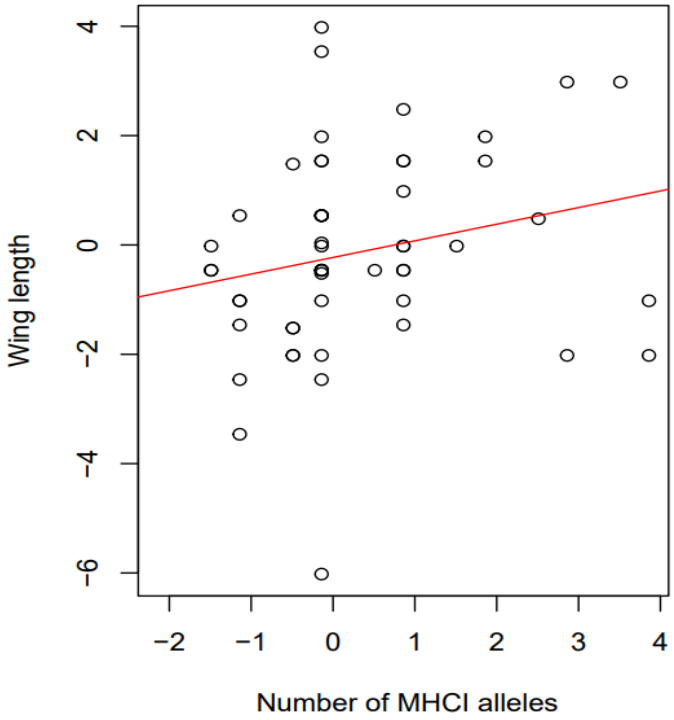


Figure 6: Scatterplot of centered number of MHCI alleles and centered wing length in the blue tit (n = 54). Red line = linear regression line (y = 0.3x - 0.14) added for visual purposes.

Table 2: Tests of association between MCHI variation and morphological variables.

MHCI variable	Morphological variable	Test type	n	r_s	p-value
Number of MHCI alleles	Tarsus length	Spearman correlation	54	0	1.00
Number of MHCI alleles	Wing length	Spearman correlation	54	0.31	0.025
Number of MHCI alleles	Mass	Spearman correlation	159	-0.04	0.60
Distance to mean	Tarsus length	Spearman correlation	54	0.04	0.76
Distance to mean	Wing length	Spearman correlation	54	-0.07	0.62
Distance to mean	Mass	Spearman correlation	57	0.05	0.74

Furthermore, there were no significant differences between the age-classes in MHCI variation (Table 3). However, there was a significant association between the age and distance to mean in males after separating for sexes, showing that older males were furthest from the mean (W = 58, p = 0.046), (Appendix 4 Supplementary Table 5). There were no significant associations between MHCI variability and measurements of reproductive success (Table 4).

Table 3: Tests of association between MHCI variation and age.

Grouping variable	Dependent variable	Test type	n SY	n ASY	W	p-value
Age	Number of MHCI alleles	Wilcoxon rank sum test	22	39	380	0.45
Age	Distance to mean	Wilcoxon rank sum test	22	39	432	0.97

Table 4: Tests of association between MHCI variation and fitness variables.

MHCI variable	Fitness variable	Sex	Test type	n	r_s	p-value
Number of MHCI alleles	Egg laying date	Female	Spearman correlation	34	0.01	0.97
Number of MHCI alleles	Clutch size	Female	Spearman correlation	34	-0.09	0.59
Number of MHCI alleles	Number fledged	Female	Pearson Correlation	23	0.18	0.42
Number of MHCI alleles	Number fledged	Male	Spearman correlation	21	0.27	0.23
Distance to mean	Egg laying date	Female	Spearman correlation	34	-0.24	0.18
Distance to mean	Clutch size	Female	Spearman correlation	34	0.29	0.10
Distance to mean	Number fledged	Female	Pearson correlation	23	-0.04	0.85
Distance to mean	Number fledged	Male	Spearman correlation	21	0.15	0.52

3.4 Associations between MHCI variation and blood parasite prevalence

There was a significant positive association between the number of MHCI alleles and the prevalence of *Haemoproteus* for both years combined (W = 302, p = 0.037) (Figure 7), though when correcting for multiple testing, it was no longer significant. However, when separating by year, only the 2001 material provided a significant positive association (Wilcoxon rank sum test; 2001: W = 23, p = 0.015. 2021: W = 132, p = 0.32) (Figure 8 and Figure 9). The logistic regression confirmed these results (centered values: Estimate = 0.49, p = 0.028; 2001: Estimate = 1.16, p = 0.032; 2021: Estimate = 0.31, p = 0.27). Neither of the remaining tests provided significant results (Table 5).

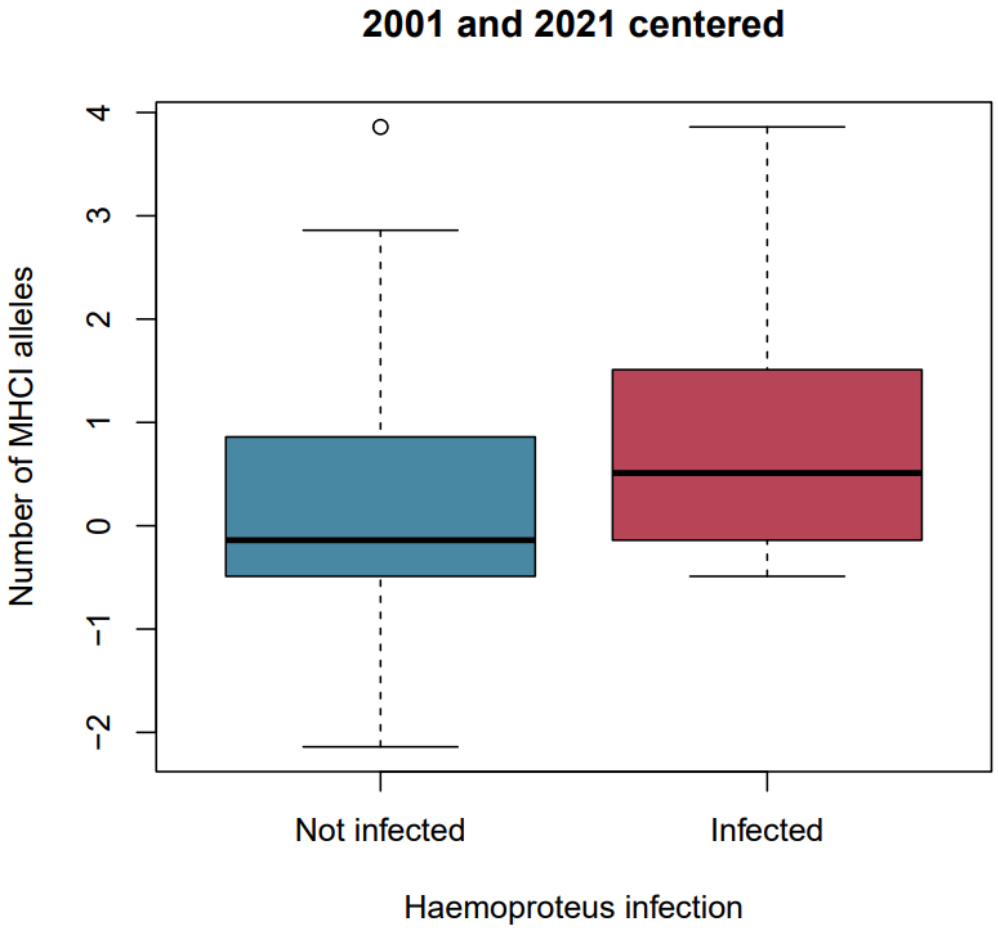


Figure 7: Boxplot of number of MHCI alleles for uninfected (blue box, n = 50) and infected (red box, n = 18) blue tits from 2001 and 2021 combined.

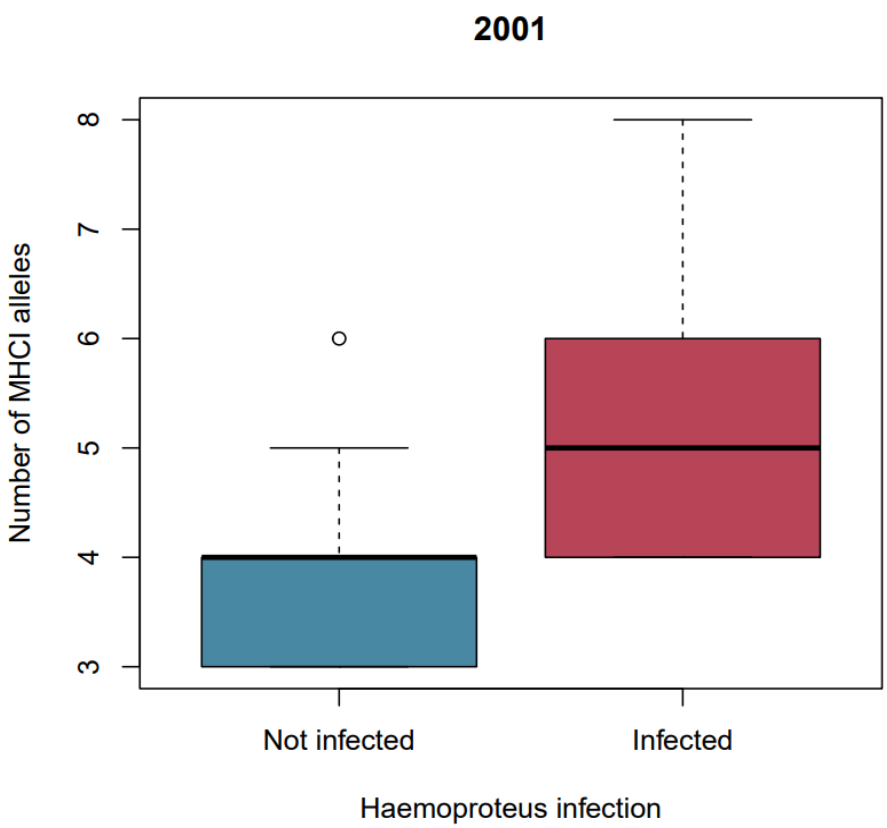


Figure 8: Boxplot of number of MHCI alleles for uninfected (blue box, n = 13) and infected (red box, n = 9) blue tits from 2001.

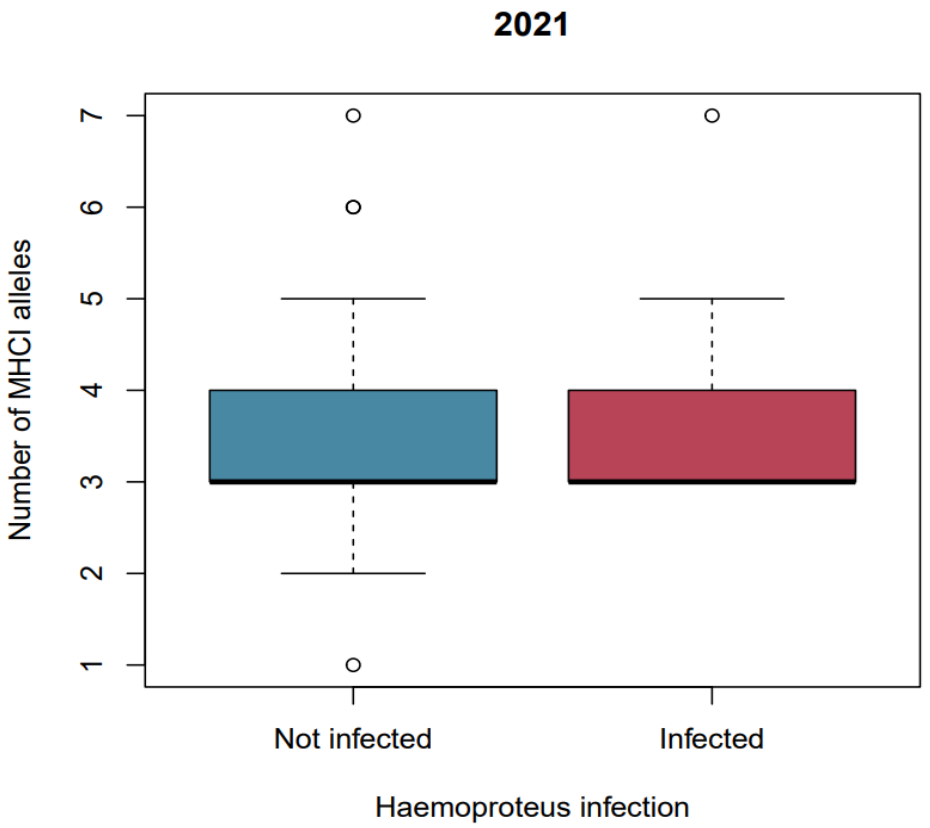


Figure 9: Boxplot of number of MHCI alleles for uninfected (blue box, n = 37) and infected (red box, n = 9) blue tits from 2021.

The association between *Haemoproteus* and the PSS sequence IWYWYVGR across both years showed that individuals carrying the allele were more likely to be infected (Fisher's exact test; p = 0.0054, the threshold p-value after correcting for multiple testing was 0.0032). This test was not quite significant after correcting for multiple testing. Since only one individual from 2021 had the allele present (ID: 107138, see Appendix 5 Supplementary Table 8b), I performed a separate test only on the 2001 data. This test showed a positive association between the presence of the allele and the presence of *Haemoproteus* (Fisher's exact test; p = 0.027), however when correcting for multiple testing this value is not significant.

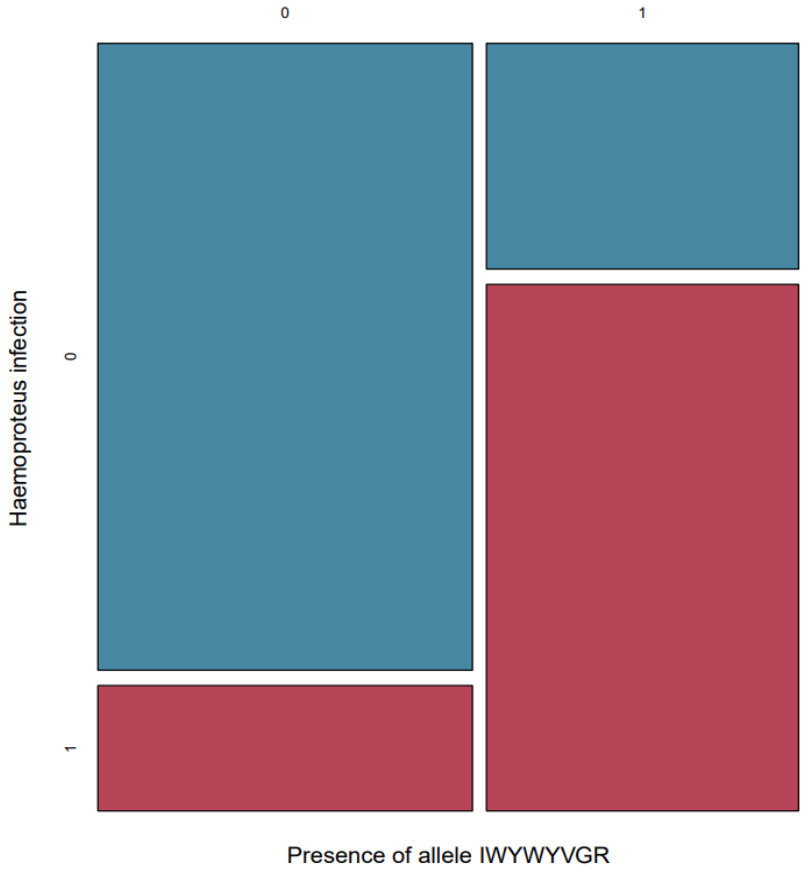


Figure 10: Mosaic plot of relationship between the presence of allele IWYWYVGR and Haemoproteus infection for uninfected and infected blue tits from 2001 (n Infected with allele present = 7, n infected with allele not present = 2, n not infected with allele present = 3, n not infected with allele not present = 10).

Table 5: Tests of association between the MHCI variables and blood parasite prevalence.

Grouping variable	Dependent variable	Test type	n infection	n not infected	W	p-value
Haemoproteus	Number of MHCI alleles	Wilcox rank sum test	18	50	302	0.037
Haemoproteus	Distance to mean	Wilcox rank sum test	18	50	468	0.80
Plasmodium	Number of MHCI alleles	Wilcox rank sum test	14	32	238	0.73
Plasmodium	Distance to mean	Wilcox rank sum test	14	32	242	0.66
Leucocytozoon	Number of MHCI alleles	Wilcox rank sum test	40	6	131	0.72
Leucocytozoon	Distance to mean	Wilcox rank sum test	40	6	96	0.42

3.5 Associations between blood parasite prevalence, morphological traits, age and reproductive success

There were no significant associations between *Haemoproteus* (both years combined) and morphological measurements, nor between *Plasmodium* and *Leucocytozoon* (2021 exclusively) and morphological measurements (Table 6). It is, however, worth mentioning that *Haemoproteus* and wing length had a close to significant association (p = 0.060), as the infected individuals tended to have longer wings.

Table 6: Tests of association between morphology variables and blood parasite prevalence.

Grouping variable	Dependent variable	Test type	n infection	n not infected	Test statistic	p-value
Haemoproteus	Tarsus length	t-test	21	45	t = -0.83	0.41
Haemoproteus	Wing length	t-test	21	45	t = -1.9	0.060
Haemoproteus	Mass	Wilcox rank sum test	20	48	W = 396	0.26
Plasmodium	Tarsus length	t-test	11	43	t = 0.55	0.59
Plasmodium	Wing length	t-test	11	43	t = 0.43	0.67
Plasmodium	Mass	Wilcox rank sum test	14	42	W = 320.5	0.62
Leucocytozoon	Tarsus length	t-test	46	8	t = 1.44	0.18
Leucocytozoon	Wing length	t-test	46	8	t = -1	0.34
Leucocytozoon	Mass	Wilcox rank sum test	48	8	W = 159	0.45

Furthermore, there were no significant difference between the age groups in blood parasite prevalence (Table 7), nor in the sex-separated tests (Appendix 4 Supplementary Table 7). Additionally, there were no significant difference in fitness-related variables with respect to the blood parasites (Table 8).

Table 7: Table showing the relationship between age and blood parasite infection rate.

Parasite genus	Test type	n infected SY	n not infected SY	n infected ASY	n not infected ASY	p-value
Haemoproteus	Fisher's exact test	6	16	16	35	0.79
Plasmodium	Fisher's exact test	1	10	13	34	0.26
Leucocytozoon	Fisher's exact test	9	2	40	7	1.00

Table 8: Tests of association between blood parasite prevalence and fitness variables.

Grouping variable	Dependent variable	Sex	Test type	Test statistic	n infected	n not infected	p-value
Haemoproteus	Clutch size	Female	Wilcoxon rank sum test	111.5	8	27	0.90
Haemoproteus	Number fledged	Female	t-test	1.05	5	19	0.34
Haemoproteus	Number fledged	Male	Wilcoxon rank sum test	115	10	22	0.85
Haemoproteus	Egg laying date	Female	Wilcoxon rank sum test	107.5	8	27	1.00
Plasmodium	Clutch size	Female	Wilcoxon rank sum test	69	9	17	0.95
Plasmodium	Number fledged	Female	t-test	-1.02	9	15	0.32
Plasmodium	Number fledged	Male	Wilcoxon rank sum test	44.5	5	27	0.23
Plasmodium	Egg laying date	Female	Wilcoxon rank sum test	79	9	15	0.51
Leucocytozoon	Clutch size	Female	Wilcoxon rank sum test	32	21	3	1.00
Leucocytozoon	Number fledged	Female	t-test	0.70	21	3	0.54
Leucocytozoon	Number fledged	Male	Wilcoxon rank sum test	79	27	5	0.56
Leucocytozoon	Egg laying date	Female	Wilcoxon rank sum test	48.5	21	3	0.15

4.0 Discussion

Overall, there was no evidence of deviation from random mating, but I found a distinction between the younger and older blue tit females with respect to non-random mating preferences and MHCI diversity. Younger females were found in pairs that had fewer PSS alleles than the average of the random pairs, whereas the older females were found with partners resulting in a higher number of PSS alleles compared to the average of the random pairs. A positive association between the number of MHCI alleles and wing length was found. There was a positive correlation between the number of MHCI alleles and the occurrence of *Haemoproteus* in the 2001 material. The MHCI allele testing also gave an almost significant result, linking a specific allele (IWYWYVGR) to the blood parasite *Haemoproteus*. Lastly, an almost significant result was also found for *Haemoproteus* infection and wing length, with longer-winged birds being more likely to be infected. No significant results were found for MHCI, measures of reproductive success, nor between blood parasites and morphological traits.

4.1.1 MHCI variation

There was a larger variety of nucleotide alleles at the individual level within 2001 (4 – 21 alleles) compared to 2021 (2 – 13 alleles). Based on the maximum number of nucleotide alleles the estimated number of loci is 11 and 7 for the IonTorrent data and MiSeq data, respectively. These were minimum estimates assuming heterozygosity for all but one allele. There was less than half the number of PSS alleles for 2001 (14) than for 2021 (33), however the average number of PSS alleles per adult individual was higher for 2001 (4.59) than for 2021 (3.48). Schut et al. (2011), Wutzler et al. (2012), Aguilar et al. (2016) and O'Connor et al. (2016) found 4, 10, 4 and 5 MHCI loci, respectively. Compared to these results, the estimated number of loci in the present study is thus generally somewhat higher than previous estimates, which could be due to differences in methodology. It is important to mention that the 2021 sample was much larger than the 2001 sample, explaining the higher total number of alleles found in 2021 compared to 2001.

The 2001 subset was sequenced using IonTorrent and the 2021 subset was sequenced using illumina. There was a huge difference in the number of reads per amplicon in the IonTorrent data (lib 1 = 37 482, lib 2 = 37 494) and the MiSeq data (lib 1 = 6 543, lib 2 = 4 426). The distinct MHCI diversity observed between the two subsets may hence be a result of different read depth, and/or could be related to the separate sequenced platform used for data generation.

From lab-work and sequencing, some individuals were lost, or filtered out due to failing in the PCR or sequencing, having amplicons with < 1500 reads and/or being partners with lost individuals. If the cut-off

limit of 1500 reads per amplicon had been lowered to 1000, I would have had six additional adults and five additional pairs in the analyses. However, the limit of 1500 reads per amplicon was set as the cut-off value due to the following: all possible alleles had > 0.2% per amplicon frequency (PAF) (Rekdal et al., 2018), and that all alleles should be present in at least three PCR-reads (Zagalska-Neubauer et al., 2010). After the filtration, I was still left with 68 adult individuals and 26 pairs. I believe this to be a sufficient amount of data for the present analyses, as well as more trustworthy due to the filtration steps (see Appendix 5 Supplementary Table 8a).

4.2 The MHCI composition of observed and random pairs

I found that younger females were with a partner resulting in a social pair that had fewer PSS alleles compared to expectations from random pairing (Figure 4), whereas it was the other way around for the older females. From testing the females age in relation to their partners number of MHCI alleles, a close to significant difference was found, indicating that younger females chose males with fewer alleles, and older females chose males with more alleles. There might be several reasons for this age specific pattern. One possibility is that older females are more experienced and dominant, thus making them better at judging male MHCI variability. In turn this will ensure them a mate with a better MHCI compatibility. As a result, younger females must settle for less optimal males. Another possibility could be that younger and older females start their egg laying at different dates. If the older female lay their eggs earlier, they might also be able to find a mate earlier ensuring a partner with an optimal MHCI variability for their offspring. In the present study, the difference between the age-groups in egg-laying dates for both years was not significant (data not shown). Due to the correlative nature of the collected data, I cannot distinguish between these explanations. Both options are based on the older females' choice (more PSS alleles than the random pairs) being the better option.

From the study by Richardson et al. (2005) it becomes clear that females with a social mate that had a low MHCI diversity were more likely to obtain EPP with a higher MHCI diversity. Similarly, Eizaguirre et al. (2009) found a deviation from random mating in the three-spined sticklebacks due to an interaction between the MHCI diversity and mate choice. Rekdal et al. (In review) found a similar pattern with mate preference and age within the bluethroat; within-pair young (WPY) from broods with no extra-pair young had a more optimal MHCII variability compared to WPY from mixed broods, but only in broods of younger males. However, unlike Rekdal et al. (2019) and the ongoing study by Rekdal et al. (In review), the present study found relationships with the number of MHCI alleles instead of the distance to optimum. There might be several reasons for this, one being differences between the species involved in the studies. Another reason

could be that in Rekdal et al. (2019) and Rekdal et al. (In review), the focus is on MHCII, whereas in the present study the focus is on MHCI.

4.3 Associations between MHCI variation, morphological traits, age and reproductive success

I found a significant positive association between the number of MHCI alleles across both study years and the length of the birds' wing across both sexes. This test was no longer significant after correcting for multiple testing, nor when separating for the sexes.

For most birds, vision is the dominant sense (Rajchard, 2009), partly due to food detection, predator detection and air navigation (Jones et al., 2007; Moore et al., 2015; Tedore & Nilsson, 2021), but also in relation to finding a potential mate (Jones et al., 2007). It is therefore reasonable to think that visual cues such as morphological traits, wing length in this case, can play an important role in mate choice. Looking back at the good genes hypothesis, wing length could in this case be considered a favorable trait indicating that they can provide favorable genes for future offspring, assuming that a higher number of alleles is favorable. However, there are relatively few studies that find significant associations between MHC variation and morphological traits. A study of yellowthroats (*Geothylpis trichas)* found an association between the black mask and a greater MHCII variation in a Wisconsin population, and between the yellow bib and MHCII variation in a New York population (Whittingham et al., 2015). Møller et al. (1998) suggested that fixed traits (morphological such as wing length, etc.) were more reliable than flexible traits (song, courtship indicators, etc.), and Candolin (2003) suggested that such traits could influence the mating decision. Møller et al. (1998) also suggested that a fixed trait could aid attention to flexible ones. Slade et al. (2017) tested for the relationship between song and MHCII diversity in song sparrows (*Melospiza melodia*) as they assumed that the song advertised the singers' genetic diversity to potential mates. However, they found no relationship but didn't exclude the possibility of an association between MHCI and song. It is also reasonable to think that there is a covariance between the MHCI variability and phenotypical factors that the birds can sense. For example, there are reports of ultraviolet sexual dimorphism and assortative mating in blue tits (Andersson et al., 1998; Hunt et al., 1999), as well as odor based-mate choice in songbirds and blue petrels (*Halobaena caerulea*), among others (Grieves et al., 2019; Leclaire et al., 2019; Strandh et al., 2012). It is plausible that female blue tits perceive the wing length as a direct indicator of the males' MHCI diversity. Another plausible scenario is that wing length is correlated with another variable the females use as a cue for mate choice.

Interestingly, there was also a significant association between age and distance to optimum for males in the sex separated tests (Appendix 4 Supplementary Table 5) where the older males were further away from the mean than younger males. There were no significant associations between age and the number of MHCI alleles, implying that older males tended to have either a low or a high number of alleles. However, this test was no longer significant after correcting for multiple testing. Still, the result is puzzling and should be subject to further study.

In this study there were no associations between MHCI variability and reproductive success, however several studies provide evidence for a significant relationship between the two variables (Bernatchez & Landry, 2003; Forsberg et al., 2007; Olsson et al., 2005; Sauermann et al., 2001; Sepil et al., 2013; Thoss et al., 2011). Sauermann et al. (2001) found that male macaques (*Macaca mulatta*) that were heterozygous for MHCII sired significantly more offspring compared to homozygotes. Olsson et al. (2005) found that male sand lizards (*Lacerta agilis*) which carried a particular restriction fragment length polymorphism MHCI genotype had greater reproductive success, compared to the males that did not carry this genotype. A link between the MHCI and reproductive fitness has been found within the great tit (*Parus major*) (Sepil et al., 2013), and between MHCII and reproductive fitness in the brown trout (*Salmo trutta*) (Forsberg et al., 2007). Thus, it is plausible that such a relationship can be found in a larger dataset of the blue tits as well.

4.4 Associations between MHC variation and blood parasite prevalence

The blood parasite composition varied greatly between the two study years, which were separated by 20 years. Both years had individuals that were infected by *Haemoproteus*, whereas *Plasmodium* and *Leucocytozoon* showed opposite patterns in 2001, with no individuals infected with *Plasmodium*, and all individuals from this year carrying *Leucocytozoon*. This difference could be due to coincidences due to small datasets, or to a real change in the parasite fauna the blue tits were exposed to. The results from this study did not support the prediction that an intermediate number of MHCI alleles corresponded with a lower blood parasite infection rate. Instead, individuals with a higher number of MHCI alleles had a greater *Haemoproteus* infection rate, both for the two years combined (Figure 7) and for the 2001 material (Figure 8), compared to the individuals with a lower number of MHCI alleles. This result might suggest a selection towards fewer MHCI alleles in this population of blue tits, though whether a low number of MHCI alleles is better than an intermediate number of MHCI alleles is not certain.

The MHCI allele IWYWYVGR had an almost significant positive association with infection by *Haemoproteus* (Figure 10, Table 5). In other words, individuals that carried this allele were more likely to be infected. From

the alleles found significant in the study by Aguilar et al. (2016), only one was found in the current study (UA104). UA104 was scored as an allele within three individuals, however they were all nestlings and thus there were no blood parasite data for these individuals.

4.5 Associations between blood parasite prevalence, morphological traits, age and reproductive success

I found a close to significant positive association between the infection of *Haemoproteus* and the length of the birds' wing across both sexes. When separating for sexes, these tests were no longer close to significant (Appendix 4 Supplementary Table 6). Morphological traits can act as honest indicators of an organisms' health and fitness as they reflect the individuals' genetic composition and possible illnesses (Hamilton & Zuk, 1982; Mays & Hill, 2004). It is therefore reasonable to think that morphological traits, such as wing length in this case, could reflect the health, and thus parasite infection, of the bird. Unlike previously, longer wings would not be considered a favorable trait, in relation to the good genes hypothesis, if the assumption is that infected individuals have worse health compared to non-infected individuals. There are relatively few studies that find significant associations between blood parasite infection and morphological traits. Associations between *Haemoproteus* infection and body condition has been found in ducks, where the body condition of infected individuals was reduced (Meixell et al., 2016). Associations between infection by *Leucocytozoon* and wing length, body mass and body condition in a variety of duck species has also been reported (Fleskes et al., 2017; Meixell et al., 2016). Infected individuals had a decreased wing length, body mass and body condition.

There is growing evidence that parasites influence host fitness components (Knowles et al., 2010; Rätti et al., 1993). These fitness-related components could be morphological traits, like wing length, or cues such as song, odor, etc., indicating the organisms ability to cope with physical demands necessary for survival or reproductive success (Puente et al., 2010). The measured difference in wing length could in this case be considered a factor reflecting the individuals fitness. However, there were no significant associations between blood parasites and reproductive success. Rätti et al. (1993) found that males infected with *Trypanosoma* had shorter tails and wings compared to the non-infected males. Interestingly, these results are the opposite of those presented in the present study. Podmokła et al. (2014), as mentioned earlier, found that infected individuals increased their reproductive investment, thus there were no difference in reproductive success of infected versus non-infected individuals. The same principles could apply to the present study. The present study also found a significant association between the age and wing length, where older individuals had longer wings (data not shown). It is thus plausible that blue tits infected with

Haemoproteus increase their terminal investment, especially as they age, resulting in no significant differences.

5.0 Conclusion

The results indicate a relatively high MHCI variability in the studied blue tit population. They show that mate choice within the blue tit is MHCI dependent, as both younger and older females have chosen mates that resulted in a different MHCI variation than expected from random pairing. However, it is difficult to say anything about causation, due to non-consistent results for the two age groups, and due to the data being correlative. Positive associations were also found between the number of MHCI alleles, on the one hand, and wing length, and *Haemoproteus* infection, on the other. Even though the correlation between *Haemoproteus* infection and wing length was not significant, the results may be connected, possibly through age as older individuals have longer wings. In addition, a link was discovered between the allele IWYWYVGR and *Haemoproteus* infection.

To summarize, the most interesting results in this study were the age-specific patterns of MHCI mate choice, and the associations between number of MHCI alleles, *Haemoproteus* infection and wing length. My results indicate that older females mate with males that have more alleles, and thus also longer wings and a greater *Haemoproteus* infection rate, whereas younger females choose males with fewer alleles, shorter wings and thus a lower infection rate. However, a greater sample across several years and locations could provide a stronger foundation for the results.

6.0 Reference list

Aguilar, J. R. D., Westerdahl, H., Puente, J. M. D. L., Tomás, G., Martínez, J., & Merino, S. (2016). MHC-I provides both quantitative resistance and susceptibility to blood parasites in blue tits in the wild. *Journal of Avian Biology*, *47*(5), 669-677.

Andersson, S., Örnborg, J., & Andersson, M. (1998). Ultraviolet sexual dimorphism and assortative mating in blue tits. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1395), 445-450.

Atkinson, C. T., Thomas, N. J., & Hunter, D. B. (Eds.). (2009). *Parasitic diseases of wild birds*. John Wiley & Sons.

Balakrishnan, C. N., Ekblom, R., Völker, M., Westerdahl, H., Godinez, R., Kotkiewicz, H., ... & Edwards, S. V. (2010). Gene duplication and fragmentation in the zebra finch major histocompatibility complex. *BMC biology*, *8*(1), 1-19. https://doi.org/10.1186/1741-7007-8-29

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289-300.

Bernatchez, L., & Landry, C. (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years?. *Journal of evolutionary biology*, *16*(3), 363-377. https://doi.org/10.1046/j.1420-9101.2003.00531.x

Candolin, U. (2003). The use of multiple cues in mate choice. *Biological reviews*, *78*(4), 575-595.

Chaix, R., Cao, C., & Donnelly, P. (2008). Is mate choice in humans MHC-dependent?. *PLoS genetics*, *4*(9), e1000184. https://doi.org/10.1371/journal.pgen.1000184

Ciloglu, A., Ellis, V. A., Bernotienė, R., Valkiūnas, G., & Bensch, S. (2019). A new one-step multiplex PCR assay for simultaneous detection and identification of avian *haemosporidian* parasites. *Parasitology Research*, *118*, 191-201. https://doi.org/10.1007/s00436-018-6153-7

Donovan, T. A., Schrenzel, M., Tucker, T. A., Pessier, A. P., & Stalis, I. H. (2008). Hepatic hemorrhage, hemocoelom, and sudden death due to *Haemoproteus* infection in passerine birds: eleven cases. *Journal of Veterinary Diagnostic Investigation*, *20*(3), 304-313. https://doi.org/10.1177/104063870802000307

Dunn, P. O., Bollmer, J. L., Freeman-Gallant, C. R., & Whittingham, L. A. (2013). MHC variation is related to a sexually selected ornament, survival, and parasite resistance in common yellowthroats. *Evolution*, *67*(3), 679-687. https://doi.org/10.1111/j.1558-5646.2012.01799.x

Edgar, R. C. (2016). SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *biorxiv*, 074161.

Eizaguirre, C., Yeates, S. E., Lenz, T. L., Kalbe, M., & Milinski, M. (2009). MHC-based mate choice combines good genes and maintenance of MHC polymorphism. *Molecular ecology*, *18*(15), 3316-3329. http://dx.doi.org/10.1111/j.1365-294X.2009.04243.x

Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M., & Ravel, J. (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*, *2*(1), 1-7.

Fleskes, J. P., Ramey, A. M., Reeves, A. B., & Yee, J. L. (2017). Body mass, wing length, and condition of wintering ducks relative to hematozoa infection. *Journal of Fish and Wildlife Management*, *8*(1), 89-100.

Forsberg, L. A., Dannewitz, J., Petersson, E., & Grahn, M. (2007). Influence of genetic dissimilarity in the reproductive success and mate choice of brown trout–females fishing for optimal MHC dissimilarity. *Journal of evolutionary biology*, *20*(5), 1859-1869.

Gibb, J. (1950). The breeding biology of the Great and Blue titmice. *Ibis*, *92*(4), 507-539.

Grieves, L. A., Gloor, G. B., Bernards, M. A., & MacDougall-Shackleton, E. A. (2019). Songbirds show odour-based discrimination of similarity and diversity at the major histocompatibility complex. *Animal behaviour*, *158*, 131-138. https://doi.org/https://doi.org/10.1016/j.anbehav.2019.10.005

Hamilton, W. D., & Zuk, M. (1982). Heritable true fitness and bright birds: a role for parasites?. *Science*, *218*(4570), 384-387.

Hellgren, O., Waldenström, J., & Bensch, S. (2004). A new PCR assay for simultaneous studies of *Leucocytozoon*, *Plasmodium*, and *Haemoproteus* from avian blood. *Journal of parasitology*, *90*(4), 797-802. https://doi.org/10.1645/GE-184R1

Hunt, S., Cuthill, I. C., Bennett, A. T., & Griffiths, R. (1999). Preferences for ultraviolet partners in the blue tit. *Animal Behaviour*, *58*(4), 809-815.

Janeway Jr, C. A., Travers, P., Walport, M., & Shlomchik, M. J. (2001). The major histocompatibility complex and its functions. In *Immunobiology: The Immune System in Health and Disease. 5th edition*. Garland Science. https://www.ncbi.nlm.nih.gov/books/NBK27156/

Johannessen, L. E., Slagsvold, T., Hansen, B. T., & Lifjeld, J. T. (2005). Manipulation of male quality in wild tits: effects on paternity loss. *Behavioral Ecology*, *16*(4), 747-754.

Jones, M. P., Pierce Jr, K. E., & Ward, D. (2007). Avian vision: a review of form and function with special consideration to birds of prey. *Journal of Exotic Pet Medicine*, *16*(2), 69-87.

Kamiya, T., O'dwyer, K., Westerdahl, H., Senior, A., & Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: The role of diversity and dissimilarity. *Molecular ecology*, *23*(21), 5151-5163. https://doi.org/10.1111/mec.12934

Kaufman, J. (2018). Unfinished business: evolution of the MHC and the adaptive immune system of jawed vertebrates. *Annual review of immunology*, *36*, 383-409. https://doi.org/10.1146/annurev-immunol-051116-052450

Klein, J., & Figueroa, F. E. L. I. P. E. (1986). Evolution of the major histocompatibility complex. *Critical reviews in immunology*, *6*(4), 295-386.

Knowles, S. C. L., Palinauskas, V., & Sheldon, B. C. (2010). Chronic malaria infections increase family inequalities and reduce parental fitness: experimental evidence from a wild bird population. *Journal of evolutionary biology*, *23*(3), 557-569. https://doi.org/https://doi.org/10.1111/j.1420-9101.2009.01920.x

Leclaire, S., Strandh, M., Dell'Ariccia, G., Gabirot, M., Westerdahl, H., & Bonadonna, F. (2019). Plumage microbiota covaries with the major histocompatibility complex in blue petrels. *Molecular ecology*, *28*(4), 833-846. https://doi.org/doi:10.1111/mec.14993

Leclaire, S., Strandh, M., Mardon, J., Westerdahl, H., & Bonadonna, F. (2017). Odour-based discrimination of similarity at the major histocompatibility complex in birds. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1846), 20162466. https://doi.org/10.1098/rspb.2016.2466

Lenz, T. L., Deutsch, A. J., Han, B., Hu, X., Okada, Y., Eyre, S., Knapp, M., Zhernakova, A., Huizinga, T. W., Abecasis, G., Becker, J., Boeckxstaens, G. E., Chen, W. M., Franke, A., Gladman, D. D., Gockel, I., Gutierrez-Achury, J., Martin, J., Nair, R. P., Nöthen, M. M., Onengut-Gumuscu, S., Rahman, P., Rantapää-Dahlqvist, S., Stuart, P. E., Tsoi, L. C., van Heel, D. A., Worthington, J., Wouters, M. M., Klareskog, L., Elder, J. T., Gregersen, P. K., Schumacher, J., Rich, S. S., Wijmenga, C., Sunyaev, S. R., de Bakker, P. I., & Raychaudhuri, S. (2015). Widespread non-additive and interaction effects within HLA loci modulate the risk of autoimmune diseases. *Nature genetics*, *47*(9), 1085-1090. https://doi.org/10.1038/ng.3379

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic acids research*, *49*(W1), W293-W296.

Lifjeld, J., & Krokene, C. (2000). Variation in the frequency of extra-pair paternity in birds: a comparison of an island and a mainland population of blue tits. *Behaviour*, *137*(10), 1317-1330.

Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., & Pallen, M. J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nature biotechnology*, *30*(5), 434-439.

Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, *27*(21), 2957-2963. https://doi.org/10.1093/bioinformatics/btr507

Mays, H. L., & Hill, G. E. (2004). Choosing mates: good genes versus genes that are a good fit. *Trends in ecology & evolution*, *19*(10), 554-559.

Meixell, B. W., Arnold, T. W., Lindberg, M. S., Smith, M. M., Runstadler, J. A., & Ramey, A. M. (2016). Detection, prevalence, and transmission of avian hematozoa in waterfowl at the Arctic/sub-Arctic interface: co-infections, viral interactions, and sources of variation. *Parasites & Vectors*, *9*, 1-18.

Milinski, M. (2006). The major histocompatibility complex, sexual selection, and mate choice. *Annu. Rev. Ecol. Evol. Syst.*, *37*, 159-186. https://doi.org/10.1146/annurev.ecolsys.37.091305.110242

Moore, B. A., Pita, D., Tyrrell, L. P., & Fernández-Juricic, E. (2015). Vision in avian emberizid foragers: maximizing both binocular vision and fronto-lateral visual acuity. *The Journal of Experimental Biology*, *218*(9), 1347-1358.

Møller, A. P., Saino, N., Taramino, G., Ferrario, S., & Galeotti, P. (1998). Paternity and multiple signaling: effects of a secondary sexual character and song on paternity in the barn swallow. *The American Naturalist*, *151*(3), 236-242.

Neefjes, J., Jongsma, M. L., Paul, P., & Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature reviews immunology*, *11*(12), 823-836. https://doi.org/10.1038/nri3084

Nowak, M. A., Tarczy-Hornoch, K., & Austyn, J. M. (1992). The optimal number of major histocompatibility complex molecules in an individual. *Proceedings of the National Academy of Sciences*, *89*(22), 10896-10899. https://doi.org/10.1073/pnas.89.22.10896

O'Connor, E. A., Strandh, M., Hasselquist, D., Nilsson, J. Å., & Westerdahl, H. (2016). The evolution of highly variable immunity genes across a passerine bird radiation. *Molecular Ecology*, *25*(4), 977-989. https://doi.org/10.1111/mec.13530

Olsson, M., Madsen, T., Wapstra, E., Silverin, B., Ujvari, B., & Wittzell, H. (2005). MHC, health, color, and reproductive success in sand lizards. *Behavioral Ecology and Sociobiology*, *58*, 289-294.

Podmokła, E., Dubiec, A., Drobniak, S. M., Arct, A., Gustafsson, L., & Cichoń, M. (2014). Avian malaria is associated with increased reproductive investment in the blue tit. *Journal of Avian Biology*, *45*(3), 219-224. https://doi.org/10.1111/j.1600-048X.2013.00284.x

Puente, J. M. D. L., Merino, S., Tomás, G., Moreno, J., Morales, J., Lobato, E., ... & Belda, E. J. (2010). The blood parasite *Haemoproteus* reduces survival in a wild bird: a medication experiment. *Biology letters*, *6*(5), 663-665. https://doi.org/10.1098/rsbl.2010.0046

Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., Bertoni, A., Swerdlow, H. P., & Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC genomics*, *13*(1), 1-13.

Rajchard, J. (2009). Ultraviolet (UV) light perception by birds: a review. *Veterinární medicína*, *54*(8), 351-359.

Rekdal, S. L. (2020). Major histocompatibility complex and mate choice in two passerine birds. [University of Oslo]. Oslo.

Rekdal, S. L., Anmarkrud, J. A., Johnsen, A., & Lifjeld, J. T. (2018). Genotyping strategy matters when analyzing hypervariable major histocompatibility complex-Experience from a passerine bird. *Ecology and Evolution*, *8*(3), 1680-1692. https://doi.org/10.1002/ece3.3757

Rekdal, S. L., Anmarkrud, J. A., Lifjeld, J. T., & Johnsen, A. (2019). Extra-pair mating in a passerine bird with highly duplicated major histocompatibility complex class II: Preference for the golden mean. *Molecular Ecology*, *28*(23), 5133-5144. https://doi.org/10.1111/mec.15273

Rekdal, S. L., Anmarkrud, J. A., Lifjeld, J. T., & Johnsen, A. (In review). Do female bluethroats without extra-pair offspring have more MHC-compatible social mates?

Richardson, D. S., Komdeur, J., Burke, T., & Von Schantz, T. (2005). MHC-based patterns of social and extra-pair mate choice in the Seychelles warbler. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1564), 759-767.

Rätti, O., Dufva, R., & Alatalo, R. V. (1993). Blood parasites and male fitness in the pied flycatcher. *Oecologia*, *96*, 410-414.

Sauermann, U., Nürnberg, P., Bercovitch, F., Berard, J., Trefilov, A., Widdig, A., Kessler, M., Schmidtke, J., & Krawczak, M. (2001). Increased reproductive success of MHC class II heterozygous males among free-ranging rhesus macaques. *Human genetics*, *108*, 249-254.

Schubert, N., Nichols, H. J., & Winternitz, J. C. (2021). How can the MHC mediate social odor via the microbiota community? A deep dive into mechanisms. *Behavioral Ecology*, *32*(3), 359-373. https://doi.org/10.1093/beheco/arab004

Schut, E., Aguilar, J. R. D., Merino, S., Magrath, M. J., Komdeur, J., & Westerdahl, H. (2011). Characterization of MHC-I in the blue tit (*Cyanistes caeruleus*) reveals low levels of genetic diversity and trans-population evolution across European populations. *Immunogenetics*, *63*, 531-542. https://doi.org/10.1007/s00251-011-0532-x

Sepil, I., Lachish, S., & Sheldon, B. C. (2013). MHC-linked survival and lifetime reproductive success in a wild population of great tits. *Molecular ecology*, *22*(2), 384-396. https://doi.org/10.1111/mec.12123

Slade, J. W. G., Watson, M. J., & MacDougall-Shackleton, E. A. (2017). Birdsong signals individual diversity at the major histocompatibility complex. *Biology Letters*, *13*(11), 20170430. https://doi.org/10.1098/rsbl.2017.0430

Slagsvold, T., Kleiven, K. W., Eriksen, A., & Johannessen, L. E. (2013). Vertical and horizontal transmission of nest site preferences in titmice. *Animal Behaviour*, *85*(2), 323-328.

Strandh, M., Lannefors, M., Bonadonna, F., & Westerdahl, H. (2011). Characterization of MHC class I and II genes in a subantarctic seabird, the blue petrel, Halobaena caerulea (Procellariiformes). *Immunogenetics*, *63*, 653-666. https://doi.org/10.1007/s00251-011-0534-8

Strandh, M., Westerdahl, H., Pontarp, M., Canbäck, B., Dubois, M.-P., Miquel, C., Taberlet, P., & Bonadonna, F. (2012). Major histocompatibility complex class II compatibility, but not class I, predicts mate choice in a bird with highly developed olfaction. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1746), 4457-4463. https://doi.org/10.1098/rspb.2012.1562

Stuglik, M. T., Radwan, J., & Babik, W. (2011). jMHC: Software assistant for multilocus genotyping of gene families using next-generation amplicon sequencing. *Molecular ecology resources*, *11*(4), 739-742. https://doi.org/10.1111/j.1755-0998.2011.02997.x

Svensson, L. (1992). *Identification guide to European passerines*. Lars Svensson.

Tedore, C., & Nilsson, D. E. (2021). Ultraviolet vision aids the detection of nutrient-dense non-signaling plant foods. *Vision Research*, *183*, 16-29.

Thoss, M., Ilmonen, P., Musolf, K., & Penn, D. J. (2011). Major histocompatibility complex heterozygosity enhances reproductive success. *Molecular ecology*, *20*(7), 1546-1557.

Tomás, G., Merino, S., Moreno, J., & Morales, J. (2007). Consequences of nest reuse for parasite burden and female health and condition in blue tits, *Cyanistes caeruleus*. *Animal Behaviour*, *73*(5), 805-814.

Tregenza, T., & Wedell, N. (2000). Genetic compatibility, mate choice and patterns of parentage: invited review. *Molecular Ecology*, *9*(8), 1013-1027.

Valkiūnas, G., & Iezhova, T. A. (2017). Exo-erythrocytic development of avian malaria and related haemosporidian parasites. *Malaria Journal*, *16*, 1-24. https://doi.org/10.1186/s12936-017-1746-7

Wettere, A. J. V. (2022). *Haemoproteus Infection in Poultry*. MSD MANUAL. Retrieved 18.01.2023 from https://www.msdvetmanual.com/poultry/bloodborne-organisms/haemoproteus-infection-in-poultry

Whittingham, L. A., Freeman-Gallant, C. R., Taff, C. C., & Dunn, P. O. (2015). Different ornaments signal male health and MHC variation in two populations of a warbler. *Molecular Ecology*, *24*(7), 1584-1595. https://doi.org/10.1111/mec.13130

Winternitz, J., Abbate, J. L., Huchard, E., Havlíček, J., & Garamszegi, L. (2017). Patterns of MHC-dependent mate selection in humans and nonhuman primates: a meta-analysis. *Molecular ecology*, *26*(2), 668-688. https://doi.org/10.1111/mec.13920

Wutzler, R., Foerster, K., & Kempenaers, B. (2012). MHC class I variation in a natural blue tit population (*Cyanistes caeruleus*). *Genetica*, *140*, 349-364. https://doi.org/10.1007/s10709-012-9679-0

Zagalska-Neubauer, M., Babik, W., Stuglik, M., Gustafsson, L., Cichoń, M., & Radwan, J. (2010). 454 sequencing reveals extreme complexity of the class II Major Histocompatibility Complex in the collared flycatcher. *BMC evolutionary biology*, *10*(1), 1-15.

Zekarias, B., Ter Huurne, A. A., Landman, W., Rebel, J., Pol, J., & Gruys, E. (2002). Immunological basis of differences in disease resistance in the chicken. *Veterinary Research*, *33*(2), 109-125. https://doi.org/10.1051/vetres:2002001

Ziegler, A., Kentenich, H., & Uchanska-Ziegler, B. (2005). Female choice and the MHC. *TRENDS in Immunology*, *26*(9), 496-502. https://doi.org/10.1016/j.it.2005.07.003

7.0 Appendix

Appendix 1

Supplementary Table 1: PCR setup and thermal profile for MHCIe3.

Temperature	Duration	Cycles
98	30 s	1
98	10 s	
65	20 s	25
72	15 s	
72	2 min	1

Supplementary Table 2: PCR setup and thermal profile for blood parasites Haemoproteus, Plasmodium and Leucocytozoon.

Temperature	Duration	Cycles
95	15 min	1
94	30 s	
59	90 s	35
72	30 s	
72	10 min	1

Appendix 2

Supplementary Table 3: A complete list of all primers used in this work. The F and R after the hyphen refers to forward or reverse, respectively. The number behind F/R reflects the id of the 12 bp barcode index motif.

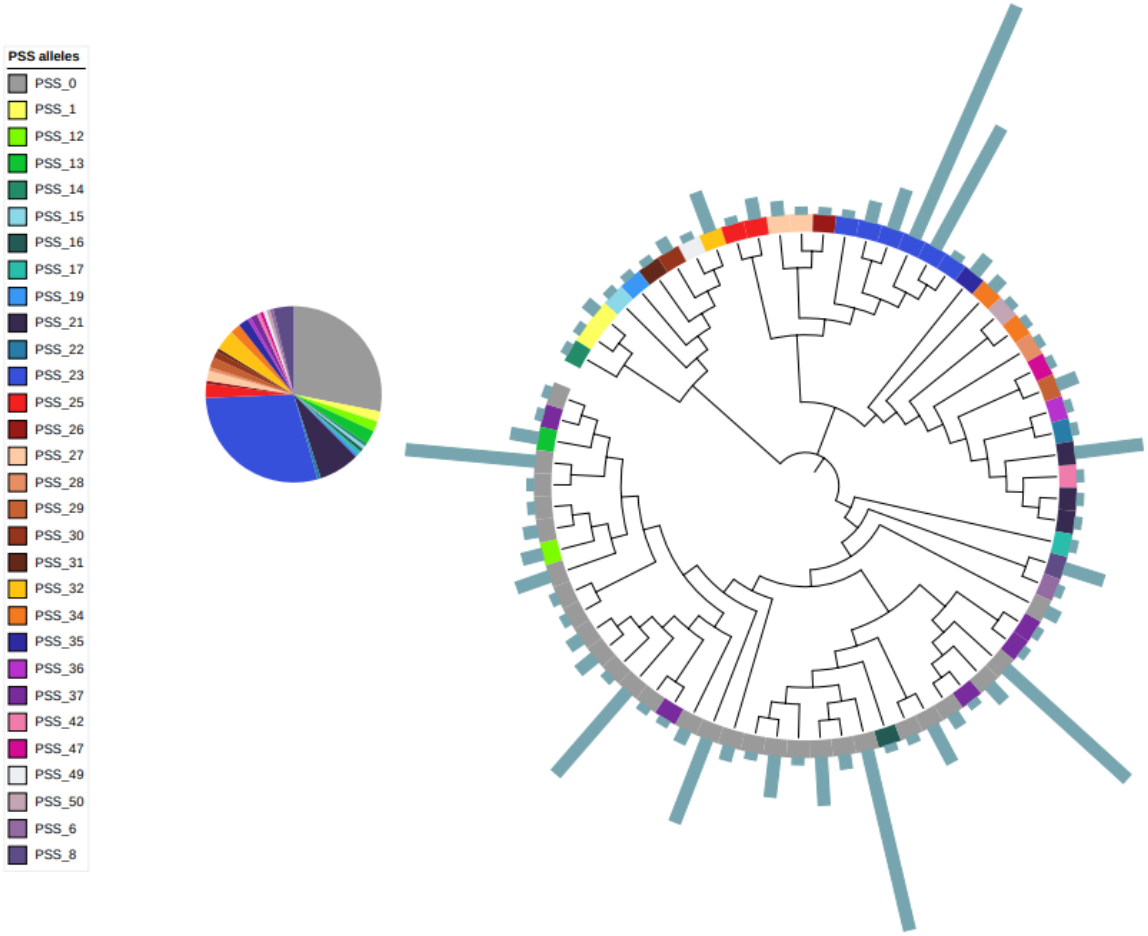
Description	Name	Sequence	Reference
MHC primer 2001	Hnalla-F1	CCTAAACTACGGTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F2	TGCAGATCCAACTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F3	CCATCACATAGGTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F4	GTGGTATGGGAGTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F5	ACTTTAAGGGTGTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F6	GAGCAACATCCTTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F7	TGTTGCGTTTCTTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F8	ATGTCCGACCAATCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F9	AGGTACGCAATTTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F10	ACAGCCACCCATTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F11	TGTCTCGCAAGCTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F12	GAGGAGTAAAGCTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F13	GTTACGTGGTTGTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F14	TACCGCCTCGGATCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F15	CGTAAGATGCCTTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F16	TACCGGCTTGCATCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F17	ATCTAGTGGCAATCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F18	CCAGGGACTTCTTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F19	CACCTTACCTTATCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F20	ATAGTTAGGGCTTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F21	GCACTTCATTTCTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F22	TTAACTGGAAGCTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	Hnalla-F23	CGCGGTTACTAATCCCCACAGGTCTCCACAC	This study

MHC primer 2001	Hnalla-F24	GAGACTATATGCTCCCCACAGGTCTCCACAC	This study
MHC primer 2001	HN46-R1	CCTAAACTACGGATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R2	TGCAGATCCAACATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R3	CCATCACATAGGATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R4	GTGGTATGGGAGATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R5	ACTTTAAGGGTGATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R6	GAGCAACATCCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R7	TGTTGCGTTTCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R8	ATGTCCGACCAAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R9	AGGTACGCAATTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R10	ACAGCCACCCATATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R11	TGTCTCGCAAGCATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R12	GAGGAGTAAAGCATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R13	GTTACGTGGTTGATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R14	TACCGCCTCGGAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R15	CGTAAGATGCCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R16	TACCGGCTTGCAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R17	ATCTAGTGGCAAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R18	CCAGGGACTTCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R19	CACCTTACCTTAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R20	ATAGTTAGGGCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R21	GCACTTCATTTCATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R22	TTAACTGGAAGCATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R23	CGCGGTTACTAAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2001	HN46-R24	GAGACTATATGCATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F1	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCTAAACTACGGTCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTAAACTACGGATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F2	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCAGATCCAACTCCCCACAGGTCTCCACAC	This study

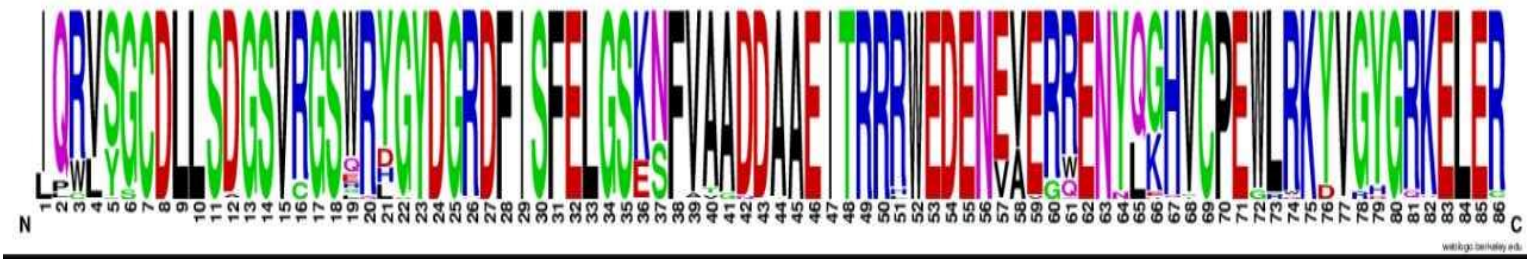
MHC primer 2021	HN46-R2	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGCAGATCCAACATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F3	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCATCACATAGGTCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R3	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCATCACATAGGATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F4	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTGGTATGGGAGTTCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R4	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGTGGTATGGGAGAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F5	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACTTTAAGGGTGTTCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACTTTAAGGGTGAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F6	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGAGCAACATCCTTTCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R6	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGCAACATCCTAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F7	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGTTGCGTTTCTGTTCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R7	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGTTGCGTTTCTTCATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F8	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATGTCCGACCAAGTTCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R8	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATGTCCGACCAATCATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F9	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGGTACGCAATTGTTCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R9	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGGTACGCAATTTCATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	HN46-R10	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACAGCCACCCATCTAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F10	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACAGCCACCCATCGATCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R11	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGTCTCGCAAGCCTAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F11	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGTCTCGCAAGCCGATCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R12	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGGAGTAAAGCCTAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F12	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGAGGAGTAAAGCCGATCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R13	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGTTACGTGGTTGGATAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F13	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTTACGTGGTTGATGATCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R14	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTACCGCCTCGGAGATAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F14	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTACCGCCTCGGAATGATCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R15	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCGTAAGATGCCTGATAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	Hnalla-F15	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGTAAGATGCCTATGATCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTACCGGCTTGCAACTCAATCCCAAATTCCCACCCACCTT	This study

MHC primer 2021	Hnalla-F16	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTACCGGCTTGCATGCGATCCCCACAGGTCTCCACAC	This study
MHC primer 2021	HN46-R17	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATCTAGTGGCAAACTCAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	HN46-R18	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCAGGGACTTCTACTCAATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	HN46-R19	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCACCTTACCTTATTCTCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	HN46-R20	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTATAGTTAGGGCTTTCTCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	HN46-R21	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGCACTTCATTTCTTCTCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	HN46-R22	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTTAACTGGAAGCCACTTCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	HN46-R23	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCGCGGTTACTAACACTTCTATCCCAAATTCCCACCCACCTT	This study
MHC primer 2021	HN46-R24	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGACTATATGCCACTTCTATCCCAAATTCCCACCCACCTT	This study
Blood parasite	HMF	ATTGGATGTCAATTACCACAATC	Ciloglu et al. 2019.
Blood parasite	LMF	TGGAACAATAATTGSATTATTTACAYT	Ciloglu et al. 2019.
Blood parasite	PMF	CCTCACGAGTCGATCAGG	Ciloglu et al. 2019.
Blood parasite	HMR	GGGAAGTTTATCCAGGAAGTT	Ciloglu et al. 2019.
Blood parasite	LMR	AACATATCATATTCCATCCATTTAGATTA	Ciloglu et al. 2019.
Blood parasite	PMR	GGAAACCGGCGCTAC	Ciloglu et al. 2019.

Appendix 3



Supplementary Figure 1: Interactive tree of life created with the online tool ITOL (Letunic & Bork, 2021) showing the overview of nucleotide alleles and how they are grouped in PSS alleles.



Supplementary Figure 2: A sequence logo where we can see on which amino acid seats the biggest diversity lies. The PSS alleles were solely based on the functionally important sites 1, 3, 5, 19, 21, 57, 60, 61 (Balakrishnan et al., 2010; Rekdal, 2020).

Appendix 4

Supplementary Table 4: Tests of association between MCHI variation and morphological variables, separated by sex.

MHCI variable	Morphological variable	Sex	Test type	n	r_s	p-value
Number of MHCI alleles	Tarsus length	male	Spearman correlation	28	0.04	0.83
Number of MHCI alleles	Wing length	male	Spearman correlation	28	0.26	0.19
Number of MHCI alleles	Mass	male	Spearman correlation	29	0.06	0.76
Number of MHCI alleles	Tarsus length	female	Spearman correlation	26	-0.05	0.82
Number of MHCI alleles	Wing length	female	Spearman correlation	26	0.34	0.09
Number of MHCI alleles	Mass	female	Spearman correlation	28	0.23	0.23
Distance to mean	Tarsus length	male	Spearman correlation	28	0.15	0.44
Distance to mean	Wing length	male	Spearman correlation	28	0.13	0.52
Distance to mean	Mass	male	Spearman correlation	29	0.32	0.09
Distance to mean	Tarsus length	female	Spearman correlation	26	-0.15	0.48
Distance to mean	Wing length	female	Spearman correlation	26	-0.23	0.27
Distance to mean	Mass	female	Spearman correlation	28	-0.23	0.25

Supplementary Table 5: Tests of association between MHCI variation and age, separated by sex.

Grouping variable	Dependent variable	Sex	Test type	n SY	n ASY	W	p-value
Age	Number of MHCI alleles	Male	Wilcoxon rank sum test	10	21	112	0.78
Age	Number of MHCI alleles	Female	Wilcoxon rank sum test	12	18	117.5	0.69
Age	Distance to mean	Male	Wilcoxon rank sum test	10	21	58	0.046
Age	Distance to mean	Female	Wilcoxon rank sum test	12	18	123.5	0.51

Supplementary Table 6: Tests of association between morphology variables and blood parasite prevalence, separated by sex.

Grouping variable	Dependent variable	Sex	Test type	n infected	n not infected	Test statistic	p-value
Haemoproteus	Tarsus length	Male	t-test	15	24	t = -1.03	0.31
Haemoproteus	Wing length	Male	t-test	15	24	t = -1.59	0.12
Haemoproteus	Mass	Male	Wilcox rank sum test	14	25	W = 15	0.39
Haemoproteus	Tarsus length	Female	t-test	6	21	t = -0.03	0.98
Haemoproteus	Wing length	Female	t-test	6	21	t = -1.13	0.28
Haemoproteus	Mass	Female	Wilcox rank sum test	6	19	W = 55	0.47
Plasmodium	Tarsus length	Male	t-test	4	26	t = 0.44	0.68
Plasmodium	Wing length	Male	t-test	4	26	t = 2.53	0.04
Plasmodium	Mass	Male	Wilcox rank sum test	5	25	W = 74	0.54
Plasmodium	Tarsus length	Female	t-test	7	17	t = 0.34	0.74
Plasmodium	Wing length	Female	t-test	7	17	t = -1.01	0.33
Plasmodium	Mass	Female	Wilcox rank sum test	9	17	W = 77.5	0.98
Leucocytozoon	Tarsus length	Male	t-test	25	5	t = 0.72	0.50
Leucocytozoon	Wing length	Male	t-test	25	5	t = -0.82	0.43
Leucocytozoon	Mass	Male	Wilcox rank sum test	26	4	W = 37	0.37
Leucocytozoon	Tarsus length	Female	t-test	21	3	t = 1.74	0.18
Leucocytozoon	Wing length	Female	t-test	21	3	t = -0.49	0.66
Leucocytozoon	Mass	Female	Wilcox rank sum test	22	4	W = 41	0.86

Supplementary Table 7: Table showing the relationship between age and blood parasite infection rate, separated by sex.

Parasite genus	Sex	Test type	n infected SY	n not infected SY	n infected ASY	n not infected ASY	p-value
Haemoproteus	Male	Fisher's exact test	5	5	11	21	0.46
Haemoproteus	Female	Fisher's exact test	1	11	5	14	0.36
Plasmodium	Male	Fisher's exact test	0	3	5	24	1.00
Plasmodium	Female	Fisher's exact test	1	7	8	10	0.19
Leucocytozoon	Male	Fisher's exact test	3	0	24	5	1
Leucocytozoon	Female	Fisher's exact test	6	2	16	2	0.56

Appendix 5

Supplementary Table 8a: Complete list of all adult individuals from the 2001 and 2021 samples, and their variables.

ID	Sex	Age	Year	Clutch size	Number fledged	Egg laying date	Tarsus	Wing	Mass	Number MHC alleles	Distance to mean
1809	Male	2	2001	9	NA	21	18.9	66	10.1	4	0.77
1810	Female	2	2001	9	Na	21	NA	NA	NA	6	1.19
1821	Male	2	2001	9	NA	20	20	68.5	11.3	7	1.55
1822	Female	NA	2001	9	NA	20	NA	NA	NA	6	1.19
1870	Male	3	2001	8	NA	21	20	71	11.4	8	1.85
1871	Female	NA	2001	8	NA	21	NA	NA	NA	4	0.77
1890	Male	2	2001	8	NA	22	20	66.5	11.1	4	0.77
1891	Female	NA	2001	8	NA	22	NA	NA	NA	5	0.64
1897	Male	2	2001	9	NA	21	20	66.5	12.4	4	0.77
1898	Female	2	2001	9	NA	21	19.1	64	10.9	5	0.64
1994	Male	2	2001	9	NA	19	20	68	11.1	6	1.19
1997	Female	NA	2001	9	NA	19	NA	NA	NA	4	0.77
2033	Male	3	2001	8	NA	19	19.2	68	11.6	3	1.26
2034	Female	NA	2001	8	NA	19	NA	NA	NA	5	0.64
2057	Male	2	2001	9	NA	21	NA	NA	NA	4	0.77
2058	Female	3	2001	9	NA	21	19	64	10.2	3	1.26
2071	Male	2	2001	7	NA	18	20	66	11.4	4	0.77
2073	Female	NA	2001	7	NA	18	NA	NA	NA	4	0.77
2080	Male	NA	2001	9	NA	11	NA	NA	NA	3	1.26
2081	Female	2	2001	9	NA	22	19.5	64	10.6	3	1.26
2102	Male	3	2001	8	NA	21	19	69.5	10.7	4	0.77
2103	Female	2	2001	8	NA	21	NA	NA	NA	5	0.64
107040	Male	3	2021	7	0	10	19	67	12	2	1.22

107041	Male	3	2021	6	4	9	18.6	69	12.5	4	0.72
107042	Male	3	2021	8	7	11	20.1	70.5	11	4	0.72
107043	Male	3	2021	5	4	12	18.5	68	11.5	4	0.72
107044	Male	3	2021	6	5	9	20	67	12	2	1.22
107045	Male	2	2021	7	0	8	18.4	62	10.5	3	0.69
107046	Male	3	2021	0	0	NA	20.2	70	11.5	NA	NA
107047	Male	3	2021	8	0	10	20.4	69	NA	NA	NA
107048	Male	3	2021	8	0	9	19	70	11.5	NA	NA
107049	Male	3	2021	8	4	8	19	67.8	11	NA	NA
107050	Male	3	2021	6	0	9	19.8	67	13	4	0.72
107051	Male	3	2021	7	6	9	18.6	68.5	11.7	NA	NA
107052	Male	3	2021	0	0	NA	18.4	68	11.9	NA	NA
107053	Male	3	2021	5	5	23	18.7	68	11.2	NA	NA
107054	Male	3	2021	7	6	9	18.8	70	12.5	5	1.23
107055	Male	3	2021	8	3	19	18.6	67.5	11	NA	NA
107056	Male	3	2021	8	5	13	18.9	66	10.4	3	0.69
107057	Male	3	2021	4	4	13	19.2	71	12	6	1.59
107058	Male	3	2021	5	0	13	19.1	69	11.6	NA	NA
107059	Male	3	2021	8	7	1	19.3	67	11.8	7	1.88
107060	Male	3	2021	1	0	6	18.6	69	11	NA	NA
107061	Male	2	2021	1	0	19	19.1	66	10.9	6	1.59
107062	Male	3	2021	8	4	8	19.5	67	11	NA	NA
107063	Male	3	2021	9	3	8	19.8	70	NA	3	0.69
107064	Male	3	2021	0	0	NA	19.7	67.5	10.4	3	0.69
107065	Male	3	2021	7	6	13	18.9	68	10.7	4	0.72
107066	Male	3	2021	8	0	4	19.9	72	12	3	0.69
107067	Male	3	2021	9	7	15	19.2	68	11.3	3	0.69

107068	Female	3	2021	5	4	12	18.6	61	12.5	2	1.22
107069	Female	3	2021	5	0	13	18	68	13	3	0.69
107070	Female	2	2021	9	7	15	18.9	65	12.5	2	1.22
107071	Female	3	2021	8	0	4	18.8	66	12.6	NA	NA
107072	Female	3	2021	NA	NA	NA	19.2	66	13.9	5	1.23
107073	Female	3	2021	8	5	13	18.4	65	12.6	3	0.69
107074	Female	3	2021	8	0	10	17.2	64	11.4	3	0.69
107075	Female	2	2021	6	0	9	18.7	63	10.6	4	0.72
107076	Female	2	2021	8	4	8	18	63	9.9	2	1.22
107077	Female	3	2021	7	6	9	19.3	64	10.9	3	0.69
107078	Female	3	2021	4	4	13	19.1	65	11.5	3	0.69
107079	Female	2	2021	9	3	8	18.4	62	10.2	2	1.22
107080	Female	3	2021	8	7	1	19.2	65	11.4	3	0.69
107088	Female	2	2021	7	6	13	19.3	64	12	3	0.69
107089	Female	3	2021	8	4	8	18.6	64	11.5	3	0.69
107090	Female	3	2021	8	7	11	19	64.5	10.9	3	0.69
107091	Female	2	2021	6	5	9	18.5	64	11	4	0.72
107092	Female	2	2021	8	3	19	18.6	62	12.5	3	0.69
107093	Female	3	2021	7	6	9	19.2	66	12.4	4	0.72
107094	Female	3	2021	5	5	23	18.8	66	12.8	3	0.69
107131	Female	3	2021	8	5	8	NA	NA	10.2	1	1.57
107132	Male	3	2021	11	10	8	NA	NA	11	4	0.72
107137	Female	3	2021	11	10	8	17.6	64	10	4	0.72
107138	Female	3	2021	8	7	11	18.3	66	10.8	4	0.72
107139	Male	3	2021	8	5	8	19.6	67	12.1	3	0.69
107145	Male	3	2021	8	7	11	20	66	11.1	7	1.88
107185	Female	2	2021	8	8	20	NA	NA	11.5	4	0.72

107186	Male	2	2021	8	8	20	NA	NA	10.4	3	0.69
107195	Female	3	2021	10	7	15	19	65	10.2	3	0.69
107203	Female	3	2021	NA	NA	NA	19.3	66	10.6	3	0.69

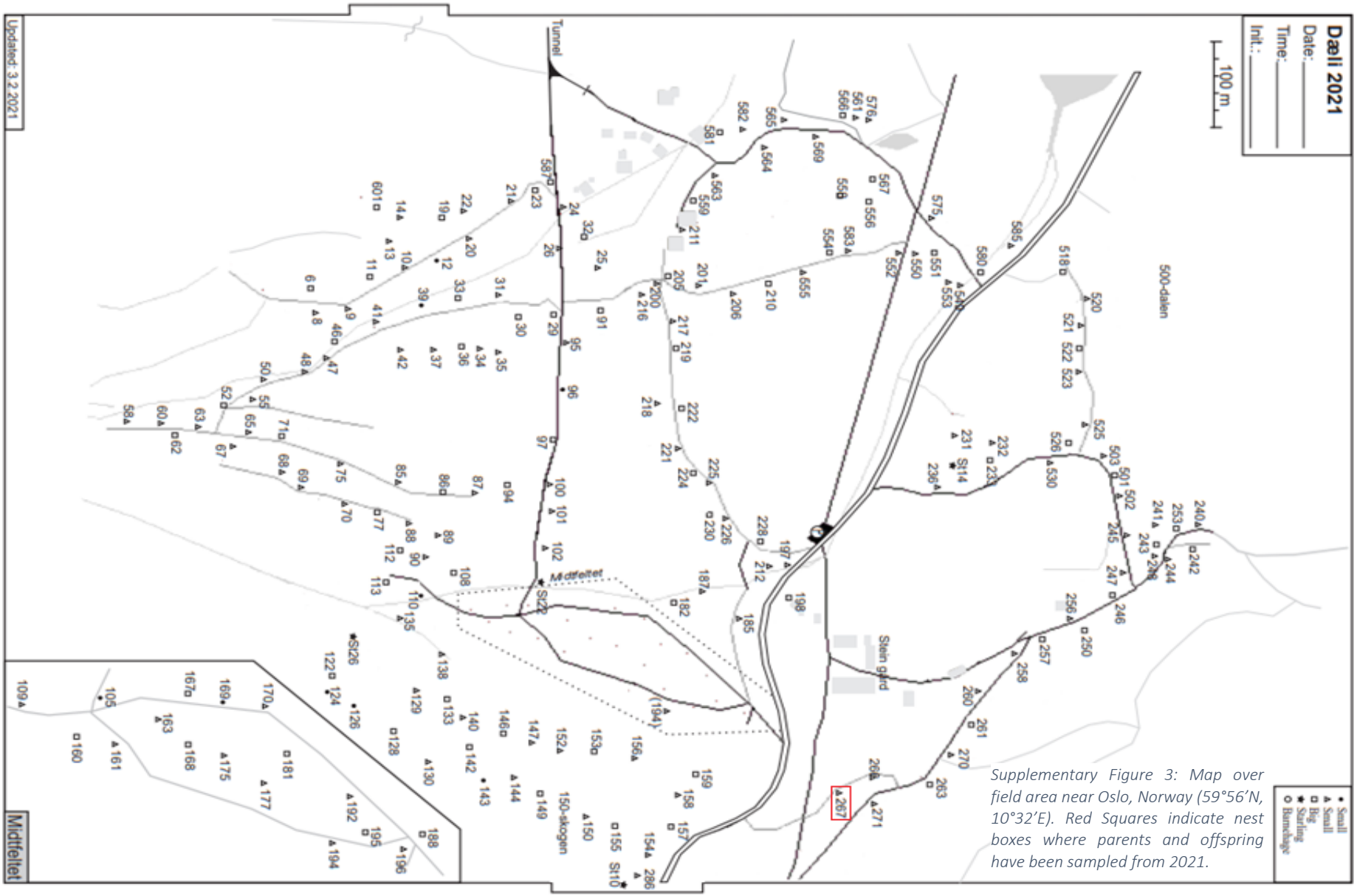
Supplementary Table 8b: Complete list of all adult individuals from the 2001 and 2021 samples, and their variables.

ID	*Haemoproteus*	*Plasmodium*	*Leucocytozoon*	Number parasites	IWYWYVGR
1809	0	0	1	1	0
1810	0	0	1	1	0
1821	1	0	1	2	1
1822	1	0	1	2	1
1870	1	0	1	2	1
1871	0	0	1	1	1
1890	1	0	1	2	0
1891	1	0	1	2	1
1897	1	0	1	2	1
1898	1	0	1	2	1
1994	1	0	1	2	1
1997	0	0	1	1	0
2033	0	0	1	1	0
2034	0	0	1	1	0
2057	1	0	1	2	0
2058	0	0	1	1	0
2071	0	0	1	1	1
2073	0	0	1	1	1
2080	0	0	1	1	0
2081	0	0	1	1	0
2102	0	0	1	1	0
2103	0	0	1	1	0
107040	0	1	1	2	0
107041	0	0	1	1	0
107042	0	0	1	1	0
107043	1	1	1	3	0
107044	0	0	1	1	0
107045	0	0	1	1	0
107046	1	0	1	2	0
107047	0	0	0	0	0
107048	1	0	1	2	0
107049	0	0	0	0	0
107050	0	0	1	1	0
107051	0	0	1	1	0
107052	0	0	1	1	0

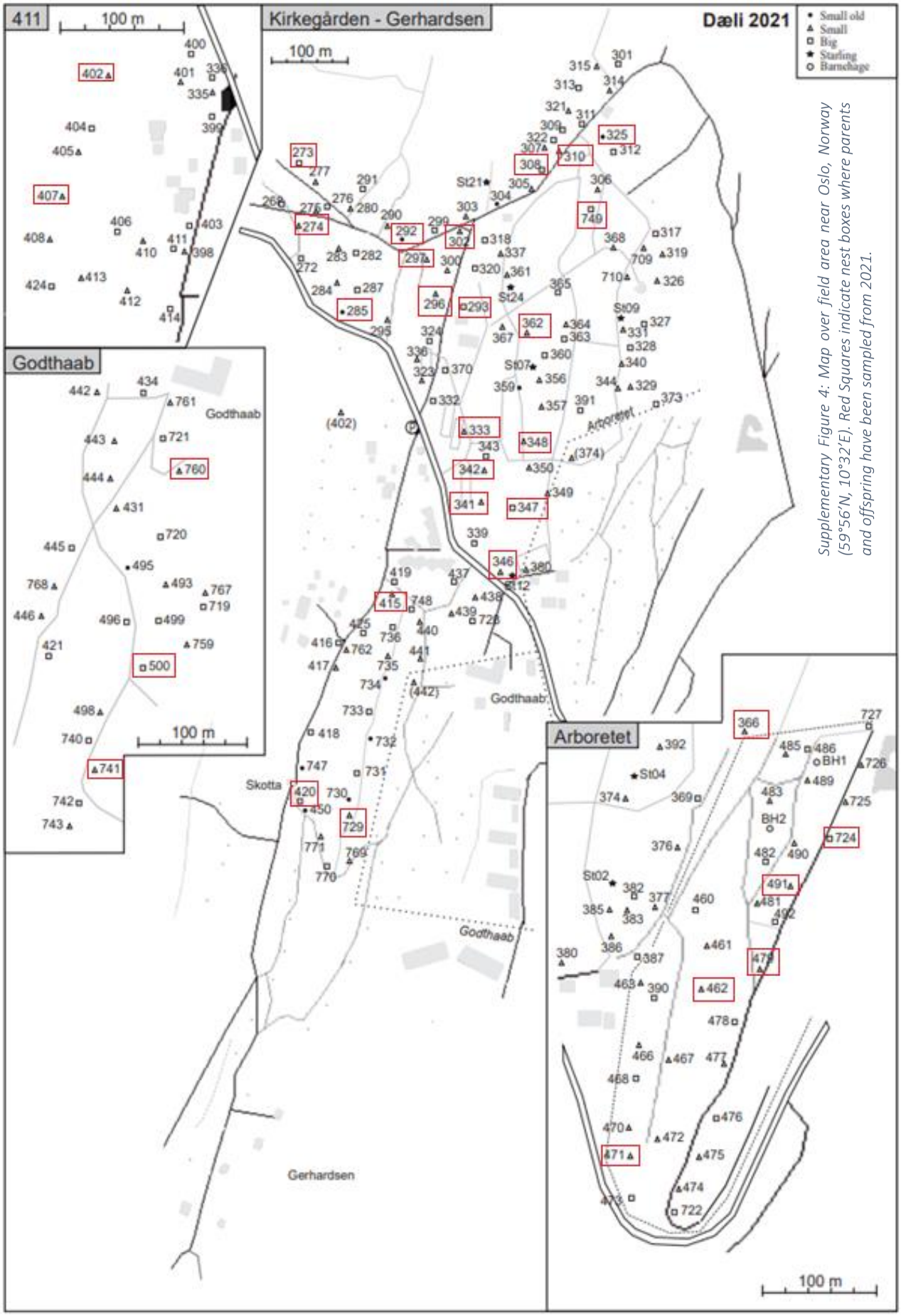
107053	1	0	1	2	0
107054	1	0	1	2	0
107055	0	0	0	0	0
107056	0	1	1	2	0
107057	0	0	1	1	0
107058	1	0	1	2	0
107059	0	0	1	1	0
107060	1	0	1	2	0
107061	0	0	1	1	0
107062	0	0	1	1	0
107063	1	0	1	2	0
107064	0	0	1	1	0
107065	0	0	1	1	0
107066	0	0	1	1	0
107067	1	0	0	1	0
107068	0	0	1	1	0
107069	0	1	1	2	0
107070	0	0	1	1	0
107071	1	0	1	2	0
107072	0	0	1	1	0
107073	0	1	1	2	0
107074	1	0	1	2	0
107075	0	0	1	2	0
107076	0	0	1	1	0
107077	0	1	0	1	0
107078	1	1	1	3	0
107079	0	0	1	1	0
107080	0	1	1	2	0
107088	0	0	1	1	0
107089	0	1	1	2	0
107090	1	0	1	2	0
107091	0	0	1	1	0
107092	0	0	0	0	0
107093	1	0	1	2	0
107094	0	0	1	1	0
107131	0	1	1	2	0
107132	0	1	1	2	0
107137	0	1	1	2	0

107138	0	0	1	1	1
107139	0	0	1	1	0
107145	1	1	0	2	0
107185	0	1	0	1	0
107186	0	0	1	1	0
107195	0	0	1	1	0
107203	0	0	0	0	0

Appendix 6



Supplementary Figure 3: Map over field area near Oslo, Norway (59°56'N, 10°32'E). Red Squares indicate nest boxes where parents and offspring have been sampled from 2021.



Supplementary Figure 4: Map over field area near Oslo, Norway (59°56'N, 10°32'E). Red Squares indicate nest boxes where parents and offspring have been sampled from 2021.

Appendix 7

Supplementary Table 9: Complete list of samples, 2021, plate 1.

107040	107048	107056	107064	107072	107080	107088	107096	107104	107112	107120	107128
107041	107049	107057	107065	107073	107081	107089	107097	107105	107112	107121	107129
107042	107050	107058	107066	107074	107082	107090	107098	107106	107114	107122	107130
107043	107051	107059	107067	107075	107083	107091	107099	107107	107115	107123	107131
107044	107052	107060	107068	107076	107084	107092	107100	107108	107116	107124	107132
107045	107053	107061	107069	107077	107085	107093	107101	107109	107117	107125	107133
107046	107054	107062	107070	107078	107086	107094	107102	107110	107118	107126	107134
107047	107055	107063	107071	107079	107087	107085	107103	107111	107119	107127	NTC1

Supplementary Table 10: Complete list of samples, 2021, plate 2. 107145 & 107146, and 107161 & 107162 was mixed up. They were therefore added again to the end of the plate.

107136	107144	107152	107160	107168	107176	107185	107193	107201	NTC2
107137	107145	107153	107161	107169	107177	107186	107194	107203	107145
107138	107146	107154	107162	107170	107178	107187	107195	107179	107161
107139	107147	107155	107163	107171	107180	107188	107196	107202	107146
107140	107148	107156	107164	107172	107181	107189	107197	107204	107162
107141	107149	107157	107165	107173	107182	107190	107198	107205	
107142	107150	107158	107166	107174	107183	107191	107199	107206	
107143	107151	107159	107167	107175	107184	107192	107200	107135	

Supplementary Table 11: Complete list of samples, 2001, plate 1.

1870	1890	1871	1891	1897	1898	1994	1997	2033	2034	2071	2073
2102	2103	1809	1810	1821	1822	2081	2058	2080	2057		