# Avian MHC

## Characterization and expression patterns of classical and non-classical MHC-I genes

Drews, Anna

2018

*Document Version:*
Other version

Link to publication

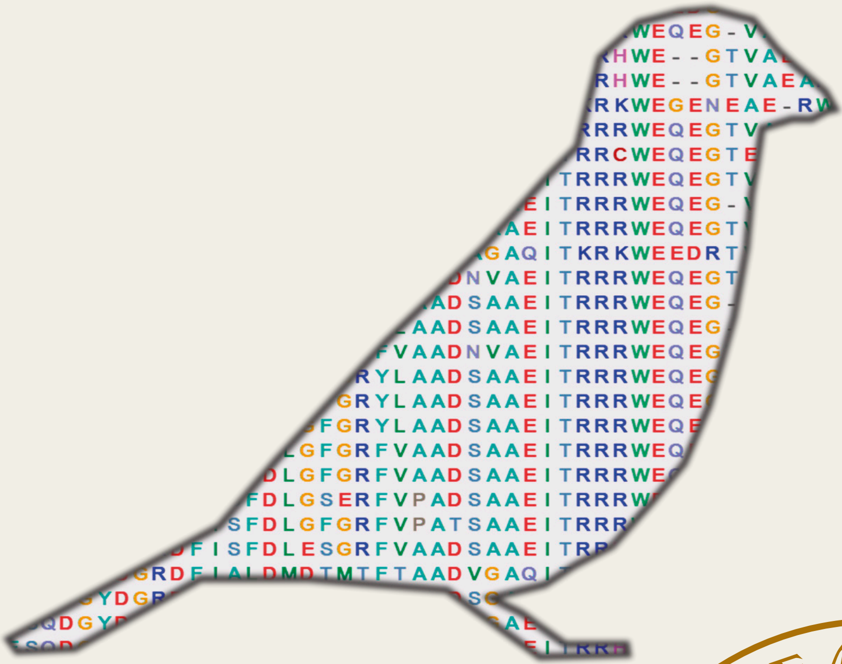
*Citation for published version (APA):*
Drews, A. (2018). *Avian MHC: Characterization and expression patterns of classical and non-classical MHC-I genes*. [Doctoral Thesis (compilation), Department of Biology].

*Total number of authors:*
1

# Avian MHC

## Characterization and expression patterns of classical and non-classical MHC-I genes

ANNA DREWS
DEPARTMENT OF BIOLOGY | LUND UNIVERSITY



# Avian MHC

## Characterization and expression patterns of classical and non-classical MHC-I genes

**Anna Drews**

LUND UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Science, Lund University, Sweden.
To be defended in the Blue Hall, Ecology Building, Sölvegatan 37, Lund, Sweden
on the 7th of December.

*Faculty opponent*
Dr. Tobias Lenz
Max Planck Institute for Evolutionary Biology, Ploen, Germany

Organization
LUND UNIVERSITY

Department of Biology
Sölvegatan 37, 223 62
Lund, Sweden

Author(s) Anna Drews

Document name
DOCTORIAL DISSERTATION

Date of issue
2018-10-29

Sponsoring organization

Title and subtitle Avian MHC - Characterization and expression patterns of classical and non-classical MHC-I genes

Abstract: The function of the vertebrate immune system is to enable recognition and elimination of microorganisms that can cause harm (pathogens). A key component in adaptive immunity is the major histocompatibility complex (MHC) that codes for molecules with antigen presentation function. There are two main types of MHC molecules, class-I (MHC-I) and class-II (MHC-II) and the focus of my thesis has been MHC-I. The number of MHC gene copies can vary greatly between populations, species and orders. Passerine birds (order Passeriformes) seem to have particularly high MHC diversity (defined as number of different MHC alleles per individual). The main aim of my thesis has been to understand the high MHC-I diversity in passerines focusing on two aspects: gene expression and presence of classical and non-classical MHC-I genes. Classical MHC molecules present antigens and trigger adaptive immune responses, whereas non-classical genes have an as yet largely unclear function in immunity. Non-classical MHC genes have lower diversity and often have lower expression. Non-classical genes have been found in several different bird orders although never been fully characterized in passerines. I investigated the presence of classical and non-classical genes using a phylogenetic approach by comparing three closely related sparrow species; house sparrow (*Passer domesticus*), Spanish sparrow (*Passer hispaniolensis*) and tree sparrow (*Passer montanus*). All three species had putatively classical and non-classical genes. All putatively non-classical alleles formed a highly supported cluster, independent of species, indicating that the presence of classical and non-classical genes predates the speciation event of these sparrows. Moreover, only one classical gene was found to be highly expressed in house sparrows and tree sparrows (Paper I). The high diversity of MHC-I alleles makes them difficult to genotype, but with high throughput sequencing (HTS) it is feasible. Initially, I used Roche 454 and then Illumina MiSeq amplicon sequencing. These two HTS methods were evaluated using house sparrow MHC-I and they both performed well (Paper II). We characterized MHC-I in a non-passerine bird, the Icelandic black-tailed godwits (*Limosa limosa islandica*) and showed that there were no putatively non-classical genes in this species despite that such genes were found in two closely related Charadriiformes species (Paper III). I partly characterized MHC-I in siskins (*Spinus spinus*) of the order Passeriformes and it had putatively non-classical genes; one highly supported cluster contained only low expression alleles that also had low diversity (Paper IV). Moreover, I found that as many as three classical genes could have high expression. Expression of MHC-I was then studied in an infection experiment with a mild and a severe avian malaria strain (Paper V). We showed that classical alleles were continuously expressed to a higher degree than non-classical alleles. Moreover, MHC-I was differently expressed in infected individuals compared to control individuals and there was a tendency for MHC-I to be more highly expressed soon after the acute phase of the severe malaria infection. In my thesis, I have shown that the presence of classical and non-classical MHC-I genes most likely is a common feature in passerine birds which has so far been overlooked. Moreover, the expression of MHC-I in passerine birds seems to differ considerably, not only between species but also between classical MHC-I alleles within individuals of the same species – an avenue for future research.

Key words MHC class I, Classical genes, Non-classical genes, Expression, Passerine birds

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language English

ISSN and key title

ISBN
Print: 978-91-7753-906-3
PDF: 978-91-7753-907-0

Recipient's notes

Number of pages

Price

Security classification



Signature

Date 2018-10-29

# Avian MHC

## Characterization and expression patterns of classical and non-classical MHC-I genes

**Anna Drews**

LUND UNIVERSITY

Coverphoto by Anna Drews

Faculty of Science
Department of Biology
ISBN (print) 978-91-7753-906-3
ISBN (PDF) 978-91-7753-907-0

Printed in Sweden by Media-Tryck, Lund University
Lund 2018

*"Sometimes the more I think, the more there is no real answer"*

A.A. Milne

# Table of Contents

# List of papers

I. **Anna Drews**, Maria Strandh, Lars Råberg and Helena Westerdahl (2017) Expression and phylogenetic analyses reveal paralogous lineages of putatively classical and non-classical MHC-I genes in three sparrow species (*Passer*). *BMC Evolutionary Biology* 17: 152

II. Haslina Razali, Emily O'Connor, **Anna Drews**, Terry Burke and Helena Westerdahl (2017) A quantitative and qualitative comparison of illumina MiSeq and 454 amplicon sequencing for genotyping the highly polymorphic major histocompatibility complex (MHC) in a preformed non-model species. *BMC Research Notes* 10: 346

III. Sara Pardal, **Anna Drews**, José A. Alves, Jaime A. Ramos, Helena Westerdahl (2017) Characterization of MHC class I in a long distance migratory wader, the Icelandic black-tailed godwit. *Immunogenetics* 69(7): 463-478

IV. **Anna Drews** and Helena Westerdahl. Characterization of putatively classical and non-classical MHC-I genes in the Eurasian siskin (*Spinus spinus*): no evidence of a single majorly expressed MHC-I gene. Manuscript

V. **Anna Drews**, Lars Råberg, Olof Hellgren, Max Lundberg, Vaidas Palinauskas and Helena Westerdahl. Expression of MHC-I genes during mild and severe avian malaria infections in Eurasian siskins (*Spinus spinus*). Manuscript

# Author Contributions

I. **AD** and HW conceived the study. **AD** performed the lab work. **AD** analyzed the data with input from HW and LR. **AD** wrote the paper with input from all authors.

II. HR, TB and HW conceived the study. HR performed the field and lab work. HR, EO, **AD** and HW analyzed the data. HR and EO wrote the paper with input from all authors.

III. SP and HW conceived the study. SP and JA performed the field work. SP and **AD** performed the lab work. SP analyzed the data with input from **AD** and HW. SP wrote the paper with input from all authors.

IV. **AD** and HW conceived the study. **AD** performed the lab work. **AD** analyzed the data with input from HW. **AD** wrote the manuscript with input from HW.

V. OH, **AD** and HW conceived the study. VP performed the experiment. **AD** performed the lab work. **AD** performed the bioinformatic analysis with input from ML. **AD** and LR analyzed the data with input from HW. **AD** wrote the manuscript with input from all authors.

# Abstract

The function of the vertebrate immune system is to enable recognition and elimination of microorganisms that can cause harm (pathogens). A key component in adaptive immunity is the major histocompatibility complex (MHC) that codes for molecules with antigen presentation function. There are two main types of MHC molecules, class-I (MHC-I) and class-II (MHC-II) and the focus of my thesis has been MHC-I. The number of MHC gene copies can vary greatly between populations, species and orders. Passerine birds (order Passeriformes) seem to have particularly high MHC diversity (defined as number of different MHC alleles per individual). The main aim of my thesis has been to understand the high MHC-I diversity in passerines focusing on two aspects: gene expression and presence of classical and non-classical MHC-I genes. Classical MHC molecules present antigens and trigger adaptive immune responses, whereas non-classical genes have an as yet largely unclear function in immunity. Non-classical MHC genes have lower diversity and often have lower expression. Non-classical genes have been found in several different bird orders although never been fully characterized in passerines.

I investigated the presence of classical and non-classical genes using a phylogenetic approach by comparing three closely related sparrow species; house sparrow (*Passer domesticus*), Spanish sparrow (*Passer hispaniolensis*) and tree sparrow (*Passer montanus*). All three species had putatively classical and non-classical genes. All putatively non-classical alleles formed a highly supported cluster, independent of species, indicating that the presence of classical and non-classical genes predates the speciation event of these sparrows. Moreover, only one classical gene was found to be highly expressed in house sparrows and tree sparrows (Paper I). The high diversity of MHC-I alleles makes them difficult to genotype, but with high throughput sequencing (HTS) it is feasible. Initially, I used Roche 454 and then Illumina MiSeq amplicon sequencing. These two HTS methods were evaluated using house sparrow MHC-I and they both performed well (Paper II). We characterized MHC-I in a non-passerine bird, the Icelandic black-tailed godwits (*Limosa limosa islandica*) and showed that there were no putatively non-classical genes in this species despite that such genes were found in two closely related Charadriiformes species (Paper III). I partly characterized MHC-I in siskins (*Spinus spinus*) of the order Passeriformes and it had putatively non-classical genes; one highly supported cluster contained only low expression alleles that also had low diversity (Paper IV). Moreover, I found that as many as three classical genes could have high expression.

Expression of MHC-I was then studied in an infection experiment with a mild and a severe avian malaria strain (Paper V). We showed that classical alleles were continuously expressed to a higher degree than non-classical alleles. Moreover, MHC-I was differently expressed in infected individuals compared to control individuals and there

was a tendency for MHC-I to be more highly expressed soon after the acute phase of the severe malaria infection.

In my thesis, I have shown that the presence of classical and non-classical MHC-I genes most likely is a common feature in passerine birds which has so far been overlooked. Moreover, the expression of MHC-I in passerine birds seems to differ considerably, not only between species but also between classical MHC-I alleles within individuals of the same species — an avenue for future research.

# Svensk sammanfattning

Alla ryggradsdjur har ett immunförsvar vars funktion är att skydda oss mot mikroorganismer som kan orsaka sjukdom, såsom bakterier och virus. Dessa sjukdomsframkallande ämnen kallas patogener. Immunförsvaret behöver kunna skilja mellan proteiner som kommer från den egna kroppen och från patogener. Molekylerna som kodas av gener i major histocompatibility complex (MHC) har en viktig del i denna igenkänning. MHC-molekylerna uttrycks på cellytan och deras funktion är att binda till små fragment av proteiner, så kallade peptider, och presentera dessa för en typ av immunförsvarceller, så kallade T-celler. Om peptiden som MHC presenterar är från en patogen så startar en immunreaktion som kommer att leda till att den infekterade cellen förstörs. Det som gör MHC så viktigt är att T-celler bara kan känna igen peptider om de sitter bundna till en MHC-molekyl. MHC kan delas upp i två klasser beroende på var peptiderna som de presenterar kommer ifrån. MHC klass I presenterar peptider som finns inuti celler (till exempel peptider från virus) medan MHC klass II presenterar peptider som finns utanför celler (till exempel peptider från bakterier). I min avhandling har jag bara fokuserat på MHC klass I.

Man har oftast bara en genkopia av varje gen men det som är speciellt med MHC-gener är att det finns många genkopior, människa har till exempel tre MHC klass I gener. En grupp av ryggradsdjur som har extremt många MHC-genkopior är sångfåglar som kan ha mellan 2 och 33 genkopior, där medelvärdet är 11 genkopior. Detta är betydligt fler genkopior än vad som har hittas hos andra grupper av fåglar och däggdjur. Varför sångfåglar har så många MHC-genkopior vet man inte och i min avhandling har jag jobbat med två aspekter som kan förklara det höga antalet MHC-gener som finns hos sångfåglar.

Den första aspekten jag har undersökt är om alla MHC-gener har samma funktion, det finns nämligen både klassiska och icke klassiska MHC-gener. Funktionen hos klassiska MHC-gener är att presentera peptider till T-celler. Icke-klassiska gener har också en funktion i immunförsvaret men den kan variera och olika icke-klassiska gener verkar ha olika funktioner. Det som kännetecknar icke-klassiska gener är att de är betydligt mindre variabla än klassiska gener och att de dessutom uttrycks till lägre grad än klassiska gener. En förklaring till sångfåglars höga antal MHC-genkopior skulle kunna vara att en stor andel är icke-klassiska. Jag har även undersökt om alla klassiska gener uttrycks lika mycket. Hos många andra fågelarter har man hittat att det bara är en genkopia av de klassiska MHC-generna som har ett riktigt högt genuttryck, oberoende av hur många genkopior som finns i genomet. Såvitt jag vet finns det inga studier som har undersökt om sångfåglars alla genkopior är uttryckta lika mycket så det är möjligt att en annan förklaring till det höga antalet genkopior är att fler gener är höguttryckta hos sångfåglar jämfört med andra grupper av fåglar.

I min avhandling har jag kommit fram till att både klassiska och icke-klassiska gener verkar vara vanliga hos sångfåglar eftersom jag har hittat dem i fyra olika arter av sångfåglar; gråsparv, pilfink, spansk sparv och grönsiska. Dessa icke-klassiska gener är alla mindre variabla och mer lika varandra jämfört med klassiska gener. Dessutom är dessa icke-klassiska gener alltid lägre uttryckta än klassiska gener. Detta tyder på att icke-klassiska gener är vanliga hos sångfåglar men proportionen mellan icke-klassiska och klassiska varierar mellan de fyra arterna som jag har undersökt; hos gråsparv och spansk sparv finns det fler icke-klassiska än klassiska gener medan hos pilfink och grönsiska så finns det fler klassiska än icke-klassiska gener. Alltså kan inte det höga antalet genkopior hos sångfåglar enbart förklaras av att de har icke-klassiska gener.

Uttrycket av de klassiska MHC-generna varierar stort hos sångfåglar, vissa gener har högt uttryck medan andra har lågt uttryck. En intressant sak som jag har kommit fram till är att antalet MHC-gener som har högt uttryck verkar variera mellan arter. Hos gråsparv och pilfink hittar jag bara en gen som har högt uttryck medan hos grönsiska hittar jag som mest tre gener. Detta tyder på att sångfåglar inte bara har en gen som är höguttryckt utan att fler kan vara höguttryckta vilket delvis skulle kunna förklara det höga antalet genkopior hos sångfåglar.

Då funktionen av MHC är att presentera peptider från patogener har jag även undersökt vad som händer med MHC uttrycket under en pågående infektion hos grönsiskor, i det här fallet två olika typer av fågelmalaria. Jag fann att klassiska gener alltid är mer höguttryckta än icke klassiska gener. Sett över hela malariainfektionen så uppreglades MHC bara väldigt svagt i blodet, men det finns en tendens att MHC är mer höguttryckt sent i infektionen, den akuta fasen av malaria, precis efter att antalet malariaparasiter i de röda blodkropparna börjar gå ner.

Sammanfattningsvis har min avhandling lyft fram två viktiga förklaringar till varför sångfåglar har så många MHC gener, nämligen att det troligtvis finns både klassiska och icke-klassiska gener och att mer än en gen kan vara höguttryckt, och därmed ha en viktig funktion i immunsvaret.

# Introduction and background

The antigen presenting molecules encoded by the major histocompatibility complex (MHC) are a key component of the vertebrate immune system and since their discovery back in 1936 (Gorer 1936a, b) these genes have fascinated scientists. One aspect that separates these genes from many other are that they are highly duplicated, and that the number of gene copies can vary quite dramatically between species.

One group that have extremely high MHC diversity in general are passerine birds belonging to the order Passeriformes, especially compared to species from other bird orders. The exact reason as to why passerine birds have this high MHC diversity is not known. The focus of my thesis is to examine aspects that could help us to understand the high MHC diversity in passerine birds.

## The vertebrate immune system

The function of the immune system is to protect us from microorganisms that can cause us harm, such as viruses, bacteria and parasites, so called pathogens (Williams 2011). The immune system can be dived into three major barriers with the function of either preventing pathogens from entering the body or if they are able to enter to eliminate them (Abbas et al. 2016). The first barrier consists of several different physical barriers, such as the skin and the mucosal layers in the respiratory and gastrointestinal tracts, that can prevent pathogens from entering the body (Cruse and Lewis 2003). The next barrier is a chemical barrier made up of, for example, many antimicrobial substances and the acidic environment in the stomach (Williams 2011). Finally, the last barrier is the large number of specialized cells that are referred to as immune cells (Murphy and Weaver 2016). One of the functions of immune cells is to distinguish between self and non-self and only trigger an immune response against foreign substances (Cruse and Lewis 2003). These immune cells can then further be divided into the innate and the adaptive immune system, depending on their function and how they recognize foreign substances (Abbas et al. 2016).

### The innate immune system

The innate immune system recognizes pathogens based on conserved structures that are only found in pathogens, via so called pattern-recognition receptors (PRR) (Murphy and Weaver 2016). After the recognition of a pathogen an immune response will be triggered (Abbas et al. 2016). The response of the innate immune system has relatively low specificity (Cruse and Lewis 2003). The innate immune system has a short response time. It takes a matter of hours after the pathogen has entered the body until it is activated (Williams 2011).

### The adaptive immune system

The adaptive immune system is more specific than that of the innate immune system (Murphy and Weaver 2016). Another difference is that it takes longer time to activate the adaptive immune system, between seven to fourteen days from the time the pathogen enters the body (Williams 2011). Finally, the adaptive immune system can produce memory cell which will lead to a stronger and faster response the second time the same pathogen is encountered (Abbas et al. 2016).

One of the important initial steps of the adaptive immune response is the so-called antigen presentation, during which small fragments of a pathogen is presented by molecules on the cell surface which will trigger an immune response (Cruse and Lewis 2003). The major histocompatibility complex codes for molecules that key role in antigen presentation (Murphy and Weaver 2016).

## The Major Histocompatibility Complex

The Major Histocompatibility Complex is a multi-gene family where the MHC molecules have a central role in antigen presentation (Neefjes et al. 2011). The function of MHC molecules is to present peptides stemming either from the host or from pathogens, to a specialized type of immune cells called T-cells (Abbas et al. 2016). An immune response should only be triggered when the presented peptide is foreign (Murphy and Weaver 2016).

The MHC genes can be divided into two main classes depending on the peptides that are presented and the type of T-cells that they are interacting with. MHC class I (MHC-I) presents intracellular peptides (*e.g.* from viruses) to cytotoxic T-cells (CD8+ T-cells) whereas MHC class II (MHC-II) presents peptides from extracellular pathogens (*e.g.* many bacterial pathogens) to T helper cells (CD4+ T-cells) (Murphy and Weaver 2016). MHC-I can be found on all nucleated cells and MHC-II can be found on specialized immune cells called antigen presenting cells (Neefjes et al. 2011). The focus of this thesis is on MHC-I.

The full MHC gene region consists of many different genes and are found on chromosome 6 in humans. The genes central to antigen presentation are found in the core MHC region including those that code for the MHC molecule (Murphy and Weaver 2016). In mammals, the MHC core region seems to be rather conserved (Kelley et al. 2005).

## The diversity of MHC-I

MHC genes have high intra-individual diversity meaning both that an individual has several MHC gene copies and that the alleles at these different gene copies often are different (*e.g.* the genes are heterozygous). Moreover, different individuals frequently have different alleles (Murphy and Weaver 2016). For example, at one MHC-I gene in humans over 5,200 different alleles have been identified worldwide (The HLA Database; Robinson et al. 2015). The MHC region is one of the most diverse gene regions found in vertebrates.

Multiple gene copies (so called paralogous) are created by gene duplication. In paralogues the selective pressure of maintaining the original gene function can be relaxed which leaves these new gene copies free to evolve different functions, a process called neo-functionalization (Ohno 1970; Eirín-López et al. 2012). In the case of MHC genes, new gene copies are thought to evolve by several different mechanisms of which one is the so called ‘birth and death process’ during which some very old gene copies retain the original function and are kept whereas other gene copies evolve slightly different functions and yet other gene copies become non-functional (Nei et al. 1997; Van Oosterhout 2009).

The high MHC diversity, *i.e.* the total number of alleles across all gene copies, that is seen for the MHC genes are thought to be maintained mainly by balancing selection (Hedrick and Thomson 1983; Hedrick 2007). There are several mechanisms that can result in balancing selection. One example is heterozygote advantage (Doherty and Zinkernagel 1975) where individuals that have different alleles at a gene copy have a higher chance at recognizing more pathogens. This would lead to a higher proportion of heterozygote individuals in a population. Another example is divergent allele advantage (Wakeland et al. 1990) and this theory says that it is even better to have alleles that are very different from each other. The more distant the alleles are the higher the possibility that they can bind to different peptides and hence give protection to a wide range of pathogens. In addition there is negative-frequency-dependent selection (Slade and McCallum 1992), which states that there are both common and rare alleles in a population and that the pathogens might evolve in order to avoid detection by the common alleles but will still be recognized by the more rare alleles. These rare alleles will then increase in frequency within the population and become common and new alleles will be rare. This will keep the variation at MHC high over time. Finally there is also fluctuating selection (Hedrick 2002), which means that because pathogens vary in abundance over time and space so will the MHC alleles, maintaining MHC gene diversity. These different examples of balancing selection are not mutually exclusive and have been investigated in both laboratory and natural populations (both MHC class I and II) (*e.g.* Doherty and Zinkernagel 1975; Wakeland et al. 1990; Slade and McCallum 1992; Ekblom et al. 2007; Hedrick 2012; Lenz et al. 2013; Sepil et al. 2013; Pierini and Lenz 2018).

## Antigen presentation of MHC-I

In more detail, what happens during antigen presentation, Figure 1, is that after a pathogen has entered a host cell and start replicating the proteins that are produced will be broken down to peptides by the proteasome (in the cytosol) (Goulder and Walker 2012). Self-proteins are also constantly broken down (Abbas et al. 2016). These peptides will then be transported into the endoplasmatic reticulum (ER) by a molecule called transporter associated with antigen processing (TAP) where the peptide is loaded onto an MHC-I molecule (Murphy and Weaver 2016). Each MHC-I molecule can bind peptides with certain characteristics and if, inside the ER, a stable MHC-I - peptide complex is formed it is transported to the cell surface where the peptide is presented to cytotoxic T-cells (Williams 2011). If the T-cell receptor is able to bind to the peptide-MHC-I complex an immune response will be trigged and the infected cell will be destroyed (Murphy and Weaver 2016).

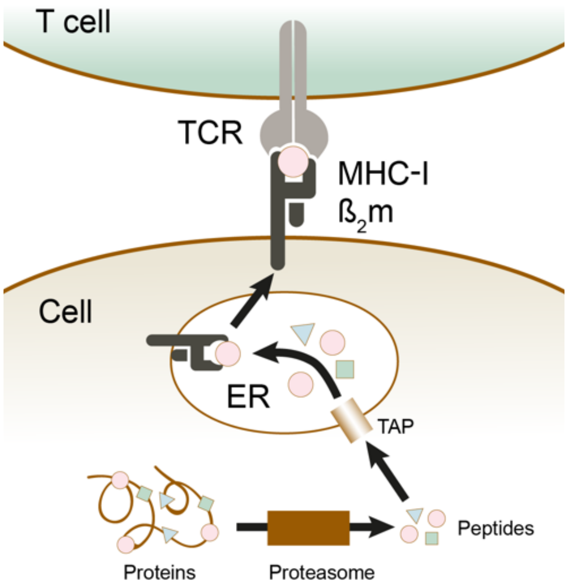



**Figure 1.** Schematic representation of intracellular antigen presentation by MHC-I. Proteins are broken down into peptides by the proteasome and the petides are then transported into the ER (endoplasmic reticulum) by TAP (transporter associated with antigen processing). The peptide is loaded onto the MHC molecule and transported to the cell surface where the peptide is presented to T-cells, if the T-cell receptor (TCR ) binds to both the peptide and the MHC molecule an immune response will be triggered. Drawn by Inger Ekström, inspired by Neefjes et al. (2011) and Abbas et al. (2016).

## The structure of MHC-I

The extracellular part of the MHC molecule is encoded by three exons (exon 2, 3 and 4) which are translated into three subparts of the MHC molecule (α1, α2 and α3; Figure 2). The MHC molecule is stabilized by another protein called beta-2 microglobulin ($\beta_2$m) (Murphy and Weaver 2016). The α1 and α2 regions form the peptide binding region (PBR) (Cruse and Lewis 2003). Each MHC molecule can bind a limited number of peptides, which ones depends on the amino acids in the PBR (Murphy and Weaver 2016). There are certain key amino acid positions in the PBR and these will determine the peptides an MHC molecule can bind (Altuvia and Margalit 2004). Because of the antigen binding ability the PBR is the most variable part of the MHC molecule (Murphy and Weaver 2016). The α3 region of the MHC molecule is more conserved compared to the α1 and α2 regions, because this region contains the binding sites for the CD8 co-receptor (Murphy and Weaver 2016). In order for the T-cell to initiate an adaptive immune response the CD8 co-receptor needs to bind to the MHC molecule (Cruse and Lewis 2003). The adaptive immune response should only be triggered when non-self peptides are presented and when several co-receptors are bound. Hence, there are multiple control steps ensuring that an immune response is not triggered by self-peptides (Murphy and Weaver 2016).

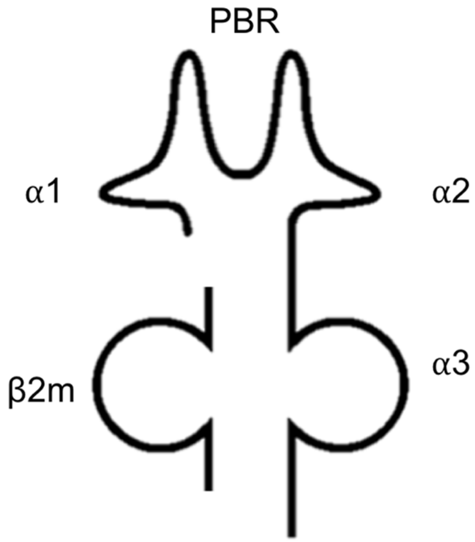



**Figure 2.** Schematic representation of the structure of the MHC-I extracellular part. The MHC-I molecule consists of three subparts encoded by three exons. α1, encoded by exon 2 and α2, encoded by exon 3 toghther make up the peptide binding region (PBR). The α3, encoded by exon 4, is a more conserved part of the molecule. The MHC molecule is stablized by another protein called beta-2 microglobulin (ß2m).

## Classical and non-classical MHC-I genes

Both MHC-I and -II genes can be divided into so called classical and non-classical MHC genes (Rodgers and Cook 2005). The classical MHC genes code for molecules that have a central role in the adaptive immune system, whereas non-classical molecules have more variable immune functions (Allen and Hogan 2013). The hallmarks of classical MHC genes are that they are highly diverse, bind and present peptides to T-cells, and are expressed in most tissues (Murphy and Weaver 2016). Non-classical MHC genes on the other hand code for molecules that resemble classical MHC molecules but rarely presents peptides to T-cells (Shawar and Vyas 1994). The non-classical genes have lower diversity, and are expressed at lower levels than classical MHC genes. Moreover, they often have restricted tissue expression (Rodgers and Cook 2005).

Non-classical genes have been identified in most of the vertebrate classes (*e.g.* in mammals, amphibians, reptiles, birds, bony fishes, cartilaginous fishes) (Kaufman et al. 1999b; Adams and Parham 2001; Lunney et al. 2009; Shiina et al. 2009; Grimholt et al. 2015). The function of non-classical genes seem to vary between species (Allen and Hogan 2013). Overall, non-classical genes are most well studied in mammals, particularly in mice and humans (Le Bouteiller and Lenfant 1996).

In humans, MHC genes are called human leukocyte antigens (HLA) and there are three classical MHC-I genes (HLA-A, -B, -C) and three non-classical genes (HLA-E, -F, -G) (Murphy and Weaver 2016). The number of different alleles per gene that have been identified at these genes varies considerably. Thousands of alleles have been identified for the classical genes whereas only slightly more than one hundred alleles have been identified for the non-classical genes, see Table 1 for specific allele numbers per gene (The HLA Database; Robinson et al. 2015). The pattern is even more extreme when comparing the number of proteins into which these alleles are translated.

**Table 1.** Number of alleles identifed at the three classical and non-classical genes in humans, collected from the HLA database (http://hla.alleles.org/nomenclature/stats.html), information is accurate as of september 2018.

| | **Clasical genes** | | | **Non-classical genes** | | |
|---|---|---|---|---|---|---|
| | HLA-A | HLA-B | HLA-C | HLA-E | HLA-F | HLA-G |
| Alleles | 4,638 | 5,590 | 4,374 | 27 | 31 | 61 |
| Proteins | 3,172 | 3,923 | 2,920 | 8 | 6 | 19 |

Although the three non-classical genes in humans have similarly low levels of diversity, the exact function of these genes seem to be variable. The most well studied non-classical gene in humans is HLA-E (Crux and Elahi 2017). HLA-E have a wider tissue expression than most non-classical genes but is always expressed to a lower degree than classical MHC-I molecules (Wei and Orr 1990). HLA-E does not present peptides from pathogens but instead seems to present peptides stemming from the signal pathway that activates MHC-I, and instead of presenting to T-cells they interact with Natural Killer cells (NK-cells) (Crux and Elahi 2017). HLA-F is the least studied of the three non-classical genes in humans and is often expressed in lymphocytes and seems to be involved in regulation of NK-cells (Lee et al. 1998; Lepin et al. 2000; Lee and Geraghty 2003; Garcia-Beltran et al. 2016). HLA-G is mainly expressed in the placenta and seems to be involved in protecting the foetus from the mother's immune system. It can interact with a number of immune cells such as antigen presenting cells, NK cells, and small populations of T cells (Vianna et al. 2007; Larsen et al. 2010).

The origin of non-classical genes has not fully been determined. There are non-classical genes that seem to be older than classical genes, although those are mainly located outside of the core MHC region, and there are non-classical genes that seem to be younger than the classical genes, and may have originated due to gene duplication of classical MHC genes (Hughes and Nei 1989; Rodgers and Cook 2005). There are examples of non-classical genes that predate speciation events. Humans and chimpanzee (*Pan troglodytes*), species that diverged 6 to 7 million years ago, form gene specific clusters at the six gene copies corresponding to HLA -A, -B, -C, -E, -F, -G (Adams and Parham 2001).

# MHC-I in birds

The first avian species where MHC-I was characterized in detail was the domestic chicken (*Gallus gallus domesticus*) (Kaufman et al. 1999b). The chicken has an MHC region that is considerably different from mammals. The whole gene region is about 20 times smaller than the human MHC region and not all genes that are found within this region in humans are found in the chicken region. Because of this, the chicken is said to have a 'minimal essential MHC' (Kaufman et al. 1999b). Regarding the of number of MHC-I genes chicken have two classical and at least two non-classical MHC-I genes (Miller et al. 1996; Kaufman et al. 1999b; Afanassieff et al. 2001).

Later on, the MHC region in more species belonging to the order Galliformes has been partly characterized, for example in turkey (*Meleagris gallopavo*) where two classical MHC-I genes were identified, quail (*Coturnix japonica*) with two classical MHC-I genes and at least two non-classical genes, black grouse (*Tetrao tetrix*) with two classical MHC-I genes, and golden pheasant (*Chrysolophus pictus*) with two classical and at least one non-classical MHC-I genes. All these species turned out to have a more or less 'minimal essential MHC' (Shiina et al. 1999, 2004; Chaves et al. 2009; Wang et al. 2012; Zeng et al. 2016). Outside the order of Galliformes, mallards belonging to the order Anseriformes, have been partly characterized and mallards also have a compact MHC region with five MHC-I genes where one is suggested to be non-classical (Moon et al. 2005). Hence, it was suggested that most birds have compact MHC regions. However, recently the crested ibis (*Nipponia nippon*) from the order Pelecaniformes was partly characterized and it does not seem to have a 'minimal essential MHC'. Instead, it had features that were more similar to the mammalian MHC region (Chen et al. 2015). Interestingly, early on it was suggested that passerines belonging to the order Passeriformes do not have a 'minimal essential MHC' but instead seem to have a highly diverse MHC based on the number of alleles identified (*e.g.* Westerdahl et al. 1999, 2000). Despite that the exact MHC-I region has not been determined in passerine birds, with the rise of high throughput sequencing methods (HTS), it is now possible to reliably genotype MHC-I also in species with high diversity like the passerines

## MHC-I genes in Passerines

Most studies in passerines have only sequenced one MHC-I exon (mostly exon 3). Because there is high similarity between gene copies, at least in exon 3, it is still not possible to determine which allele that belongs to which gene copy when simultaneously amplifying over all gene copies (*e.g.* Sepil et al., 2012; Karlsson & Westerdahl, 2013; O'Connor et al., 2016). Hence, the number of genes in passerines are determined based on the number of different alleles identified and since MHC genes often are heterozygous the number of genes is estimated by dividing the maximum number of alleles with two.

Even though passerines in general have high MHC-I diversity there is considerable copy number variation (CNV) between species within Passeriformes, Figure 3. O'Connor et al. (2016) reported between four and 19 MHC-I gene copies per species across Passerida (containing *e.g.* warblers, flycatchers and pipits) based on genomic MHC-I exon 3 sequences in open reading frame. The highest number of MHC-I gene copies reported to date is in the sedge warbler (*Acrocephalus schoenobaenus*) where up to 65 MHC-I alleles have been identified in one individual, which suggests that sedge warblers can have at least up to 33 gene copies (Biedrzycka et al. 2017a). In contrast, the zebra finch (*Taeniopygia guttata*), with only two MHC-I genes seems to be an outlier among passerine birds (Ekblom et al. 2011). There is also CNV within passerine species, for example between 7-10 MHC-I genes have been identified in house sparrows (*Passer domestics*) (O'Connor et al. 2016). The exact purpose of passerines multiple MHC-I gene copies is not known. Throughout this thesis two aspects of MHC-I diversity have been considered, the presence of both classical and non-classical genes as well as the expression patterns of the multiple MHC-I genes.

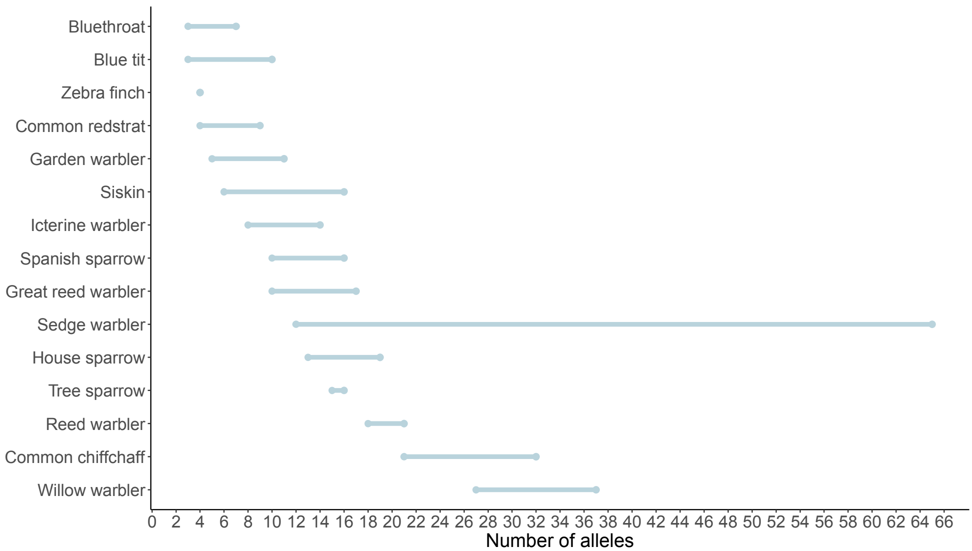



**Figure 3.** Number of MHC-I exon 3 alleles identifed in 15 passerine species. Data from O'Connor et al 2016, Biedrzycka et al. 2017a and Ekblom et al 2011. All data is from HTS methods execpt zebra finch where two genes have been determined based on genetic mapping.

## Classical and non-classical alleles in birds

When the chicken MHC region was characterized a separate independent cluster with MHC-like genes was identified which was named MHC-Y and was determined to contain at least two non-classical MHC-I genes (Briles et al. 1993; Miller et al. 1994). When these genes belonging to the MHC-Y region were further studied it was shown that the alleles belonging to MHC-Y were more similar to each other than to alleles belonging to the classical gene cluster, called MHC-B in chicken (Afanassieff et al. 2001). Moreover, there is an amino acid substitution in the PBR which makes the non-classical MHC-I molecules in chicken unlikely to bind peptides in the same way as classical MHC (Afanassieff et al. 2001). Also, the non-classical genes have restricted tissue expression (Hunt et al. 2006).

Later on, when characterizing more Galliformes species, non-classical genes were also identified in quail, turkey and golden pheasant (*Chrysolophus pictus*) (Shiina et al. 2004; Chaves et al. 2009; Zeng et al. 2016). Although the exact number of genes is not always determined, non-classical genes seem to be a common feature in Galliformes birds. In fact, non-classical genes seem to be very common in birds in general and have been identified in multiple species, such as, the mallard (*Anas platyrhynchos*) (Moon et al. 2005), the crested ibis (Chen et al. 2015) and red-billed gull (*Larus scopulinus*) (Cloutier et al. 2011).

As for passerine birds, there is a single study that has suggested that house sparrows could have non-classical genes. Karlsson and Westerdahl (2013) identified MHC-I alleles with a 6 bp deletion in exon 3. These alleles also formed a distinct cluster in a phylogenetic tree. Moreover, these alleles had lower genetic diversity compared to alleles without the 6 bp deletion strongly suggesting that these alleles belong to non-classical genes. Hence it is possible that non-classical genes are an overlooked feature of passerine birds.

## Expression of MHC-I genes in birds

Another explanation for multiple MHC-I genes in passerine birds could be that not all of them are equally expressed. For example, in humans, the three classical genes (HLA-A, -B, -C) are all expressed but there is slight variation in expression levels between these genes. HLA-A and -B genes are expressed to a higher degree than HLA-C genes, leading to a large variation in gene expression (Apps et al. 2015).

A different expression pattern has been seen in birds. Out of the two classical genes in chicken only one is highly expressed, called the major gene, whereas the other gene, called the minor, has low level of expression. (Kaufman et al. 1995, 1999a, 1999b; Wallny et al. 2006; Shaw et al. 2007). One highly expressed gene has also been identified in several other bird species, such as quail, mallard (both in domestic and wild mallards), red-billed gull and crested ibis (Mesa et al. 2004; Shiina et al. 2006; Cloutier et al. 2011; Chen et al. 2015; Fleming-Canepa et al. 2016). Interestingly, in some of these species more than two gene copies occur so this expression pattern is not only caused by the number of genes. For example, mallards have five gene copies, four classical and one non-classical, still only one gene copy has high expression whereas the other four gene copies have low or no expression (Mesa et al. 2004; Moon et al. 2005). It was suggested that the difference in expression seen between these gene copies are regulated by microRNA (Chan et al. 2016). Moreover, Chan et al. (2016) showed that the two alleles at the highly expressed gene copy were also differently expressed, caused by difference in promoter activation.

Within passerines there are few studies that have considered the expression of MHC-I genes, but there are a few examples of how many MHC-I genes that could be expressed. In great read warblers (*Acrocephalus arundinaceus*), eight different full length transcripts have been identified in cDNA in a single individual, indicating that at least four genes are expressed, and in blue tit (*Cyanistes caeruleus*) five MHC exon 3 transcripts, *i.e.* 3 genes, were found (Westerdahl et al. 1999; Schut et al. 2011). These findings show that more than one gene is expressed in passerine birds. Since most of the multiple alleles identified in passerine birds seem to be functional based on the available sequence data (Bonneaud et al. 2004; Schut et al. 2011; Sepil et al. 2012; Karlsson and Westerdahl 2013; O'Connor et al. 2016; Roved et al. 2018) it is possible that passerine birds express more than one MHC-I gene. However, no studies have looked at the level of expression of different MHC alleles within passerine birds, information that is needed in order to determine if passerine birds also have a majorly expressed MHC-I gene.

Kaufman (1999) hypothesized that the reason for a major gene in the chicken is that this MHC-I gene has coevolved with TAP, the protein responsible for transporting peptides into the ER. The TAP molecule is encoded by two genes, TAP1 and TAP2 and in the chicken genome the major MHC gene is located next to TAP2 (Kaufman et al. 1999a). The idea is that TAP will only transport peptides that the particular major MHC molecule can bind (Kaufman 2015a). This co-evolution will then result in high TAP diversity and that certain TAP-MHC haplotypes will be inherited together (Walker et al. 2005, 2011; Kaufman 2015b). To date, in all species where a majorly expressed MHC gene have been identified it is also located next to a TAP gene, providing support for the idea that in birds MHC-I and TAP have co-evolved (Mesa et al. 2004; Shiina et al. 2006; Cloutier et al. 2011; Chen et al. 2015; Fleming-Canepa et al. 2016).

Interestingly, a tight linkage between MHC-I and TAP has not only been reported from birds but also from other vertebrates such as sharks, teleost fishes, and amphibians (Flajnik et al., 1999; Kelley et al., 2005). In both salmon (*Salmo salar*) and the frog *Xenopus*, one majorly expressed MHC gene has been identified and it is also located next to TAP (Grimholt et al., 2002; Ohta et al., 2003). This suggests that a strong linkage between MHC and TAP could be the ancestral stage. On the contrary, in mammals where all MHC genes seems to be expressed, MHC and TAP are not found next to each other and hence cannot co-evolve (Kaufman 2015a; Murphy and Weaver 2016). Interestingly, not all birds seem to have just one highly expressed MHC-I gene, for example in one red knot individual as many as five alleles, *i.e.* 3 genes, have been identified as highly expressed (Buehler et al. 2013). Where TAP is located within passerines is not known but the preliminary data from the zebra finch genome have placed TAP and MHC on the same chromosome but it has not been possible to determine the distance between them (Balakrishnan et al. 2010; Ekblom et al. 2011).

# Thesis aims

The overall aim of my thesis was to gain a deeper understating of the high MHC-I diversity seen in passerine birds. In particular, I wanted to determine if passerine birds, similarly to many other bird species, have both classical and non-classical genes as well as if only one MHC gene copy is highly expressed.

In paper I, the aim was to characterize putatively classical and non-classical MHC genes in three sparrow species in order to investigate MHC-I organization in closely related species. We also wanted to compare how many alleles that were expressed, as well as determine the relative expression level of these alleles, *i.e.* determine the number of alleles that have high and low expression.

In paper II, the aim was to compare two high throughput sequencing methods that are commonly used when genotyping MHC-I in passerine birds.

In paper III, the aim was to characterize MHC-I genes in a non-passerine bird, the black-tailed godwit, and determine if putatively classical and non-classical genes are also common in the order Charadriiformes. Moreover, by determining the diversity of MHC-I in a non-passerine it can later on be compared to passerines.

In paper IV, the aim was to further determine how common putatively classical and non-classical genes are in passerines by characterizing MHC-I in a novel species, the Eurasian siskin. The aim also was to determine the number of expressed alleles as well as the proportion that had high expression.

In paper V, continuing on the work from paper IV, the aim was to determine how the expression of MHC-I, both putatively classical and non-classical alleles, changed throughout an infection, in this case both a mild and severe avian malaria infection.

# General methodology

## Study species

Throughout this thesis I have manly studied MHC-I in bird species from the order Passeriformes but also in one species of the order Charadriiformes.

### Sparrows (*Passer*)

I have studied three species of sparrows (*Passer*): house sparrow (*P. domesticus*), Spanish sparrow (*P. hispaniolensis*) and tree sparrow (*P. montanus*), Figure 4. All three species can be found in both urban and rural environments but their native range differs; house sparrows and tree sparrows can be found throughout most of Eurasia whereas Spanish sparrows are only found around the Mediterranean Sea and in south-west Asia (Mullarney et al. 2006). The house sparrows and Spanish sparrows are more closely related to each other than to the tree sparrow (birdtree.org; Jetz et al. 2012). House sparrow and Spanish sparrow split from each other three million years ago and tree sparrows split from the other two species seven million years ago. The house sparrows used in this thesis were caught in one of two locations, either at Löberöd, Skåne, Sweden or on Lundy Island, located in the Bristol Channel, UK. Tree sparrows were caught in Löberöd, Skåne, Sweden and the Spanish sparrows were from a captive population at University of Oslo, Norway.

**Figure 4**. The three sparrows studied during this thesis; house sparrow (left), Spanish sparrow (middle), tree sparrow (right). Photo: Julio Neto

## Eurasian siskin (*Spinus spinus*)

The Eurasian siskin (*Spinus spinus*) is a small passerine bird, Figure 5, belonging to the finch family (Fringillidae) and the native range of siskins are rural areas all throughout Eurasia (Mullarney et al. 2006). Based on phylogenetic analysis, siskins and sparrows are quite distantly related, the species split from each other 29 million years ago (birdtree.org; Jetz et al. 2012). The siskins used in this thesis were caught on the Curonian Spit in the Baltic Sea (Kaliningrad region, Russia).

**Figure 5.** Eurasin siskin (*Spinus spinus*). Photo: Vaidas Palinauskas

## Icelandic black-tailed godwit (*Limosa limosa islandica*)

The Icelandic black-tailed godwit (*Limosa limosa islandica*) is a migratory wader that is found in the western Palearctic (Delany et al. 2009). It breeds predominantly on Iceland, and outside the breeding season, godwits move along the Atlantic coast, preferring brackish habitats such as sheltered estuaries, lagoons, and large intertidal mudflats. Godwits winter in temperate countries (mostly Iberia) (Gill et al. 2007; Delany et al. 2009; Alves et al. 2010, 2013).

**Figure 6.** The icelandic black-tailed godwit (*Limosa limosa islandica*). Photo: Sara Pardal

# Studying MHC-I genes

One of the advantages of studying molecular genetic markers in birds is that their red blood cells are nucleated (Glomski and Pica 2011), meaning that by taking a small blood samples (25 microliter) it is possible to get the entire bird genome or in my case one particular type of genes, MHC-I.

Throughout this thesis I have used two different types of samples, genomic DNA (gDNA) and RNA. With genomic DNA it is possible to determine which MHC alleles an individual has in the genome and then by studying RNA it is possible to determine which of these alleles that are expressed (transcribed). RNA is single stranded and DNA is double stranded and the base thymine is exchanged for uracil in RNA (Freeland et al. 2011). In order to study and analyse the RNA it needs to be transformed into complementary DNA (cDNA). This process takes the single stranded RNA and transforms it to double stranded cDNA and also converts the uracil back to thymine.

## Sequencing MHC-I alleles

The MHC-I genes are highly duplicated, especially in passerines, which makes it hard to sequence them (Biedrzycka et al. 2017b). I have mainly worked with just one exon, exon 3, in the MHC-I genes. Since alleles from several different gene copies are co-amplified, the DNA sequence cannot be read from a single Polymerase Chain Reaction (PCR) amplification. An additional method is needed that separates the different alleles and I have used three different methods to sequence the MHC-I alleles.

First, I have used cloning and Sanger sequencing. During cloning the PCR product from amplification of the MHC-I alleles is ligated into a vector and then transformed into E. coli bacteria. The bacteria then grow and each bacterial colony contains a unique DNA sequence. With this method you can separate out the different MHC-I alleles. However, this is a very time-consuming method and hence difficult to use when studying many individuals. Still, cloning and Sanger sequencing is a good starting point when studying MHC-I in new species since you can obtain long DNA sequences spanning up to 800 base pairs, *i.e.* exon 2-4 in MHC-I genes. These long DNA sequence reads will help with primer design in more detailed studies of exon 3.

With the rise of high throughput sequencing (HTS) methods, and more specifically HTS amplicon sequencing, it has become easier to genotype MHC-I in many individuals simultaneously (*e.g.* Strandh et al. 2011; Sepil et al. 2012; O'Connor et al. 2016; Biedrzycka et al. 2017a). There are several different HTS amplicon methods that all in various ways, using PCR amplicon products, determine the specific alleles for each individual. I have used two different techniques; Roche 454 and Illumina Miseq amplicon sequencing. Briefly, you run an amplification with primers that will amplify MHC-I exon 3 and these amplicons are marked with index sequences which will enable you to later on separate out different individuals. The advantage with HTS amplicon sequencing is that during the DNA sequencing all MHC-I exon 3 sequences are separated and individually sequenced. Then, with the help of various bioinformatic tools, it is possible to determine the alleles of each individual, basically counting how many times each MHC-I sequence has been sequenced, also referred to as read depths. Since I have been sequencing both gDNA and cDNA, I can determine which alleles an individual has in the genome and which of these alleles that are expressed. Also, by comparing the read depth of each allele within an individual, I can determine the relative expression of specific MHC-I alleles (*i.e.* high or low expression) but the relative expression cannot be compared between individuals.

Finally, I wanted to determine the relative expression of different MHC-I alleles across individuals. This can be done using RNA sequencing (RNAseq). However, there are certain problems with RNAseq when working with highly duplicated genes like MHC-I. During RNAseq only small fragments are sequenced, in this case 100 base pairs (bp), and in order to determine which genes these transcripts belong to they are either mapped to an existing genome or a *de novo* transcriptome is assembled. But since there are multiple MHC-I genes that are partly similar since they share motifs, there are major problems with using standard methods for de novo assembly. In order to circumvent these problems and determine the relative expression of each MHC-I allele, I combined two different methods, HTS amplicon sequencing and RNAseq. With amplicon sequencing it is possible to determine, with high accuracy, which alleles are expressed, while the relative expression can be determined with RNAseq. Since I aimed to normalize the MHC-I expression relative to the entire transcriptome, I removed all MHC-I exon 3 transcripts within a *de novo* assemble transcriptome and replaced them with the MHC-I exon 3 alleles that were determined with HTS amplicon sequencing. This was done on an individual level so that a unique transcriptome was created for each individual. Then the RNAseq transcripts were mapped to these unique transcriptomes, thereby making it possible to determine the relative expression of each MHC-I allele.

# Results and discussion

## Paper I

### Expression and phylogenetic analyses reveal paralogous lineages of putatively classical and non-classical MHC-I genes in three sparrow species (*Passer*)

The presence of classical and non-classical MHC-I genes has never been confirmed in passerines but have been suggested for house sparrows (Karlsson and Westerdahl 2013). We wanted to further investigate the presence of putatively classical and non-classical MHC-I genes by comparing MHC-I in three closely related sparrow species; house sparrow, Spanish sparrow and tree sparrow.

We compared the relative expression per MHC-I allele within individuals in order to determine if all alleles were expressed to the same degree. We amplified MHC-I with four different primer combinations to estimate the entire MHC-I diversity in each individual and then sequenced exon 3 using 454 amplicon sequencing.

The phylogenetic analysis of the sparrow MHC-I exon 3 sequences revealed one cluster with high bootstrap support (92%) and low diversity that contained alleles from all the tree sparrow species. These findings indicated that the alleles in this cluster are putatively non-classical and that the presence of putatively non-classical alleles pre-dates the speciation events of these tree sparrow species. The number of putatively classical and non-classical alleles varied between the tree sparrow species; house sparrow had on average 4±2 putatively classical alleles and 8±3 putatively non-classical alleles, Spanish sparrow had 5±1 putatively classical alleles and 11±2 putatively non-classical alleles and tree sparrow had 10±3 putatively classical alleles and 5±2 putatively non-classical alleles. About 50% of both putatively classical and non-classical alleles were expressed in all species. Interestingly, the number of expressed classical alleles varied less between species than the gDNA alleles. However, the relative expression per allele within an individual did vary when comparing putatively classical and non-classical alleles. A high variation in degree of expression was found for putatively classical alleles, whereas the variation of expression of putatively non-classical alleles was low, Figure 7. Only few alleles, maximum two putatively classical MHC-I alleles, were ever highly expressed. Overall these results show that three different sparrow species have putatively non-classical alleles and suggests that non-classical alleles could be a common feature also in passerine birds.

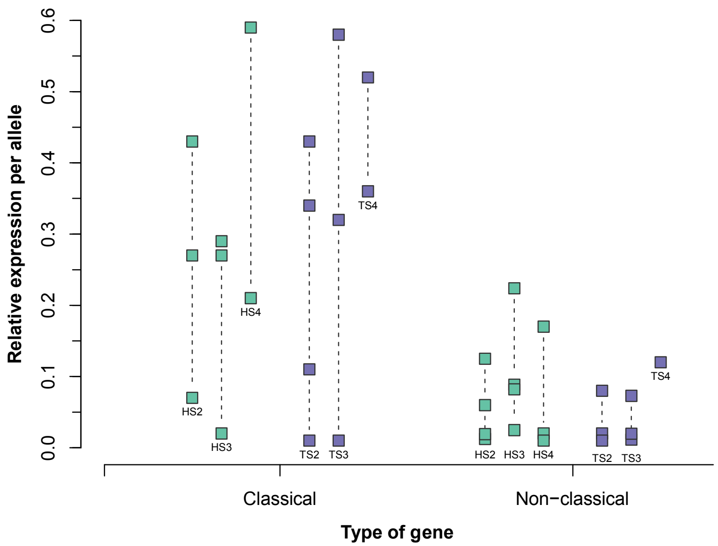



**Figure 7**. Variance in relative read depth per allele of putatively classical and non-classical alleles in three house sparrow (HS2, HS3, and HS4; indicated in green) and three tree sparrow individuals (TS2, TS3 and TS4; indicated in purple). There is a significant difference in variance in expression (measured as relative read depth per allele) between putatively classical and non-classical alleles (Levine'stest: F = 5.20, p =0.005).

# Paper II

## A quantitative and qualitative comparison of illumina MiSeq and 454 amplicon sequencing for genotyping the highly polymorphic major histocompatibility complex (MHC) in a non-model species

There are two main methods used for MHC-I genotyping with HTS, Roche 454 and Illumina Miseq amplicon sequencing (*e.g.* O'Connor et al. 2016; Biedrzycka et al. 2017a). The aim of this study was to compare these two methods.

Overall, we found that although Miseq gives higher read depth the same alleles are identified with the two methods as long as the diversity of the amplicon was rather low, *i.e.* the number of alleles. As long as an individual has fewer than six alleles the two methods identify the same alleles, but if an individual has six alleles or more there is a higher risk that the 454 method will fail in genotyping all alleles correctly. Still, in our data set 98% of the amplicon genotypes were the same irrespectively of HTS method used. We concluded that 454 and Miseq perform equally well in low diversity amplicons but that for high diversity amplicons Miseq outperforms 454 (Figure 8).

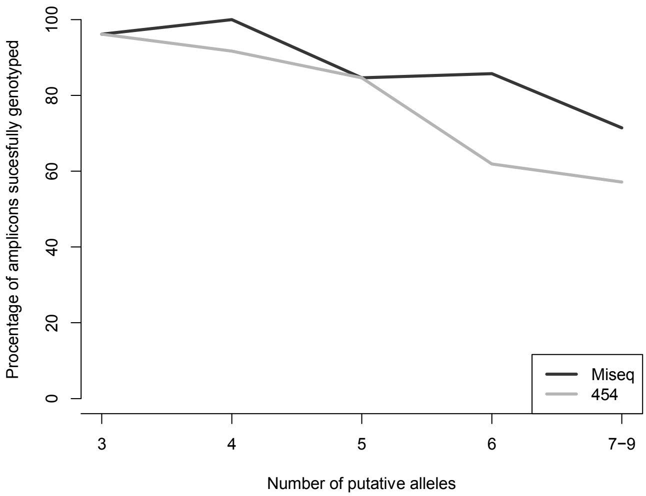



**Figure 8.** The percentage of amplicons successfully genotyped using MiSeq (black) and 454 (grey), dived based on the total number of true alleles.

# Paper III

## Characterization of MHC class I in a long distance migratory wader, the Icelandic black-tailed godwit

In order to study putatively classical and non-classical MHC-I alleles also in a non-passerine bird we decided to partly characterize MHC-I in a Charadriiformes species, the Icelandic black tailed godwit. We used DNA Sanger sequencing of long transcripts (exon 2-4) and then genotyped 84 individuals with Illumina MiSeq amplicon sequencing. In total 47 nucleotide exon 3 alleles were identified and each godwit individual had between two and seven MHC-I alleles. Hence, godwits have lower MHC-I diversity than most passerine birds. Furthermore, the MHC-I diversity in black-tailed godwits was also compared to that of two other Charadriiformes species; the red knot (*Calidris canutus*), and the red-billed gull (*Larus scopulinus)* (Cloutier et al. 2011; Buehler et al. 2013). The diversity and divergence of the black-tailed godwit MHC-I alleles to a large extent fell between the estimates for red knot and red-billed gull. Possibly explained by difference in pathogen pressure, black-tailed godwits are constricted to a more pathogen poor environment compared to red knots, that had the highest MHC-I diversity of the Charadriiformes species.

Interestingly, the occurrence of putatively non-classical alleles has been suggested in both red knot and in red-billed gull, but we found no evidence for non-classical alleles in black-tailed godwit based on comparative phylogenetic analysis across the three Charadriiformes species, Figure 9. One explanation would be that non-classical genes may not be fixed in all bird species and that just because a few species within an order have non-classical genes it is not given than all species belonging to that order have non-classical genes. An alternatively explanation would be that non-classical genes do exist in black-tailed godwits but are more distant so that we did not co-amplify them when sequencing classical genes.

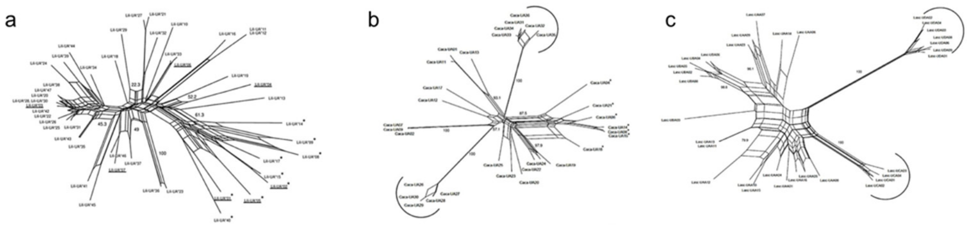



**Figure 9.** Neighbour-net phylogenetic networks of MHC-I exon 3 nucleotide alleles from a) Icelandic black-tailed godwits (*Limosa limosa islandica*), b) red knots (*Calidris canutus*; Buehler et al. 2013), and c) red-billed gulls (*Larus scopulinus*; Cloutier et al. 2011). Non-classical alleles form distinct clusters in red-billed gulls and red knots, indicated with brackets. No such cluster were found in black-tailed godwits.

# Paper IV

## Characterization of putatively classical and non-classical MHC-I genes in the Eurasian siskin (*Spinus spinus*), but no evidence of a single majorly expressed MHC-I gene

The Eurasian siskin is a small passerine bird. In this study we partly characterized MHC-I with Sanger sequencing followed by genotyping MHC-I in 18 individuals, as well as determining the relative expression of these MHC-I alleles with Illumina Miseq using three different primer combinations. In total 88 MHC-I exon 3 alleles were identified in the 18 individuals and out of these 84 were expressed. Furthermore, the expression level could be classified as high or low for 80 alleles. When a phylogenetic analysis was combined with degree of expression, one significant cluster with high support (bootstrap: 987) which only contained low expression alleles was revealed,

Figure 10. The alleles in this cluster also had lower genetic diversity (number of nucleotide alleles, number of amino acid alleles, amino acid P-distance and number of positively selected sites). Strikingly, no sites were found to be under positive selection in this cluster, whereas six sites where found to be under positive selection among the rest of the alleles.

The number of putatively classical and non-classical MHC-I alleles in siskins was 8±2 and 3±1, respectively. Interestingly, almost all alleles were expressed, 97% of putatively classical alleles and 89% of putatively non-classical alleles, indicating a different expression pattern than previously described in sparrows. This difference became even more evident when comparing the relative expression of putatively classical and non-classical alleles within siskin individuals. Similar to sparrows, there was large variation in degree of expression for putatively classical alleles and low variation in degree of expression of non-classical alleles. However, unlike sparrows, where 1-2 alleles were highly expressed, on average five alleles were highly expressed per individual in siskins. These results suggest that siskins also have both classical and non-classical alleles and indicates that more than one classical MHC-I gene have high expression.

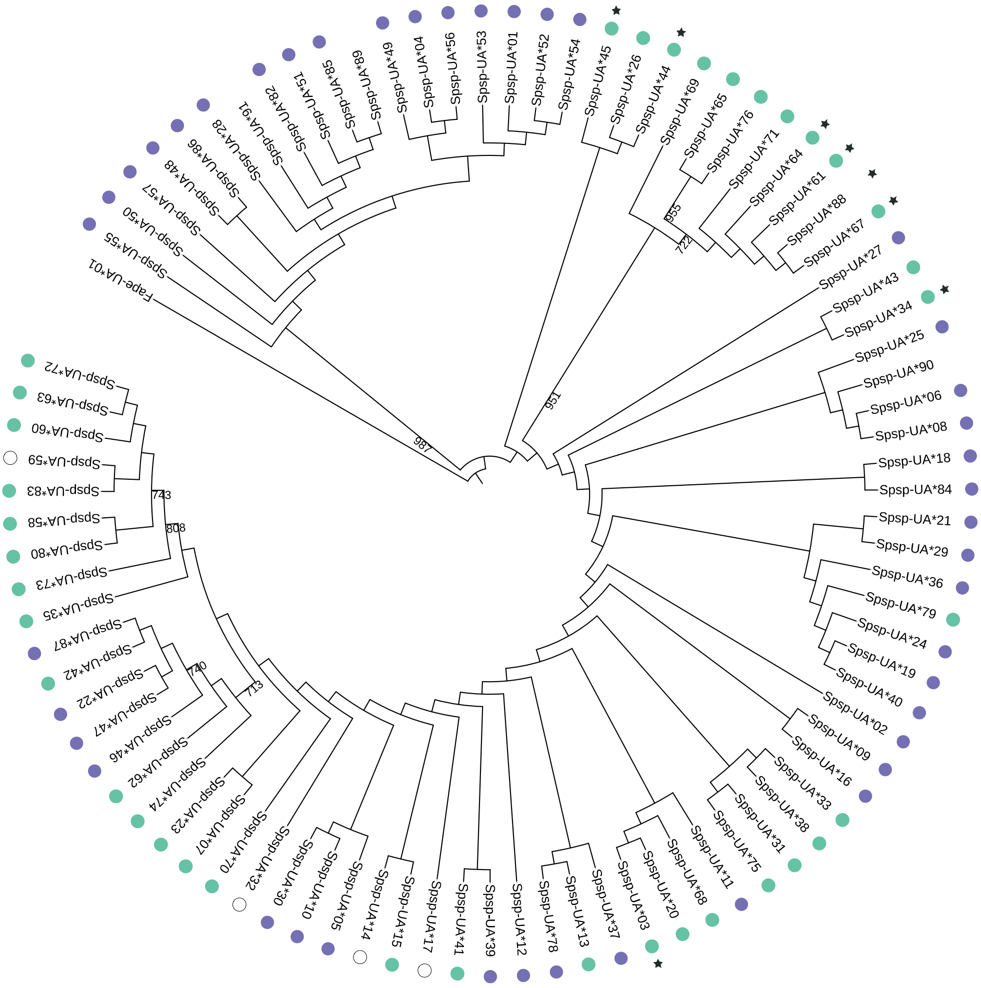



**Figure 10.** Maximum likelihood tree based on 88 MHC class I exon 3 nucleotide sequences from siskins. Green circles represent alleles with high expression, purple circles represent alleles with low expression, white circles represent alleles with undetermined expression and alleles without circles are not expressed. Stars (*) indicates alleles that had a 3 bp insertion. All putatively non-classical alleles (N=18) are found in one cluster with strong bootstrap support (987). Branch length is unscaled, *i.e.* all branches are the same length and bootstarp values higher than 700 are displayed.

# Paper V

## Expression of MHC-I genes during mild and severe avian malaria infections in Eurasian siskins (*Spinus spinus*)

The function of MHC-I is to enable cytotoxic T-cells to recognize intracellular pathogens (Murphy and Weaver 2016). The expression of MHC-I could therefore be expected to vary throughout an infection with such pathogens. To study this in more detail, siskins were infected with avian malaria parasites. Avian malaria is a common disease in passerines and as with other malaria parasites avian malaria parasites infect and multiply in red blood cells (Valkiūnas 2005). Since red blood cells of birds are nucleated, (Glomski and Pica 2011), it is possible that these cells can express MHC-I as a response to malaria infection.

Siskins were randomly assigned into three different treatment groups, infected with a severe malaria strain (SGS1; n=9), infected with a mild malaria strain (GRW4; n=5) or placed in a control group (uninfected individuals, n=5) and then blood samples were collected at four times throughout the infection period (day 0, 8, 20, 36 post inoculation). The relative expression of MHC-I was determined by combining Illumina amplicon sequencing and Illumina RNAseq. The advantage of this method is that the relative expression per allele could be determined throughout the infection period. Since siskins have both putatively classical and non-classical MHC-I genes we could determine the change in expression of both types of MHC-I genes throughout an infection period.

Classical MHC genes were continuously expressed to a higher degree than non-classical genes, but interestingly both types of genes responded to the malaria infections in a similar way (figure 11). Overall, infected and uninfected individuals expressed MHC-I differently at the different time points during the infection and there was a tendency ($p=0.06$) for MHC-I to be more highly expressed at day 20, just after the peak of infection, in individuals infected with the severe malaria lineage. Since only a week response of MHC-I gene expression in blood was found during the avian malaria infection, further studies are needed in order to determine the full importance of MHC-I expression during the blood-stages of avian malaria parasites in passerine birds.

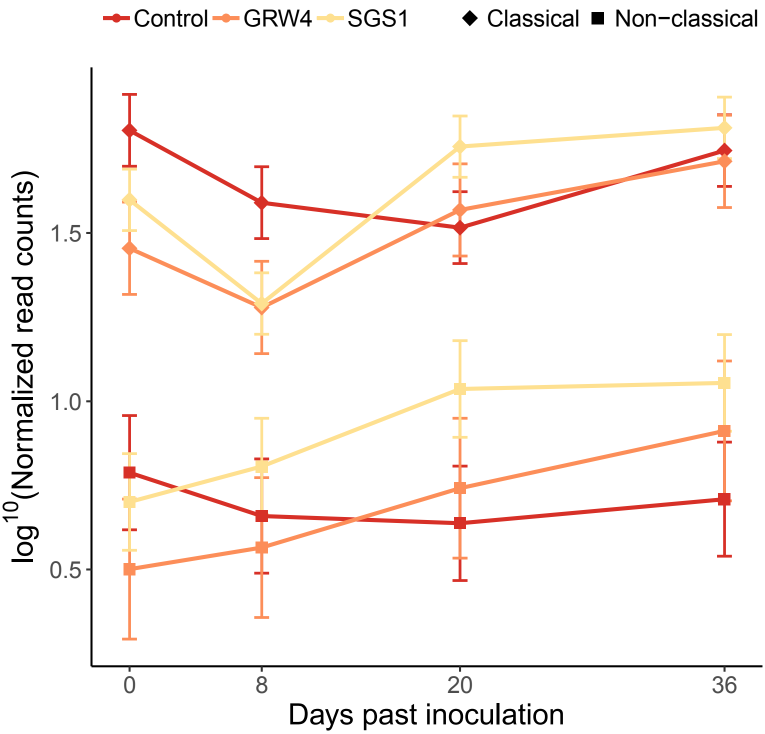



**Figure 11.** Average MHC-I expression shown as least-squares mean ±SE from repeated-measures analysis for the three different treatments, SGS1, GRW4 and control, and separately for classical and non-classical alleles.

# Conclusions

In this thesis I have examined the high diversity of MHC-I in passerine birds from two aspects, the presence of putatively classical and non-classical genes and the expression of classical genes.

I have concluded that putatively non-classical MHC-I genes seem to be a common feature in passerine birds, as all four species examined here had putatively non-classical genes. These non-classical genes could be recognized as having low genetic diversity, having low expression and forming a highly supported cluster in a phylogenetic analysis. Interestingly, the proportion of putatively classical and non-classical genes varied between the passerine species. In house sparrow and Spanish sparrow most genes identified were non-classical, whereas in tree sparrow and siskin most genes were classical. This indicates that although non-classical genes could be present in most passerine species it alone cannot explain the high diversity seen in passerines. However, it is an aspect of passerine MHC-I that should be examined before other ecological and evolutionary studies are carried out.

When a non-passerine species (a wader, the black-tailed godwit) was examined, no putatively non-classical genes were identified although putatively non-classical genes had previously been found in closely related species. This lack of non-classical genes could mean that black-tailed godwits simply do not have them or alternatively that their non-classical genes are too diverged from classical genes to be co-amplified. In order to determine if non-classical MHC-I genes are found in most bird species further studies are needed.

In most bird species studied to date only one MHC-I gene is highly expressed. Here, I have, for the first time, tried to determine the expression levels of putatively classical genes in passerine birds. Within passerines the number of highly expressed genes seem to vary between species, in house sparrow and tree sparrow only one gene seems to be highly expressed but in siskins three genes were identified as having high expression. This suggests that the number of highly expressed genes varies between species. If more than one MHC-I gene copy is highly expressed in passerines, it could help to explain the high diversity found in passerine birds.

How classical and non-classical genes are expressed throughout an avian malaria infection was also examined and interestingly the expression of both types changed throughout the infection period which could mean that non-classical genes also are important during a malaria infection, however, more studies are needed.

To summarize I have shown that the presence of both classical and non-classical genes as well as how these genes are expressed are important aspects of passerine MHC-I diversity and should be considered when studying MHC-I in passerines.

# References


Abbas AK, Lichtman AH, Pillai S (2016) Basic immunology: functions and disorders of the immune system. Elsevier, St. Louis, Missouri

Adams E, Parham P (2001) Species-specific evolution of MHC class I genes in the higher primates. Immunol Rev 183:41–64. doi: 10.1038/icb.1996.62

Afanassieff M, Goto RM, Ha J, et al (2001) At least one class I gene in restriction fragment pattern-Y (Rfp-Y), the second MHC gene cluster in the chicken, is transcribed, polymorphic, and shows divergent specialization in antigen binding region. J Immunol 166:3324–33. doi: 10.4049/jimmunol.166.5.3324

Allen RL, Hogan L (2013) Non-Classical MHC Class I Molecules (MHC-Ib). eLS 1–12. doi: 10.1002/9780470015902.a0024246

Altuvia Y, Margalit H (2004) A structure-based approach for prediction of MHC-binding peptides. Methods 34:454–459. doi: 10.1016/j.ymeth.2004.06.008

Alves JA, Gunnarsson TG, Hayhow DB, et al (2013) Costs, benefits, and fitness consequences of different migratory strategies. Ecology 94:11–17. doi: 10.1890/12-0737.1

Alves JA, Lourenço PM, Piersma T, et al (2010) Population overlap and habitat segregation in wintering Black-tailed Godwits Limosa limosa. Bird Study 57:381–391. doi: 10.1080/00063651003678475

Apps R, Meng Z, Del Prete GQ, et al (2015) Relative Expression Levels of the HLA Class-I Proteins in Normal and HIV-Infected Cells. J Immunol 194:3594–3600. doi: 10.4049/jimmunol.1403234

Balakrishnan CN, Ekblom R, Völker M, et al (2010) Gene duplication and fragmentation in the zebra finch major histocompatibility complex. BMC Biol 8:29. doi: 10.1186/1741-7007-8-29

Biedrzycka A, O'Connor E, Sebastian A, et al (2017a) Extreme MHC class i diversity in the sedge warbler (Acrocephalus schoenobaenus); Selection patterns and allelic divergence suggest that different genes have different functions. BMC Evol Biol 17:1–12. doi: 10.1186/s12862-017-0997-9

Biedrzycka A, Sebastian A, Migalska M, et al (2017b) Testing genotyping strategies for ultra-deep sequencing of a co-amplifying gene family: MHC class I in a passerine bird. Mol Ecol Resour 17:642–655. doi: 10.1111/1755-0998.12612

birdtree.org A Global Phylogeny of Birds. https://birdtree.org/

Bonneaud C, Sorci G, Morin V, et al (2004) Diversity of Mhc class I and IIB genes in house sparrows (Passer domesticus). Immunogenetics 55:855–865. doi: 10.1007/s00251-004-0648-3

Briles WE, Goto RM, Auffray C, Miller MM (1993) A polymorphic system related to but genetically independent of the chicken major histocompatibility complex. Immunogenetics 37:408–414. doi: 10.1007/BF00222464

Buehler DMDM, Verkuil YIYI, Tavares ESESES, Baker AJAJAJ (2013) Characterization of MHC class i in a long-distance migrant shorebird suggests multiple transcribed genes and intergenic recombination. Immunogenetics 65:211–225. doi: 10.1007/s00251-012-0669-2

Chan WF, Parks-Dely JA, Magor BG, Magor KE (2016) The Minor MHC Class I Gene UDA of Ducks Is Regulated by Let-7 MicroRNA. J Immunol 197:1212–20. doi: 10.4049/jimmunol.1600332

Chaves LD, Krueth SB, Reed KM (2009) Defining the turkey MHC: sequence and genes of the B locus. J Immunol 183:6530–6537. doi: 10.4049/jimmunol.0901310

Chen L-C, Lan H, Sun L, et al (2015) Genomic organization of the crested ibis MHC provides new insight into ancestral avian MHC structure. Sci Rep 5:7963. doi: 10.1038/srep07963

Cloutier A, Mills JAJAJ a., Baker AJAJAJ (2011) Characterization and locus-specific typing of MHC class i genes in the red-billed gull (Larus scopulinus) provides evidence for major, minor, and nonclassical loci. Immunogenetics 63:377–394. doi: 10.1007/s00251-011-0516-x

Cruse M, Lewis R (2003) Atlas of Immunology, Second Edi. CRC Press, Baton Rouge

Crux NB, Elahi S (2017) Human Leukocyte Antigen (HLA) and immune regulation: How do classical and non-classical HLA alleles modulate immune response to human immunodeficiency virus and hepatitis C virus infections? Front Immunol 8:1–26. doi: 10.3389/fimmu.2017.00832

Delany S, Scott D, Dodman T, Stroud D (2009) An atlas of wader populations in Africa and western Eurasia. Wageningen : Wetlands International

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

Eirín-López JM, Rebordinos L, Rooney AP, Rozas J (2012) The Birth-and-Death Evolution of Multigene Families Revisited. In: Repetitive DNA. pp 170–196

Ekblom R, Sæther SA, Jacobsson P, et al (2007) Spatial pattern of MHC class II variation in the great snipe (Gallinago media). Mol Ecol 16:1439–1451. doi: 10.1111/j.1365-294X.2007.03281.x

Ekblom R, Stapley J, Ball AD, et al (2011) Genetic mapping of the major histocompatibility complex in the zebra finch (Taeniopygia guttata). Immunogenetics 63:523–30. doi: 10.1007/s00251-011-0525-9

Fleming-Canepa X, Jensen SM, Mesa CM, et al (2016) Extensive Allelic Diversity of MHC Class I in Wild Mallard Ducks. J Immunol 197:783–94. doi: 10.4049/jimmunol.1502450

Freeland JR, Kirk H, Petersen S (2011) Molecular ecology. Oxford ; Hoboken, NJ : Wiley-Blackwell, 2011

Garcia-Beltran WF, Hölzemer A, Martrus G, et al (2016) Open conformers of HLA-F are high-affinity ligands of the activating NK-cell receptor KIR3DS1. Nat Immunol 17:1067–1074. doi: 10.1038/ni.3513

Gill JA, Langston RHW, Alves JA, et al (2007) Contrasting trends in two Black-tailed Godwit populations : a review of causes and recommendations. Wader Study Gr Bull 43–50

Glomski AC, Pica A (2011) The avian erythrocyte: its phylogenetic odyssey. CRC Press, Boca Raton

Gorer PA (1936a) The detection of a hereditary antigenic difference in the blood of mice by means of human group a serum. J Genet 32:17. doi: 10.1007/BF02982499

Gorer PA (1936b) The Detection of Antigenic Differences in Mouse Erythrocytes by the Employment of Immune Sera. Br J Exp Pathol 17:42–50

Goulder PJR, Walker BD (2012) Review HIV and HLA Class I : An Evolving Relationship. Immun Rev 37:426–440. doi: 10.1016/j.immuni.2012.09.005

Grimholt U, Tsukamoto K, Azuma T, et al (2015) A comprehensive analysis of teleost MHC class I sequences. BMC Evol Biol 15:15–32. doi: 10.1186/s12862-015-0309-1

Hedrick PW (2007) Balancing selection. Curr Biol 17:230–231. doi: 10.1016/j.cub.2007.01.012

Hedrick PW (2002) Pathogen resistance and genetic variation at MHC loci. Evolution 56:1902–1908

Hedrick PW (2012) What is the evidence for heterozygote advantage selection ? Trends Ecol Evol 27:698–704. doi: 10.1016/j.tree.2012.08.012

Hedrick PW, Thomson G (1983) Evidence for balancing selection at HLA. Genetics 104:449–56

Hughes ALL, Nei M (1989) Evolution of the major histocompatibility complex: independent origin of nonclassical class I genes in different groups of mammals. Mol Biol Evol 6:559–79

Hunt HD, Goto RM, Foster DN, et al (2006) At least one YMHCI molecule in the chicken is alloimmunogenic and dynamically expressed on spleen cells during development. Immunogenetics 58:297–307. doi: 10.1007/s00251-005-0074-1

Jetz W, Thomas GH, Joy JB, et al (2012) The global diversity of birds in space and time. Nature 491:444–448. doi: 10.1038/nature11631

Karlsson M, Westerdahl H (2013) Characteristics of MHC Class I Genes in House Sparrows Passer domesticus as Revealed by Long cDNA Transcripts and Amplicon Sequencing. J Mol Evol 77:8–21. doi: 10.1007/s00239-013-9575-y

Kaufman J (1999) Co-evolving genes in MHC haplotypes: the “rule” for nonmammalian vertebrates? Immunogenetics 50:228–36. doi: 10.1007/s002510050597

Kaufman J (2015a) What chickens would tell you about the evolution of antigen processing and presentation. Curr Opin Immunol 34:35–42. doi: 10.1016/j.coi.2015.01.001

Kaufman J (2015b) Co-evolution with chicken class I genes. Immunol Rev 267:56–71. doi: 10.1111/imr.12321

Kaufman J, Jacob J, Shaw I, et al (1999a) Gene organisation determines evolution of function in the chicken MHC. Immunol Rev 167:101–117. doi: DOI: 10.1111/j.1600-065X.1999.tb01385.x

Kaufman J, Milne S, Göbel TW, et al (1999b) The chicken B locus is a minimal essential major histocompatibility complex. Nature 401:923–925. doi: 10.1038/44856

Kaufman J, Völk H, Wallny H-JJ, et al (1995) A “minimal essential Mhc” and an “unrecognized Mhc”: Two extremes in selection for polymorphism. Immunol Rev 143:63–88. doi: 10.1111/j.1600-065X.1995.tb00670.x

Kelley J, Walter L, Trowsdale J (2005) Comparative genomics of major histocompatibility complexes. Immunogenetics 56:683–695. doi: 10.1007/s00251-004-0717-7

Larsen MH, Hylenius S, Andersen A-MN, Hviid TVF (2010) The 3 -untranslated region of the HLA-G gene in relation to pre-eclampsia: revisited. Tissue Antigens 75:253–261. doi: 10.1111/j.1399-0039.2009.01435.x

Le Bouteiller P, Lenfant F (1996) Antigen-presenting function(s) of the non-classical HLA-E, -F and -G class I molecules: The beginning of a story. Res Immunol 147:301–313. doi: 10.1016/0923-2494(96)89643-X

Lee N, Geraghty DE (2003) HLA-F surface expression on B cell and monocyte cell lines is partially independent from tapasin and completely independent from TAP. J Immunol 171:5264–5271. doi: 10.4049/jimmunol.171.10.5264

Lee N, Llano M, Carretero M, et al (1998) HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. Proc Natl Acad Sci USA 95:5199–5204. doi: 10.1073/pnas.95.9.5199

Lenz TL, Mueller B, Trillmich F, et al (2013) Divergent allele advantage at MHC-DRB through direct and maternal genotypic effects and its consequences for allele pool composition and mating

Lepin EJM, Bastin JM, Allan DSJ, et al (2000) Functional characterization of HLA-F and binding of HLA-F tetramers to ILT2 and ILT4 receptors. Eur J Immunol 30:3552–3561. doi: 10.1002/1521-4141(200012)30:12<3552::AID-IMMU3552>3.0.CO;2-L

Lunney JK, Ho CS, Wysocki M, Smith DM (2009) Molecular genetics of the swine major histocompatibility complex, the SLA complex. Dev Comp Immunol 33:362–374. doi: 10.1016/j.dci.2008.07.002

Mesa CM, Thulien KJ, Moon D a, et al (2004) The dominant MHC class I gene is adjacent to the polymorphic TAP2 gene in the duck, Anas platyrhynchos. Immunogenetics 56:192–203. doi: 10.1007/s00251-004-0672-3

Miller MM, Goto R, Bernot a, et al (1994) Two Mhc class I and two Mhc class II genes map to the chicken Rfp-Y system outside the B complex. Proc Natl Acad Sci U S A 91:4397–4401

Miller MM, Goto RM, Taylor RL, et al (1996) Assignment of Rfp-Y to the chicken major histocompatibility complex NOR microchromosome and evidence for high-frequency recombination associated with the nucleolar organizer region. Proc Natl Acad Sci U S A 93:3958–3962

Moon DA, Veniamin SM, Parks-Dely JA, Magor KE (2005) The MHC of the duck (Anas platyrhynchos) contains five differentially expressed class I genes. J Immunol 175:6702–6712. doi: 10.4049/jimmunol.175.10.6702

Mullarney K, Svensson L, Zetterström D, Grant PJ (2006) Bird guide. The most complete field guide to the birds of britain and europe. HarperCollins Publishers Ltd, London

Murphy K, Weaver C (2016) Janeway's Immunobiology, 9th edn. Garland Science, New York

Neefjes J, Jongsma ML, Paul P, Bakke O (2011) Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat Rev Immunol 11:823–836. doi: 10.1038/nri3084

Nei M, Gu X, Sitnikova T (1997) Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci 94:7799–7806

O'Connor EA, Strandh M, Hasselquist D, et al (2016) The evolution of highly variable immunity genes across a passerine bird radiation. Mol Ecol 25:977–989. doi: 10.1111/mec.13530

Ohno S (1970) Evolution by Gene Duplication. Springer Berlin Heidelberg, Berlin, Heidelberg, Heidelberg

Pierini F, Lenz TL (2018) Divergent Allele Advantage at Human MHC Genes: Signatures of Past and Ongoing Selection. Mol Biol Evol 35:2145–2158. doi: 10.1093/molbev/msy116

Robinson J, Halliwell JA, Hayhurst JD, et al (2015) The IPD and IMGT/HLA database: Allele variant databases. Nucleic Acids Res 43:D423-31. doi: 10.1093/nar/gku1161

Rodgers JR, Cook RG (2005) MHC class Ib molecules bridge innate and acquired immunity. Nat Rev Immunol 5:459–471. doi: 10.1038/nri1635

Roved J, Hansson B, Tarka M, et al (2018) Evidence for sexual conflict over major histocompatibility complex diversity in a wild songbird. Proc R Soc B Biol Sci 285:. doi: 10.1098/rspb.2018.0841

Schut E, Aguilar JRJR De, Merino S, et al (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. doi: 10.1007/s00251-011-0532-x

Sepil I, Lachish S, Hinks AE, Sheldon BC (2013) Mhc supertypes confer both qualitative and quantitative resistance to avian malaria infections in a wild bird population. Proc Biol Sci 280:20130134. doi: 10.1098/rspb.2013.0134

Sepil I, Moghadam HK, Huchard E, et al (2012) Characterization and 454 pyrosequencing of Major Histocompatibility Complex class I genes in the great tit reveal complexity in a passerine system. BMC Evol Biol 12:68. doi: 10.1186/1471-2148-12-68

Shaw I, Powell TJTJ, Marston DADD a, et al (2007) Different evolutionary histories of the two classical class I genes BF1 and BF2 illustrate drift and selection within the stable MHC haplotypes of chickens. J … 178:5744–5752. doi: 10.4049/jimmunol.178.9.5744

Shawar S, Vyas J (1994) Antigen presentation by major histocompatibility complex class IB molecules. Annu Rev Immunol 12:839–880. doi: 10.1146/annurev.iy.12.040194.004203

Shiina T, Hosomichi K, Hanzawa K (2006) Comparative genomics of the poultry major histocompatibility complex. Anim Sci J 77:151–162. doi: 10.1111/j.1740-0929.2006.00333.x

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

Shiina T, Shimizu C, Oka A, et al (1999) Gene organizsation of the quail major histocompatibility complex (MhcCoja) class I gene region. Immunogentics 49:384–394. doi: 10.1007/s002510050511

Shiina T, Shimizu S, Hosomichi K, et al (2004) Comparative genomic analysis of two avian (quail and chicken) MHC regions. J Immunol 172:6751–63. doi: 10.4049/jimmunol.172.11.6751

Slade RW, McCallum HI (1992) Overdominant vs. Frequency-Dependent Selection at MHC Loci. Genetics 132:861–862. doi: 10.1097/SLA.0b013e318230a1b4

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. doi: 10.1007/s00251-011-0534-8

The HLA Database HLA Alleles Numbers. http://hla.alleles.org/nomenclature/stats.html. Accessed 25 Oct 2018

Valkiūnas G (2005) Avian malaria parasites and other haemosporidia. CRC Press, Boca Raton

Van Oosterhout C (2009) A new theory of MHC evolution: Beyond selection on the immune genes. Proc R Soc B Biol Sci 276:657–665. doi: 10.1098/rspb.2008.1299

Vianna P, Dalmáz CA, Veit TD, et al (2007) Immunogenetics of pregnancy: Role of a 14-bp deletion in the maternal HLA-G gene in primiparous pre-eclamptic Brazilian women. Hum Immunol 68:668–674. doi: 10.1016/j.humimm.2007.05.006

Wakeland E, Boehme S, She J, et al (1990) Ancestral polymorphisms of Mhc class-II genes—divergent allele advantage. Immunol Res 9:115–122

Walker BA, Hateren A Van, Milne S, et al (2005) Chicken TAP genes differ from their human orthologues in locus organisation , size , sequence features and polymorphism. 232–247. doi: 10.1007/s00251-005-0786-2

Walker BA, Hunt LG, Sowa AK, et al (2011) The dominantly expressed class I molecule of the chicken MHC is explained by coevolution with the polymorphic peptide transporter (TAP) genes. Proc Natl Acad Sci 108:8396–8401. doi: 10.1073/pnas.1019496108

Wallny H-J, Avila D, Hunt LG, et al (2006) Peptide motifs of the single dominantly expressed class I molecule explain the striking MHC-determined response to Rous sarcoma virus in chickens. Proc Natl Acad Sci U S A 103:1434–1439. doi: 10.1073/pnas.0507386103

Wang B, Ekblom R, Strand TM, et al (2012) Sequencing of the core MHC region of black grouse (Tetrao tetrix) and comparative genomics of the galliform MHC. BMC Genomics 13:553. doi: 10.1186/1471-2164-13-553

Wei X, Orr HT (1990) Differential expression of HLA-E, HLA-F, and HLA-G transcripts in human tissue. Hum Immunol 29:131–142. doi: 10.1016/0198-8859(90)90076-2

Westerdahl H, Wittzell H, von Schantz T (1999) Polymorphism and transcription of Mhc class I genes in a passerine bird, the great reed warbler. Immunogenetics 49:158–170

Westerdahl H, Wittzell H, von Schantz T (2000) Mhc diversity in two passerine birds: no evidence for a minimal essential Mhc. Immunogenetics 52:92–100. doi: 10.1007/s002510000256

Williams A (2011) Immunology: mucosal and body surface defences. John Wiley & Sons, Chichester, West Sussex ; Hoboken, NJ

Zeng Q, Zhong G, He K, et al (2016) Molecular characterization of classical and nonclassical MHC class I genes from the golden pheasant (Chrysolophus pictus ). Immunogenetics 43:8–17. doi: 10.1111/iji.12245

# Acknowledgments

There are many people that have helped and inspired me during my time as a PhD-student and I could not have finished it without you. One of the things that I have considered as an extra bonus during my time as a PhD-student is the fact that I got to meet so many wonderful people from all over the world.

**Helena**, my supervisor, I heard one of your lectures on MHC during my master program and though it sounded really interesting and that you seemed like a good supervisor so I contacted you in order to do a master project on MHC which in the end turned out to be the starting point of my PhD. I still find MHC interesting although, sometimes complicated to work with and that you are a good supervisor. I have always felt lucky that you were my supervisor. Thank you for everything you have done for me over these last years. You have encouraged me, inspired me and pushed me when needed. I could not have asked for a better supervisor.

**Lars**, my assistant supervisor, thank you for all your help during my PhD work. Especially, now towards then end. Thank you for reading things in the middle of the night and for all your valuable comments, feedback and discussion. You have almost convinced me that repeated measures analysis is fun!

I ended up having two examiners, **Thomas and Dennis,** thank you both for all you have done; for taking an interest in my project, for asking me how things are going and for making sure things are running smoothly

**Björn,** thank you for agreeing to be my scientific mentor and being part of my committee.

Just after I started, I joined the GENECO mentoring program and was asked if I had any mentors and at the time I could not think of any but now looking back I realized I had two. **Maria**, thank you for being such an inspiration. You are always happy and always have some nice and cool new result to show and discuss and I have enjoyed this so much! **Emily**, you have also been an inspiration for me. Thank you for always taking the time to discuss things with me, I am not sure if I would have figured out how to filter MHC data when I first started without you. And thank you for always answering all my English questions.

**MEEL** have been a great group to be a part of and thank you to all the wonderful people in it, past and present, for making it such a warm and welcoming environment!

Thank you to all my **colleagues** in the Ecology building for making it such an inspiring and encouraging place to work.

The **GENECO** research school was a great opportunity for me  to interact with other PhD students and seniors to discuss science. A special thanks to all the wonderful people in my GENECO mentor cohort.

Thank you to all the people at Linnaeus university in Kalmar for making me feel very welcome when I visited you.

To all my fellow PhD students: **Elin**, **Jacob**, **Xi**, **Hanna**, **Julian**, **Pablo**, **Fredrick**, **Dafne**, **Ann-Kathrine**, **Lokesh**, **Marco**, **Katrine**, **Anna**, **Beatriz**, **John**, **Utku, Kajsa, Linus, Gabriel, Jothi** thank you for making my PhD studies such a fun time!

There are many people I would like to thank that have helped me one way or another throughout my PhD studies and to mention just a few of you:

**Elin**, my wonderful friend and office mate for most of my PhD period. Thank for all your support and encouragement and for all the fun times, both at work and outside! I feel lucky that we could do our PhD work at the same time, it's been a journey with peaks and valleys and you have meant so much to me through all of it!

**Jane**, thank you for everything! For teaching me about lab-work when I started, for always listening to me, for always asking how I am, for all the things we have done both at work and outside, simply for being such a good friend!

**Anu**, you are amazing! Thank you for always being there and always helping. I feel lucky to count you as my friend.

**Anna**, you always bring so much good energy with you whenever I meet you. Thank you for being you and for your support! I still cannot believe that it took for us both to start working at the second floor for us to meet but I'm glad we did.

**Luz**, thank you for all your encouragement and all the fun time. And of course, for teaching me how to say: gorrión!

**Sara**, you and I struggled to figure out that first Illumina run, it was nice not to have to do that one my own and in the end it all turned out fine. Thank you for everything you have done!

**Olof**, thank you for the siskin samples, and for all the discussions about MHC and malaria.

**Max**, thank you for slowly and carefully explaining bioinformatics to me

**Vaidas**, thank you for all your help and answering all my malaria questions.

**Inger,** thank you for the nice picture showing antigen presentation.

**Fredrik H,** thank you for helping us to catch the sparrows

**Julio**, thank you for all the wonderful sparrow pictures.

And finally, my amazing **family**, near and far, thank you for all your support and encouragement and for simply always being there for me whatever I need!

# List of papers

I. **Drews A**, Strandh M, Råberg L and Westerdahl H. (2017) Expression and phylogenetic analyses reveal paralogous lineages of putatively classical and non-classical MHC-I genes in three sparrow species (*Passer*). *BMC Evolutionary Biology* 17: 152

II. Razali H, O'Connor E, **Drews A**, Burke T and Westerdahl H. (2017) A quantitative and qualitative comparison of illumina MiSeq and 454 amplicon sequencing for genotyping the highly polymorphic major histocompatibility complex (MHC) in a preformed non-model species. *BMC Research Notes* 10: 346

III. Pardal S, **Drews A**, Alves JA., Ramos JA., Westerdahl H. (2017) Characterization of MHC class I in a long distance migratory wader, the Icelandic black-tailed godwit. *Immunogenetics* 69(7): 463-478

IV. **Drews A** and Westerdahl H. Characterization of putatively classical and non-classical MHC-I genes in the Eurasian siskin (*Spinus spinus*): no evidence of a single majorly expressed MHC-I gene. Manuscript

V. **Drews A**, Råberg L, Hellgren O, Lundberg M, Palinauskas V and Westerdahl H. Expression of MHC-I genes during mild and severe avian malaria infections in Eurasian siskins (*Spinus spinus*). Manuscript

Printed by Media-Tryck, Lund 2018 NORDIC SWAN ECOLABEL 3041 0903

Lund University

Department of Biology
Faculty of Science
Lund University

ISBN 978-91-7753-906-3
ISSN 1652-8220