OULU 2008
A 514

ACTA UNIVERSITATIS OULUENSIS

Lumi Viljakainen

EVOLUTIONARY GENETICS OF IMMUNITY AND INFECTION IN SOCIAL INSECTS

FACULTY OF SCIENCE,
DEPARTMENT OF BIOLOGY,
BIOCENTER OULU,
UNIVERSITY OF OULU

ACTA UNIVERSITATIS OULUENSIS
A Scientiae Rerum Naturalium 514

LUMI VILJAKAINEN

EVOLUTIONARY GENETICS OF IMMUNITY AND INFECTION IN SOCIAL INSECTS

Academic dissertation to be presented, with the assent of the Faculty of Science of the University of Oulu, for public defence in Kuusamonsali (Auditorium YB210), Linnanmaa, on November 21st, 2008, at 12 noon

OULUN YLIOPISTO, OULU 2008

Acta Univ. Oul. A 514, 2008

Supervised by
Professor Pekka Pamilo
Associate Professor Jürgen Gadau

Reviewed by
Professor Ross H. Crozier
Professor Anneli Hoikkala

ISBN 978-951-42-8927-9 (Paperback)
ISBN 978-951-42-8928-6 (PDF)
http://herkules.oulu.fi/isbn978951428928 6/
ISSN 0355-3191 (Printed)
ISSN 1796-220X (Online)
http://herkules.oulu.fi/issn03553191/

Cover design
Raimo Ahonen

OULU UNIVERSITY PRESS
OULU 2008

Viljakainen, Lumi, Evolutionary genetics of immunity and infection in social insects

Faculty of Science, Department of Biology, University of Oulu, P.O.Box 3000, FI-90014 University of Oulu, Finland; Biocenter Oulu, University of Oulu, P.O.Box 5000, FI-90014 University of Oulu, Finland

Acta Univ. Oul. A 514, 2008

Oulu, Finland

Abstract

In social insects a major cost of social life is the high number of pathogens found in large societies and the greater likelihood of transmission of pathogens among closely related individuals. The aim of this thesis was to investigate the effect of high pathogen pressure on the molecular evolution of genes involved in the innate immune system in social insects. In addition, the transmission dynamics of the intracellular bacteria *Wolbachia* in wood ants was examined.

By comparing DNA sequences from diverse species of ants and honeybees it was shown that the immune genes in hymenopteran social insects have evolved rapidly. However, by using codon-based likelihood models of evolution positive selection was detected in only two ant genes. This may reflect behaviourally based colony-level defences that can reduce selective pressure on the immune genes.

The transmission modes of *Wolbachia* were studied by comparing DNA sequence variation of the bacteria with that of the host ants. First, it was found that all the studied ants carry *Wolbachia*. Second, *Wolbachia* have been transmitted both vertically from mother to offspring and horizontally between individuals of the same as well as of different species.

Keywords: ants, honeybees, insect immunity, molecular evolution, positive selection, Wolbachia

Viljakainen, Lumi, Yhteiskuntahyönteisten immuunipuolustuksen ja bakteeri-infektioiden evoluutiogeneettinen tutkimus
Luonnontieteellinen tiedekunta, Biologian laitos, Oulun yliopisto, PL 3000, 90014 Oulun yliopisto; Biocenter Oulu, Oulun yliopisto, PL 5000, 90014 Oulun yliopisto
Acta Univ. Oul. A 514, 2008
Oulu

Tiivistelmä

Yhteiskuntahyönteisten (muurahaiset, ampiaiset, mehiläiset ja termiitit) ekologisen menestyksen kääntöpuolena on ollut jatkuva riesa taudinaiheuttajista, joita suurissa yhteisöissä tavataan runsaammin kuin yksittäin elävissä eliöissä. Taudinaiheuttajien tuoman paineen myötä yhteiskuntahyönteisille on kehittynyt käyttäytymiseen perustuvia puolustusmekanismeja täydentämään kaikille monisoluisille eliöille yhteistä synnynnäistä, fysiologista immuniteettia. Nämä puolustusmekanismit ovat todiste siitä, että taudeilla on ollut suuri merkitys yhteiskuntahyönteisten käyttäytymisen evoluutiossa. Toisaalta taudinaiheuttajien vaikutuksista synnynnäiseen immuunipuolustukseen tiedetään hyvin vähän.

Väitöstutkimuksen ensisijainen kohde oli taudinaiheuttajien merkitys yhteiskuntahyönteisten synnynnäisen immuunipuolustuksen evoluutiossa. Tutkimuksessa tarkasteltiin, miten immuunijärjestelmän geenit ovat ajan mittaan muuttuneet. Tulokset osoittivat että muutoksia, jotka johtavat proteiinien aminohappojen vaihtumiseen on tapahtunut tiuhempaan tahtiin muurahaisilla ja mehiläisillä kuin yksittäin elävällä banaanikärpäsellä. Merkkejä erityisen voimakkaasta luonnonvalinnasta löydettiin kuitenkin yllättävän pienestä määrästä geenejä. Tämä voi johtua siitä, että käyttäytymiseen perustuvat puolustusmekanismit lieventävät taudinaiheuttajien vaikutusta synnynnäiseen immuunipuolustukseen.

Väitöstutkimukseen sisältyi myös hyönteisten solunsisäisen bakteerin, *Wolbachian*, siirtymismekanismien kartoitus kekomuurahaisilla. *Wolbachia* on loinen, joka siirtyy yleensä äidiltä jälkeläisille munasolussa. Leviäminen voi tapahtua myös horisontaalisesti lajitoverien ja jopa eri lajien edustajien kesken. Geenisekvensseihin perustuvassa tutkimuksessa kaikista muurahaisista löytyi *Wolbachia*-bakteereja, ja samasta yksilöstä saattoi löytyä useaa eri bakteerikantaa. Koska muurahaislajien väliset geneettiset erot olivat paljon suurempia kuin erot niissä elävien bakteerien välillä, voitiin päätellä että bakteerien pääasiallinen leviämistapa tutkituilla muurahaisilla on ollut horisontaalinen.

Asiasanat: hyönteisten immuunipuolustus, mehiläinen, molekyylievoluutio, muurahainen, positiivinen valinta, Wolbachia

Acknowledgements

I owe my deepest gratitude to my supervisor Professor Pekka Pamilo. I appreciate the independence that you have given me in the projects, and the support you have provided whenever I have needed it. It has been a privilege to work with you; it is not easy to find as benevolent a scientist as you are.

Another important person for the completion of this thesis has been our lab technician Riitta Jokela. Thank you for the magnificent work you have done! Many thanks also to the past and current members of the ant group: Marianne Elias, Kukka Haapaniemi, David Hughes, Hannaleena Mäki-Petäys, Jonna Saapunki and Anu Vihavainen, it has been a joy to work with you! Marjut Kreivi and Niina Tero who shared the lab with us were also good company during the years. Thanks also to Christian Bernasconi for cheering us up every now and then by visiting our lab and bringing us delicious Swiss chocolate!

I wish to thank all the people in the Department of Biology in Oulu University for creating such a nice working environment. The people in the plant genetics group: Sonja Kujala, Johanna Leppälä and Tanja Pyhäjärvi, thank you also for your friendship! I have greatly enjoyed our lunch and coffee break discussions, whether they concerned science or something else completely. During the last year the Monday PhD coffee club has provided valuable peer support.

I am grateful to Professor Ross Crozier and Professor Anneli Hoikkala for reviewing this thesis. I would also like to thank my co-authors Jay Evans, Martin Hasselmann, Max Reuter and Olav Rueppel for sharing ideas and for support and encouragement.

My most loving thanks belong to my family and relatives. Sirpa ja Ismo Brunni, olette mahtavat isovanhemmat Eemulle, kiitos korvaamattomasta avustanne vuosien varrella. Äiti, kiitos vankkumattomasta tuesta ja kannustuksesta, olet tärkeä. Matti, ilman sinua tämä väitöskirja ei olisi valmistunut aikataulussa, kiitos siitä ja kaikesta muusta. Eemu-kulta, mikään muu maailmassa ei tuota minulle niin paljon iloa kuin sinä, sen avulla olen jaksanut paremmin.

Oulu, September 2008 Lumi Viljakainen

Abbreviations

GNBP	Gram-negative bacteria-binding protein
MYA	million years ago
PCR	polymerase chain reaction
PGRP	peptidoglycan recognition protein
RACE	rapid amplification of cDNA ends
π	nucleotide diversity

List of original papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals:

I Viljakainen L & Pamilo P (2005) Identification and molecular characterization of *defensin* gene from the ant *Formica aquilonia*. Insect Mol Biol 14: 335–338.

II Viljakainen L & Pamilo P (2008) Selection on an antimicrobial peptide defensin in ants. J Mol Evol. In press.

III Viljakainen L, Evans JD, Hasselmann M, Rueppell O, Tingek S & Pamilo P (2008) Rapid evolution in immune proteins of social insects. Manuscript.

IV Viljakainen L, Reuter M & Pamilo P (2008) *Wolbachia* transmission dynamics in *Formica* wood ants. BMC Evol Biol 8: 55.

Contents

1 Introduction

1.1 Evolutionary genetics

Evolution is a process in which organisms adapt to spatially and temporally varying environments. The raw material for adaptations is a change in DNA sequence, a mutation. In a population mutations create genetic variation, which in turn is shaped by genetic drift and selection. Drift and selection determine the fate of the new mutation driving it to either extinction or fixation. In the long-term genetic differences between populations may accumulate sufficiently to give rise to a new species. Evolutionary genetics is a broad field of study that aims to clarify genetic changes, natural forces and population parameters that have resulted in differences between populations and among species.

1.2 Molecular evolution

The study of molecular evolution focuses on deciphering the process of evolution at the molecular level, namely DNA, RNA and protein. The rate of molecular evolution is a function of deleterious, neutral and advantageous mutations, their selection coefficients and the effective population size (Fay & Wu 2003). For four decades the salient theoretical basis of molecular evolutionary studies has been the neutral theory of molecular evolution introduced by Kimura (1968), see also King & Jukes (1969). The neutral theory of molecular evolution states that most nucleotide and amino acid substitutions are selectively neutral. The theory does not rule out adaptation, but emphasises the predominant role of random genetic drift driving the mutations to extinction or fixation. According to the neutral theory functionally important amino acids will be constrained over time, whereas neutral sites are allowed to accumulate nucleotide substitutions seen as a faster rate of evolution.

Ohta (1973) extended the neutral theory to include slightly deleterious mutations (and later on, also slightly advantageous mutations) and named it “the nearly neutral theory”. The major difference between these theories is the effect of population size. The substitution process of neutral mutations is not affected by population size, whereas nearly neutral mutations become fixed faster in smaller populations. This is because small populations are more affected by drift that can hinder selection for alleles that have small selection coefficients. Consequently,

the rate of molecular evolution is expected to negatively correlate with population size.

When inferring natural selection that has shaped the evolution of DNA sequences in the past, the neutral theory of molecular evolution provides the null hypothesis against which alternative hypothesis can be tested. Accordingly, nucleotide sites having a rate of evolution less than that expected for neutral sites have been under purifying selection and sites having a rate of evolution greater than that expected for neutral sites are evolving under positive selection. The best known examples of positive selection involve host-pathogen interactions driving positive selection in either the pathogen or in the host (Stahl & Bishop 2000, Worobey *et al.* 2007).

It was generally thought, under the paradigm of neutral theory, that positive selection is very rare in nature. However, the detection of positive selection has been restricted because the approaches used to infer selection from sequence data were based on the assumption that the nonsynonymous substitution rate is higher than the synonymous substitution rate throughout the entire protein. This is rarely the case as positive selection usually affects only a few sites in the protein whereas in the majority of sites purifying selection dominates (Nielsen 2005). The development of likelihood methods (Kosakovsky Pond *et al.* 2005, Yang 2007) based on codon substitution models (Goldman & Yang 1994, Muse & Gaut 1994), has facilitated the research focusing on finding signs of positive selection.

1.3 Insect immunity

The insect immune system has been studied in great detail in the fruit fly *Drosophila melanogaster*, a widely used model organism in genetics. Insect immunity involves only innate responses and lacks adaptive immunity, which has evolved in the ancestors of jawed vertebrates (Gillespie *et al.* 1997). Characteristics of adaptive immunity, such as somatic recombination of immune receptor genes and clonal expansion of activated lymphocytes, are thus absent in insects (Hoffmann 2003). Despite the fact that insect immunity lacks the extent of memory and specificity typical to adaptive immunity (though see Sadd & Schmid-Hempel 2006), insects are surprisingly resistant to microbial infections (Hoffmann & Reichhart 2002).

Insect innate immune responses are in essence similar among insect taxa (Gillespie *et al.* 1997). Upon infection innate defence mechanisms are activated comprising cellular and humoral responses. Phagocytosis by special blood cells,

melanisation and encapsulation of large invaders constitute the cellular response. Humoral responses involve chains of events that begin in the hemolymph, which is an insect correlate of blood and circulates in the open body cavity. These responses result in the production of antimicrobial peptides via two distinct signalling pathways, Toll and Imd, which have been reported to also act synergistically (reviewed by Ferrandon *et al.* 2007).

1.3.1 Toll pathway

Activation of the Toll pathway (Fig. 1) is triggered by the recognition of Gram-positive bacteria and fungi (Ferrandon *et al.* 2007). There are two types of recognition proteins, PGRPs (peptidoglycan-recognition proteins), and GNBPs (Gram-negative bacteria-binding proteins). It was previously thought that GNBPs bind lipopolysaccharide (LPS) on the surface of Gram-negative bacteria. However, the recognition of both PGRPs and GNBPs is based on different types of bacterial peptidoglycan, and despite the name, GNBPs can recognise Gram-positive bacteria (Leulier *et al.* 2003).

PGRPs form a large evolutionary conserved family. The *Drosophila* genome encodes several PGRPs that all share a PGRP domain of 160 amino acids (Werner *et al.* 2000). Some of the PGRPs have enzymatic activity while others have lost that function and operate solely in recognising microbes (Mellroth *et al.* 2003, Bischoff *et al.* 2006). The Toll pathway is activated by two recognition PGRPs, PGRP-SA and PGRP-SD, and GNBP1 that all bind to Gram-positive bacteria (Gobert *et al.* 2003, Wang *et al.* 2006). The Toll pathway can also be triggered by GNBP3 and Persephone, which detect fungal components (Gottar *et al.* 2006, Ligoxygakis *et al.* 2002).

The recognition phase is followed by a complex signalling cascade, which involves several interacting molecules (reviewed by Ferrandon *et al.* 2007). The extracellular ligand Späzle binds to the transmembrane receptor Toll resulting in conformational changes in the receptor. The binding triggers formation of a complex around intracellular region of Toll involving a death domain containing molecules MyD88 (myeloid differentiation primary-response gene 88), Tube and Pelle. This is followed by phosphorylation of Cactus and subsequently a transfer of the transcription factor DIF (Dorsal-related immunity factor) to the nucleus.

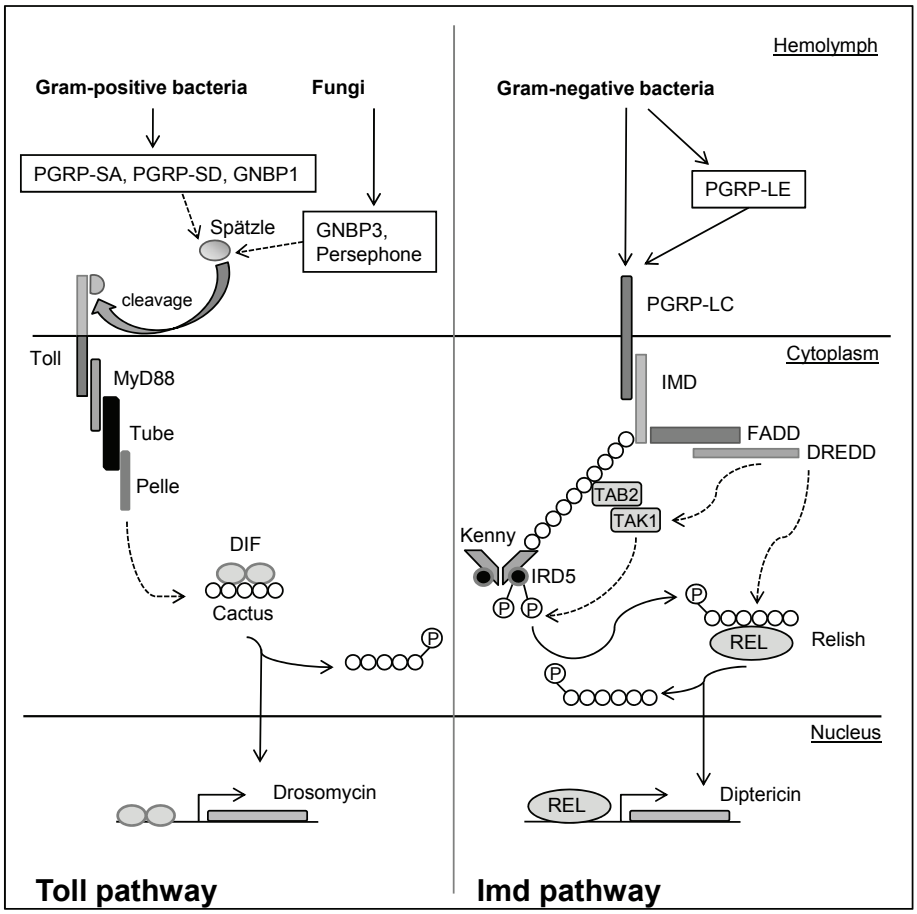

Fig. 1. Humoral response of insect immune system involves Toll and Imd pathways. This scheme is based on knowledge of *Drosophila*, and is modified from Hoffmann (2003) and Ferrandon *et al.* (2007).

The signalling cascade culminates in the expression of genes encoding antimicrobial peptides in the nucleus. Antimicrobial peptides are produced primarily in the fat body tissue, which is equivalent to the mammalian liver (Hoffmann 2003). Several peptides with antimicrobial properties have been characterised in *Drosophila*: defensin is anti-Gram-positive peptide; cecropins, drosocin, attacins and diptericins act against Gram-negative bacteria; drosomycin and metchnikowin have an antifungal function (Bulet *et al.* 1999). Of these

peptides, antifungal drosomycin and metchnikowin and anti-Gram-positive defensin are induced by the Toll pathway.

The molecules involved in the Toll signalling cascade presented above concern adult fruit flies, whereas the composition slightly differs in embryos. It is also notable that some of the factors along the Toll pathway do not function only in immune defence, but are also crucial in the development of the fly embryo (Moussian & Roth 2005).

1.3.2 Imd pathway

The Imd pathway (Fig. 1) is mainly involved in defence against Gram-negative bacteria. PGRP-LC and PGRP-LE bind these bacteria and activate the Imd signalling. PGRP-LC is a transmembrane receptor which interacts, probably via an unidentified adaptor (Kaneko *et al.* 2006), with IMD (immune deficiency) that is a death domain-containing molecule. IMD mediates signals to two distinct cascades (reviewed by Ferrandon *et al.* 2007). The first cascade involves TAK1 (transforming growth factor beta-activated kinase 1) and TAB2 (TAK1-binding protein 2), which together activate the IKK (IκB kinase) signalling complex. This complex is formed by the catalytic subunit IRD5 (immune-response deficient 5) and a regulatory subunit Kenny, and results in the phosphorylation of Relish. The phosphorylation triggers the second cascade, which involves IMD, FADD (FAS-associated death domain) and DREDD (death-related ced-3/Nedd2-like protein) resulting in the cleavage of Relish. The subunits of Relish are released to enter the nucleus, where they act as the key transcription factors of the Imd pathway (Hedengren *et al.* 1999) inducing the production of anti-Gram-negative peptides.

1.4 Social insects

All ants (Hymenoptera, Formicidae) and termites (Isoptera) and certain bees and wasps (Hymenoptera) are social insects. Division of labour, cooperative brood care and overlap in generations are characteristics of social insects. Division of labour is most striking in reproduction, where only one or a few queens lay eggs. In hymenopteran social insects all the other females are sterile workers which take care of diverse tasks from brood care and foraging to defence of the nest. Males contribute to colony life by mating with the queens, after which the males die. Hymenopteran social insects are haplodiploid; females are produced from

fertilised eggs and thus are diploid, whereas males are produced from unfertilised eggs and thus are haploid.

1.4.1 Social immunity

The characteristics of social insects have been suggested to represent a significant cost from the perspective of immunity (Schmid-Hempel 1998). Large colonies with warm temperatures and interacting individuals are propitious breeding grounds for pathogens. However, in addition to individual defences social insects have evolved a multitude of collective defences (reviewed by Boomsma *et al.* 2005 and Cremer *et al.* 2007). For example wood ants incorporate into their nest conifer resin which has been demonstrated to have antibacterial and antifungal properties (Christe *et al.* 2003, Chapuisat *et al.* 2007). Other preventive means include allogrooming (mutual cleaning) and collective recognition and removal of infected individuals.

If all these measures fail and an infection starts to spread in the colony there are still some mechanisms to restrain the contagion. Division of labour of the workers means that they form groups that are specialised in certain tasks. Working in specialised groups decreases interactions between individuals in a colony thereby reducing the spread of the infection. Hygienic behaviour is also important in repression of the disease encompassing removal of individuals that have died from an infection as well as removal of other waste. Finally, increasing genetic variation among colony members by multiple mating by the queen (polyandry) or forming colonies with multiple reproducing queens (polygyny) may hinder the spread of the infection in the colony, as some individuals may be resistant to the disease.

1.4.2 Pathogens infecting social insects

Pathogens and parasites that infect social insects have been listed by Schmid-Hempel (1998), and that database has been recently updated (Boomsma *et al.* 2005). The honeybee *Apis mellifera* is the best-studied social insect in this respect due to its economical importance as a pollinator and a producer of honey. In summary, social insects suffer from various micro- and macroparasites including bacteria, fungi, viruses, protozoa, nematodes, helminths, mites and parasitic insects. The most widely studied bacteria in insects are *Wolbachia*.

1.4.3 Wolbachia

Wolbachia are Gram-negative bacteria infecting arthropods and filarial nematodes. They are endosymbionts that are maternally inherited in the cytoplasm of the egg, but also horizontal transfer between individuals of the same species and even of different species have been reported (Vavre *et al.* 1999, Huigens *et al.* 2000, Huigens *et al.* 2004, Ahrens & Shoemaker 2005). *Wolbachia* are reproductive parasites promoting their own transmission by causing cytoplasmic incompatibility, parthenogenesis, male killing and feminisation, but may also be obligate mutualists (Werren 1997, Hoerauf *et al.* 1999, Dedeine *et al.* 2001). The specific effects in hymenopteran social insects are not known, although it has been suggested that *Wolbachia* most likely causes cytoplasmic incompatibility and male killing in ants (Keller *et al.* 2001, Van Borm *et al.* 2001, Wenseleers 2001, Wenseleers *et al.* 2002).

1.5 Aims of the study

An essential feature of social insects is the formation of colonies in which related individuals behave cooperatively. This, in addition to reproductive division of labour, has without doubt had a large influence on the evolutionary and ecological success of social insects. A major cost of social life is the high amount of pathogens in large groups and the more likely transmission of pathogens among closely related individuals. Social insects have evolved many colony-level defences, e.g. hygienic behaviour, and the existence of these traits manifests the importance of pathogens in shaping the evolution of social insects. However, very little is known about how pathogen pressure has affected the evolution of the innate immune system.

The principal aim of the study was to investigate the effect of the high pathogen pressure on the molecular evolution of genes involved in the innate immune system. I started by identifying and characterising the first immune gene, *defensin*, from ants (I), continued with a study of molecular evolution of *defensin* in several ant taxa (II) and extended the topic to cover several immune genes in ants and honeybees (III). Paper III also included the identification of four immune genes from *Myrmica* ants.

The second part of the thesis dealt with the endosymbiotic bacteria *Wolbachia* in ants. These bacteria are able to manipulate the behaviour of the host in order to promote their own transmission. *Wolbachia* are transmitted among the hosts either

vertically from mother to offspring or horizontally between individuals that can be from the same or even different species. I have examined the occurrence and transmission dynamics of *Wolbachia* in *Formica rufa* group wood ants from a large geographical scale covering most of the Eurasian distribution of these wood ants (IV).

2 Material and methods

2.1 Sampling

2.1.1 Ants

Papers II and III used phylogenetic information and the aim was to include many different species to improve the power of a codon-based selection test. However, the applicability of PCR (polymerase chain reaction) primers set the limit to the extent of sampling. In paper II PCR primers were designed based on the *defensin* sequence from *Formica aquilonia* belonging to the subfamily Formicinae. Therefore most samples in paper II represented this subfamily with an approximate divergence time of 75 MYA (million years ago) between the most distant genera *Lasius* and *Formica* (Brady *et al.* 2006, Moreau *et al.* 2006). In addition, two *Myrmica* species were included, which belong to the subfamily Myrmicinae. The *Myrmica defensin* was identified separately as described in paper II. The subfamilies Formicinae and Myrmicinae are separated by more than 100 million years (Brady *et al.* 2006, Moreau *et al.* 2006) and thus the sampling in paper II had considerable phylogenetic depth.

The ant samples in paper III involved several species of *Myrmica* and the outgroup *Manica rubida*, which is the closest relative to *Myrmica* (Astruc *et al.* 2004, Brady *et al.* 2006). Additionally, sequences of *Solenopsis invicta* were retrieved from an EST database (Fourmidable Ant Sequence Database) to add depth to the sampling. The approximated divergence time of *Myrmica* and *Solenopsis* is 90 MYA (Moreau *et al.* 2006).

The data in paper IV consisted of six species of the *Formica rufa* group wood ants that have diverged approximately 0.5 MYA (Goropashnaya *et al.* 2004).

2.1.2 Honeybees

Five species of honeybees were used in paper III, *Apis mellifera*, *A. cerana*, *A. koschevnikovi*, *A. dorsata* and *A. andreniformis*. The divergence times for these species are not resolved, but the origin of the genus *Apis* has been estimated to be in the early Oligocene 30 – 40 MYA (Engel 1998, Arias & Sheppard 2005).

2.2 Molecular methods

2.2.1 Immune gene identification in ants

The target ant immune genes were selected along the well-characterised Toll and Imd pathways of the *Drosophila* innate immune system and the identification process used degenerate primers in 3´ rapid amplification of cDNA ends (RACE). The degenerate primers were designed with the program CODEHOP (Rose *et al.* 1998). The primer design is based on conserved protein blocks in protein multiple alignments. The protein sequence of *Formica rufa* defensin (Taguchi *et al.* 1998) was used in paper I and protein sequences of several immune genes from *Apis mellifera* and *Drosophila melanogaster* along with protein sequences deduced from the *Solenopsis invicta* EST database (Fourmidable Ant Sequence Database) in paper III. In the reverse transcription stage of the primer design a codon usage table of *Lasius niger* was applied.

In paper I RNA was extracted from untreated individual ants, whereas in paper III the immune system of the ants was induced by pricking the abdomen with a needle and RNA was extracted from a pool of individuals. Successful 3´RACE products were cloned and sequenced. Using the sequences of 3´RACE products gene-specific primers were designed either manually (I) or with the program Primer3Plus (Untergasser *et al.* 2007) (III) for subsequent use in 5´RACE. PCR and sequencing primers were designed based on the aligned 3´ and 5´RACE products. The immune genes were amplified and sequenced from genomic DNA in order to use the information of the intron sequences in the phylogeny reconstruction.

2.2.2 Sequencing and cloning

The data consists of DNA sequences. All the PCR primers used for DNA amplification, as well as the sequencing primers, have been designed for these studies, except in paper IV, in which general primers for amplifying a gene encoding for *Wolbachia* surface protein (*wsp*) were used (Braig *et al.* 1998). The PCR products of the immune genes were sequenced directly or first cloned and then sequenced, and the sequencing was done in both 3´and 5´ directions. At least three bacterial colonies per clone were sequenced, the quality of the sequences was checked manually after which the sequence strands were assembled into

contigs. PCR products of the *Wolbachia wsp* were all cloned, after which multiple colonies were sequenced and processed as above.

2.3 Statistical analyses

2.3.1 Sequence polymorphism and divergence

In order to investigate sequence polymorphism and divergence the gene sequences were aligned using the help of the deduced protein sequence. There are several methods to examine the extent of divergence between DNA sequences, which is the foundation for the study of molecular evolution. The simplest measure is the number of polymorphic sites, which is the number of nucleotide sites having two or more variants. In within species sequence data a useful measure is nucleotide diversity, π, which is the average number of nucleotide substitutions per site between two sequences.

In a species-level comparison the divergence between two sequences may be large and thus, if the DNA sequence encodes a protein, it is informative to separate the substitutions that result in a change in the encoded amino acid (nonsynonymous substitutions) from those that do not (synonymous substitutions). Then, the divergence between the two nucleotide sequences may be represented as the rate of synonymous substitutions per synonymous site (d_S) and the rate of nonsynonymous substitutions per nonsynonymous site (d_N).

A coding sequence contains more sites in which a nonsynonymous substitution may potentially occur than sites in which a synonymous substitution may occur. Also, the number of synonymous and nonsynonymous sites varies depending on the gene. Moreover, in a pairwise comparison of sequences originating from different species it is likely that some of the nucleotide sites have changed multiple times. Several methods have been developed to estimate the number of synonymous and nonsynonymous sites and also to avoid underestimating the divergence by correcting for multiple hits. In the estimation of d_N and d_S I have used the method of Pamilo & Bianchi (1993) and Li (1993).

2.3.2 Phylogeny reconstruction

The evolutionary relationships between species may be studied by reconstructing a phylogenetic tree using molecular data. There are several methods to choose

from to reconstruct a phylogeny including distance, parsimony and likelihood methods and most recently the Bayesian inference. An essential part of phylogenetic analysis in all the methods except parsimony is choosing the substitution model. The function of using a substitution model is to correct for multiple changes that have occurred at an individual nucleotide site.

In papers II and III phylogenetic trees were generated for the ants because no published phylogenies exist for the species involved. Phylogenetic information was needed in the codon-based methods to infer selection as well as in the relative rate test. I have chosen Bayesian inference as the method for phylogeny reconstruction implemented in the program MrBayes version 3.1 (Ronquist & Huelsenbeck 2003). In paper II sequence data from only one gene, *defensin*, were used and the obtained tree is rather a gene tree than a species tree. In paper III data from four genes were used which should improve the chance of reconstructing a species tree. The data sets in both papers I and II also involved the intron sequences of the genes. Substitution models for each gene separately were chosen with programs MrModeltest version 2.2 (Nylander 2004) and PAUP* version 4.0b10 (Swofford 2002).

2.3.3 Detection of positive selection

Positive selection in papers II and III was tested for by using various codon-substitution models based on the original model by Goldman & Yang (1994). These models are implemented in a program codeml in the PAML 4 package, and they use phylogenetic information (the topology of the tree) in a comparison of protein-coding genes (Yang 2007). The models are based on the neutral theory of molecular evolution. Therefore, a nonsynonymous substitution rate (d_N) greater than a synonymous substitution rate (d_S) is a sign of positive selection. The models consider a codon, rather than the whole gene, as a unit of evolution. I used site models that allow ω ($\omega = d_N/d_S$ = nonsynonymous / synonymous rate ratio) to vary among codons (Nielsen & Yang 1998). Detection of codons with $\omega > 1$ is considered as an indication of positive selection. Basically, a model that does not allow $\omega > 1$ is compared with a model that does, and a likelihood ratio test is applied to determine the model that best fits the data.

2.3.4 Relative rate test

To infer differences in evolutionary rates (both in d_N and d_S) a relative rate test implemented in the program RRTree (Robinson-Rechavi & Huchon 2000) was applied in papers II and III. The relative rate test is based on a comparison of the evolutionary distance between one target of interest and an outgroup with the distance between another target of interest and the outgroup (Sarich & Wilson 1973, Wu & Li 1985). RRTree uses phylogenetic information, namely the topology of the tree, so that evolutionary distances between two groups of sequences forming separate phylogenetic lineages can be compared.

2.3.5 Correction for multiple testing

To avoid false positives in the relative rate test and the codon-based methods of detecting positive selection, a correction for multiple testing was performed. I have used a false discovery rate method implemented in the program QVALUE (Storey & Tibshirani 2003). For each *p*-value the program estimates a *q*-value, which is a measure of significance in terms of the false discovery rate. For example, if the *q*-value for a certain case is 0.05, it means that 5% is the expected proportion of false positives incurred if this certain case is considered significant.

3 Results and discussion

3.1 Isolation of immune genes from ants

Identification of an antimicrobial peptide gene *defensin* from the wood ant *Formica aquilonia* is described in paper I, and the identification of four immune genes (*PGRP-SC2*, *GNBP1*, *abaecin* and *defensin*) from the red ant *Myrmica ruginodis* in paper III. Originally in the paper III the aim was to isolate 14 ant immune genes, for which 2–10 degenerate primers per gene were designed. Of these 14 genes, five were successfully identified from *Myrmica* and five from *Formica* ants (Table 1). Due to problems in RACE the immune genes identified from *Formica* are all largely partial either in the 3' or 5' end and thus not used in this thesis. Also *hymenoptaecin* from *Myrmica* was excluded because of difficulties in the identification process (several distinct PCR products) and the time limit associated with the project.

Table 1. Target immune genes in ants including the number of primers designed for each gene and successful isolations.

Gene	Number of primers	Successful isolation	
		Myrmica	*Formica*
PGRP-SA	3[a]	x	x
GNBP1	5	x	
PGRP-LC	3		
Toll	8		x
MyD88	10		
Tube	6		
Pelle	7		x
cactus	6		
Dorsal	7		x
Ird5	2		
TAK-1	10		x
defensin	5	x	x[b]
abaecin	2	x	
hymenoptaecin	2	x	

[a] With these primers two genes were identified, *PGRP-SC2* from *Myrmica* and *PGRP-SA* from *Formica*

[b] Identification described in paper I

The genes identified represent the three stages of the immune pathways: recognition of pathogens, intracellular signalling and production of effector proteins. PGRP-SA, PGRP-SC2 and GNBP1 are recognition proteins; Toll, Pelle, Dorsal and TAK-1 are intracellular signalling proteins; defensin, abaecin and hymenoptaecin are antimicrobial peptides, or in other words, effector proteins (Fig. 1 and Table 1). The functions of these proteins have been resolved in previous studies. The ant protein sequences deduced shared high levels of amino acid identity with the known proteins (43–75% according to blastp searches), so they most likely have the same function, although experimental tests are needed to ascertain this.

3.2 Evolutionary rates in the immune genes of social insects

The evolutionary rates in the immune genes of both ants and honeybees were high, as indicated by the d_N/d_S ratios (II, III, Table 2). The ratios were estimated with two distinct methods: based on pairwise estimates using the average of nonsynonymous substitution rates divided by the average of synonymous substitution rates, and a "global" d_N/d_S ratio estimated by the program codeml in PAML (Table 2). The d_N/d_S ratio calculated from the pairwise distance estimates was 0.725 in *defensin*. The global d_N/d_S ratio of *defensin* varied from 0.83 to 1.43 depending on the gene region (Table 4 in II). The d_N/d_S ratio in the *Myrmica* immune genes was 0.519 calculated from the combined pairwise estimates of the four immune genes. The median of the global d_N/d_S ratios was 0.500. In honeybees the mean of the pairwise distances resulted in a slightly lower d_N/d_S ratio of 0.455, and a median of the global ratios of 0.257. As a comparison non-immunity gene sequences were studied in honeybees, and they showed a significantly lower median of d_N/d_S ratios of 0.010 ($P = 0.005$).

Table 2. Evolutionary rates in immune genes of ants and honeybees.

Locus	Codons	Taxa	Global d_N/d_S[a]	Average d_N/d_S[b]
Ant immune genes[c]				
PGRP-SC2	169	10	0.525	
GNBP1	240	10	0.500	0.519
abaecin	52	11	0.400	
defensin (III)	77	5	0.500	
defensin (II)	82	20	0.841	0.725
Honeybee immune genes[c]				
PGRP-SA	145	5	0.203	
PGRP-SC2	162	5	0.239	
GNBP1 (LG9)	389	5	0.257	
GNBP1 (LG4)	386	5	0.200	
cactus-1	287	5	0.294	0.455
Relish	347	5	0.194	
Imd	509	5	0.318	
hymenoptaecin	155	5	0.299	
defensin-1	110	4	0.300	
Honeybee non-immunity genes[c]				
EF1-alpha	157	5	0	
Itpr[d]	321	5	0.022	0.134
nad2[d]	126	5	0.009	
cox2[d]	225	5	0.010	

[a] estimated using codeml in PAML

[b] derived from average pairwise d_N and d_S estimates

[c] from paper III

[d] sequences are from the GenBank

It has been suggested that social insects could have more rapid evolutionary rates because of small effective population sizes (resulting from a low ratio of the number of reproductive individuals to the total biomass), and therefore the more rapid fixation of nearly neutral mutations by drift enhancing the nonsynonymous substitution rate (Bromham & Leys 2005). This effect would be genome-wide. However, comparison of synonymous and nonsynonymous substitution rates in the immune genes with the corresponding rates in non-immunity genes in honeybees showed that the average rate of synonymous substitutions was about the same for the immune genes ($d_S = 0.101$) and the nuclear non-immunity genes ($d_{S\ \text{nuclear}} = 0.097$) (III). Therefore, in honeybees the difference in d_N/d_S ratios between the immune genes and the non-immunity genes is caused by a higher rate

of nonsynonymous substitutions in the immune genes. It is likely that the same holds true for ants, but a similar comparison of immune genes with non-immunity genes is required to confirm this.

In ants the median of the d_N/d_S ratios was almost twice that of the value in honeybees, the difference being significant (Mann-Whitney *U*-test, $P = 0.005$). It would be tempting to explain the higher evolutionary rate in ants by the difference in their life history compared to honeybees. For example, Boomsma *et al.* (2005) propose that bees could have a lower risk of contracting diseases as they forage on relatively hygienic nectar and pollen, whereas ants are omnivorous. However, based of the available data it is premature to draw clear conclusions.

The median of the global d_N/d_S ratios (Sackton *et al.* 2007) of the same *Drosophila* immune genes as involved in paper III was 0.110. This is significantly lower than the rate of immune gene evolution in both ants and honeybees (Mann-Whitney *U*-test, $P = 0.001$ for honeybee-*Drosophila* and $P = 0.016$ for ant-*Drosophila* comparison) (III). Taken together, it is clear that pathogen pressure causes the occurrence of many more amino acid substitutions in the immune genes of honeybees compared to non-immunity genes. Also, evidently the pathogen pressure is higher in social insects compared to *Drosophila*, as indicated by the significantly higher d_N/d_S ratios in social insects.

3.3 Positive selection in the immune genes of social insects

Positive selection was detected in only two out of the five immune genes studied in ants, whereas none of the nine immune genes studied in honeybees showed signs of positive selection (II, III). The positively selected ant immune genes were an antimicrobial peptide *defensin* and *PGRP-SC2*, which in *Drosophila* functions as a regulator of the Imd pathway. Positive selection in *defensin* was studied in both papers II and III, the former covering 29 ant species representing subfamilies Formicinae and Myrmicinae and the latter covering several *Myrmica* species and *Solenopsis invicta* from the subfamily Myrmicinae. Evidence of positive selection was found only in paper II.

There are several possible explanations for finding positive selection in *defensin* in paper II, but not in paper III. First, in paper II a larger number of species was studied, and thus positive selection may have remained undetected in paper III due to a lack of power in the selection test, not only in the case of *defensin*, but also in the other immune genes. Second, in paper II few lineages were shown to have a significantly higher rate of nonsynonymous substitutions,

and the overall result likely depended on these lineages. The reason for finding selection in *defensin* in paper II, but not in paper III, may thus be that in the data of paper II a sufficient number of lineages (but not necessarily all) has been under positive selection. For example in *Drosophila* positive selection has been shown in the intracellular signalling molecule *Relish* in the *D. melanogaster* lineage, though not in the other species of the *D. melanogaster* group (Sackton *et al.* 2007). This kind of branch-specific test requires a prior assumption of finding positive selection in certain lineages, but there was no motivation for such expectations in either paper II or III.

It is important to notice that only a fraction of all the insect immune genes has been studied in this thesis. In *Drosophila* 245 immune-related genes have been identified (Sackton *et al.* 2007). Positive selection was studied in 226 of these genes in six species of the *D. melanogaster* group, and evidence for selection was found only in 10% of the genes. Notably, none of the genes involved here belong to this group. However, in termites positive selection has been found in a defensin-like antimicrobial peptide termicin (Bulmer & Crozier 2004). In addition, two GNPBs and Relish have also evolved under positive selection in termites indicating that these proteins likely have been involved in host-pathogen arms races (Bulmer & Crozier 2006). Thus, the lack of positive selection in the majority of the immune genes studied here cannot be explained by these genes commonly not being involved in interactions promoting selection.

Positive selection in *defensin* and *PGRP-SC2* was not particularly strong. Only two codon sites out of the studied 82 in *defensin* were indicated as positively selected (II). Moreover, in *PGRP-SC2* even a smaller fraction of codon sites, only one out of 169, was consistently inferred to have been affected by selection. It is puzzling that on the one hand the immune genes in hymenopteran social insects are evolving at a fast rate, but on the other hand the immune genes do not show signs of strong positive selection. Based on the evolutionary rates in the immune genes the pathogen pressure in social insects seems to be higher when compared with *Drosophila*, but apparently not strong enough to cause many immune genes to evolve under positive selection. Therefore, there has to be some measures that lower the pathogen pressure.

3.4 Role of colony-level defences in the evolution of immune genes

Social insects have evolved many colony-level defences which are likely to keep pathogen pressure at a moderate level. Cremer *et al.* (2007) have classified these

defences into three stages: 1) protection against pathogen acquisition from the environment and import into the colony (e.g. honeybees eliminate infected individuals at the nest entrance), 2) preventing the establishment of a pathogen in the nest (nest hygiene), and 3) inhibition of pathogen spread (e.g. spatial and behavioural compartmentalisation).

Also important in the inhibition of pathogen spread is the level of genetic variation within colonies. Colony-level resistance has been shown to increase by genetic diversity resulting from polyandry (Hughes & Boomsma 2004) and mixing of genetic lines (Reber *et al.* 2008) in ants, and from outbreeding in termites (Calleri *et al.* 2006). Also, strong genotypic host-parasite interactions have been demonstrated in bumble bees (Schmid-Hempel & Reber Funk 2004). Even though social defences may lower pathogen pressure, evidence of genetic variation in colony survival and strain-specific effects suggest selection on physiological mechanisms. The present results show similar d_N/d_S ratios in all the genes from different parts of the immune defence pathways, and positive selection in immune genes has been detected in both ants and termites (II, Bulmer & Crozier 2004). Based on these results, it seems that selection on the amino acid composition of the immune-related genes has been an important part in the fight against pathogens by social insects. Therefore, it seems that social defence mechanisms have their limitations.

3.5 Wide-spread *Wolbachia* infection in wood ants

An example where both colony-level and individual defences may have failed was found in paper IV. *Wolbachia* bacteria were screened from multiple populations of six species of the *Formica rufa* group wood ants covering most of their geographical distribution area in Eurasia. All 32 samples examined were found to contain *Wolbachia*. These bacteria are reproductive parasites with their main route of transmission occurring vertically from the mother to the offspring in the cytoplasm of the egg (Werren 1997). Therefore, when transmitted vertically there may not be a chance for the host to get rid of an established *Wolbachia* infection, at least not by the innate immune pathways, in which microbial recognition takes place in the extracellular hemolymph.

However, in paper IV it was shown that, in addition to vertical transmission, frequent horizontal transmission has also occurred in the ants studied. Identical *Wolbachia* strains were found associated with distant mitochondrial DNA haplotypes among *F. rufa* group ants as well as with *F. rufa* group ants and a more

distantly related *F. exsecta* (Reuter & Keller 2003), with an estimated divergence time of 3.5–4 MYA (Goropashnaya 2003). During the process of horizontal transmission *Wolbachia* have to survive in an extracellular environment (Rasgon *et al.* 2006), and thus in this phase the recognition of *Wolbachia* by the host could be possible. Therefore, the wide-spread *Wolbachia* infection in wood ants indicates that the bacteria have successfully evaded the host immune system during several horizontal transmission events.

Siozios *et al.* (2008) have reviewed multiple studies in which it was shown that *Wolbachia* are able to escape the host immune system without actively suppressing the host immune response. On the other hand mutualistic relationships between *Wolbachia* and the host have been reported in filarial nematodes (Taylor *et al.* 2005) and in a parasitic wasp (Dedeine *et al.* 2001). The actual effect of *Wolbachia* on the ant hosts is not known, although cytoplasmic incompatibility and male-killing have been suggested (Keller *et al.* 2001, Van Borm *et al.* 2001, Wenseleers 2001, Wenseleers *et al.* 2002). Whether the extensive occurrence of *Wolbachia* in wood ants could be explained by bacterial evasion of the host immune response or whether there could even be positive interactions between the ant host and *Wolbachia* remains to be resolved in future studies.

4 Concluding remarks

In conclusion I have identified five immune genes in ants and shown that all these genes, along with the nine studied honeybee immune genes, have evolved rapidly. This is the first demonstration in hymenopteran social insects of high pathogen pressure having an impact on the evolution of immune system genes. Despite the high evolutionary rates, a clear indication of positive selection was found in only two ant immune genes. It may be that the behaviourally based colony-level defences in social insects reduce the selection pressure on immune genes.

Both the individual- and colony-level defences have failed to prevent infection with the endosymbiotic bacteria *Wolbachia* – not only in ants, which were studied here, but also in many other arthropods and filarial nematodes studied earlier. *Wolbachia* double infections were found in all of the six wood ant species studied covering most of their geographical distribution area in Eurasia. It was shown that *Wolbachia* infection within this host clade has been shaped by both horizontal and vertical transmission of the bacteria.

Further comparative studies on immune gene evolution in social insects involving larger number of immune genes from species that represent different levels of sociality and different phylogenetic lineages are needed to examine whether there is an association between the level of sociality and the evolutionary rate in immune genes. It would also be interesting to determine the actual effect of *Wolbachia* in ants to shed light on the wide-spread *Wolbachia* infection in wood ants.

References

Ahrens ME & Shoemaker D (2005) Evolutionary history of *Wolbachia* infections in the fire ant *Solenopsis invicta*. BMC Evol Biol 5(1): 35.

Arias MC & Sheppard WS (2005) Phylogenetic relationships of honey bees (Hymenoptera:Apinae:Apini) inferred from nuclear and mitochondrial DNA sequence data. Mol Phylogenet Evol 37(1): 25–35.

Astruc C, Julien JF, Errard C & Lenoir A (2004) Phylogeny of ants (Formicidae) based on morphology and DNA sequence data. Mol Phylogenet Evol 31(3): 880–893.

Bischoff V, Vignal C, Duvic B, Boneca IG, Hoffmann JA & Royet J (2006) Downregulation of the *Drosophila* immune response by peptidoglycan-recognition proteins SC1 and SC2. PLoS Pathog 2(2): e14.

Boomsma JJ, Schmid-Hempel P & Hughes WOH (2005) Life histories and parasite pressure across the major groups of social insects. In: Fellowes MDE, Holloway GJ & Rolff J (eds) Insect Evolutionary Ecology. Wallingford: CABI: 139–175.

Brady SG, Schultz TR, Fisher BL & Ward PS (2006) Evaluating alternative hypotheses for the early evolution and diversification of ants. Proc Natl Acad Sci USA 103(48): 18172–18177.

Braig HR, Zhou W, Dobson SL & O'Neill SL (1998) Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. J Bacteriol 180(9): 2373–2378.

Bromham L & Leys R (2005) Sociality and the rate of molecular evolution. Mol Biol Evol 22(6): 1393–1402.

Bulet P, Hetru C, Dimarcq JL & Hoffmann D (1999) Antimicrobial peptides in insects; structure and function. Dev Comp Immunol 23(4–55): 329–344.

Bulmer MS & Crozier RH (2006) Variation in positive selection in termite GNBPs and Relish. Mol Biol Evol 23(2): 317–326.

Bulmer MS & Crozier RH (2004) Duplication and diversifying selection among termite antifungal peptides. Mol Biol Evol 21(12): 2256–2264.

Calleri DV, McGrail Reid E, Rosengaus RB, Vargo EL & Traniello JF (2006) Inbreeding and disease resistance in a social insect: effects of heterozygosity on immunocompetence in the termite *Zootermopsis angusticollis*. Proc Biol Sci 273(1601): 2633–2640.

Chapuisat M, Oppliger A, Magliano P & Christe P (2007) Wood ants use resin to protect themselves against pathogens. Proc Biol Sci 274(1621): 2013–2017.

Christe P, Oppliger A, Bancalà F, Castella G & Chapuisat M (2003) Evidence for collective medication in ants. Ecol Lett 6(168): 19.

Cremer S, Armitage SA & Schmid-Hempel P (2007) Social immunity. Curr Biol 17(16): R693–702.

Dedeine F, Vavre F, Fleury F, Loppin B, Hochberg ME & Bouletreau M (2001) Removing symbiotic *Wolbachia* bacteria specifically inhibits oogenesis in a parasitic wasp. Proc Natl Acad Sci USA 98(11): 6247–6252.

Engel MS (1998) Fossil honey bees and evolution in the genus *Apis* (Hymenoptera: Apidae). Apidologie 29: 265–281.

Fay JC & Wu CI (2003) Sequence divergence, functional constraint, and selection in protein evolution. Annu.Rev.Genomics Hum Genet 4: 213–235.

Ferrandon D, Imler JL, Hetru C & Hoffmann JA (2007) The *Drosophila* systemic immune response: sensing and signalling during bacterial and fungal infections. Nat Rev Immunol 7(11): 862–874.

Fourmidable Ant Sequence Database. http://fourmidable.unil.ch/.

Gillespie JP, Kanost MR & Trenczek T (1997) Biological mediators of insect immunity. Annu Rev Entomol 4269: 611–643.

Gobert V, Gottar M, Matskevich AA, Rutschmann S, Royet J, Belvin M, Hoffmann JA & Ferrandon D (2003) Dual activation of the *Drosophila* Toll pathway by two pattern recognition receptors. Science 302(5653): 2126–2130.

Goldman N & Yang Z (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol 11(5): 725–736.

Goropashnaya A (2003) Phylogeographic structure and genetic variation in *Formica* ants. Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 912.

Goropashnaya AV, Fedorov VB & Pamilo P (2004) Recent speciation in the *Formica rufa* group ants (Hymenoptera, Formicidae): inference from mitochondrial DNA phylogeny. Mol Phylogenet Evol 32(1): 198–206.

Gottar M, Gobert V, Matskevich AA, Reichhart JM, Wang C, Butt TM, Belvin M, Hoffmann JA & Ferrandon D (2006) Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. Cell 127(7): 1425–1437.

Hedengren M, Asling B, Dushay MS, Ando I, Ekengren S, Wihlborg M & Hultmark D (1999) Relish, a central factor in the control of humoral but not cellular immunity in *Drosophila*. Mol Cell 4(5): 827–837.

Hoerauf A, Nissen-Pahle K, Schmetz C, Henkle-Duhrsen K, Blaxter ML, Buttner DW, Gallin MY, Al-Qaoud KM, Lucius R & Fleischer B (1999) Tetracycline therapy targets intracellular bacteria in the filarial nematode *Litomosoides sigmodontis* and results in filarial infertility. J Clin Invest 103(1): 11–18.

Hoffmann JA (2003) The immune response of *Drosophila*. Nature 426(6962): 33–38.

Hoffmann JA & Reichhart JM (2002) *Drosophila* innate immunity: an evolutionary perspective. Nat Immunol 3(2): 121–126.

Hughes WO & Boomsma JJ (2004) Genetic diversity and disease resistance in leaf-cutting ant societies. Evolution 58(6): 1251–1260.

Huigens ME, de Almeida RP, Boons PAH, Luck RF & Stouthamer R (2004) Natural interspecific and intraspecific horizontal transfer of parthenogenesis-inducing *Wolbachia* in *Trichogramma* wasps. Proc Biol Sci 271(1538): 509.

Huigens ME, Luck RF, Klaassen RH, Maas MF, Timmermans MJ & Stouthamer R (2000) Infectious parthenogenesis. Nature 405(6783): 178–179.

Kaneko T, Yano T, Aggarwal K, Lim JH, Ueda K, Oshima Y, Peach C, Erturk-Hasdemir D, Goldman WE, Oh BH, Kurata S & Silverman N (2006) PGRP-LC and PGRP-LE have essential yet distinct functions in the *Drosophila* immune response to monomeric DAP-type peptidoglycan. Nat Immunol 7(7): 715–723.

Keller L, Liautard C, Reuter M & Brown WD (2001) Sex ratio and *Wolbachia* infection in the ant *Formica exsecta*. Heredity 87(2): 227–233.

Kimura M (1968) Evolutionary rate at the molecular level. Nature 217(5129): 624–626.

King JL & Jukes TH (1969) Non-Darwinian evolution. Science 164(881): 788–798.

Kosakovsky Pond SL, Frost SD & Muse SV (2005) HyPhy: hypothesis testing using phylogenies. Bioinformatics 21(5): 676–679.

Leulier F, Parquet C, Pili-Floury S, Ryu JH, Caroff M, Lee WJ, Mengin-Lecreulx D & Lemaitre B (2003) The *Drosophila* immune system detects bacteria through specific peptidoglycan recognition. Nat Immunol 4(5): 478–484.

Li WH (1993) Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J Mol Evol 36(1): 96–99.

Ligoxygakis P, Pelte N, Hoffmann JA & Reichhart JM (2002) Activation of *Drosophila* Toll during fungal infection by a blood serine protease. Science 297(5578): 114–116.

Mellroth P, Karlsson J & Steiner H (2003) A scavenger function for a *Drosophila* peptidoglycan recognition protein. J Biol Chem 278(9): 7059–7064.

Moreau CS, Bell CD, Vila R, Archibald SB & Pierce NE (2006) Phylogeny of the ants: diversification in the age of angiosperms. Science 312(5770): 101–104.

Moussian B & Roth S (2005) Dorsoventral axis formation in the *Drosophila* embryo–shaping and transducing a morphogen gradient. Curr Biol 15(21): R887–99.

Muse SV & Gaut BS (1994) A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Mol Biol Evol 11(5): 715–724.

Nielsen R (2005) Molecular signatures of natural selection. Annu Rev Genet 39: 197–218.

Nielsen R & Yang Z (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148(3): 929–936.

Nylander JAA (2004) MrModeltest 2.2.

Ohta T (1973) Slightly deleterious mutant substitutions in evolution. Nature 246(5428): 96–98.

Pamilo P & Bianchi NO (1993) Evolution of the *Zfx* and *Zfy* genes: rates and interdependence between the genes. Mol Biol Evol 10(2): 271–281.

Rasgon JL, Gamston CE & Ren X (2006) Survival of *Wolbachia pipientis* in cell-free medium. Appl Environ Microbiol 72(11): 6934–6937.

Reber A, Castella G, Christe P & Chapuisat M (2008) Experimentally increased group diversity improves disease resistance in an ant species. Ecol Lett 11(7): 682–689.

Reuter M & Keller L (2003) High levels of multiple *Wolbachia* infection and recombination in the ant *Formica exsecta*. Mol Biol Evol 20(5): 748–753.

Robinson-Rechavi M & Huchon D (2000) RRTree: relative-rate tests between groups of sequences on a phylogenetic tree. Bioinformatics 16(3): 296–297.

Ronquist F & Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19(12): 1572–1574.

Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM & Henikoff S (1998) Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. Nucleic Acids Res 26(7): 1628–1635.

Sackton TB, Lazzaro BP, Schlenke TA, Evans JD, Hultmark D & Clark AG (2007) Dynamic evolution of the innate immune system in *Drosophila*. Nat Genet 39(12): 1461–1468.

Sadd BM & Schmid-Hempel P (2006) Insect immunity shows specificity in protection upon secondary pathogen exposure. Curr Biol 16(12): 1206–1210.

Sarich VM & Wilson AC (1973) Generation time and genomic evolution in primates. Science 179(4078): 1144–1147.

Schmid-Hempel P (1998) Parasites in social insects. Princeton NJ USA, Princeton University Press.

Schmid-Hempel P & Reber Funk C (2004) The distribution of genotypes of the trypanosome parasite, *Crithidia bombi*, in populations of its host, *Bombus terrestris*. Parasitology 129(Pt 2): 147–158.

Siozios S, Sapountzis P, Ioannidis P & Bourtzis K (2008) *Wolbachia* symbiosis and insect immune response. Insect Science 15(1): 89–100.

Stahl EA & Bishop JG (2000) Plant–pathogen arms races at the molecular level. Curr Opin Plant Biol 3(4): 299–304.

Storey JD & Tibshirani R (2003) Statistical significance for genomewide studies. Proc Natl Acad Sci USA 100(16): 9440–9445.

Swofford DL (2002) PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sunderland, Massachusetts USA, Sinauer Associates.

Taguchi S, Bulet P & Hoffmann JA (1998) A novel insect defensin from the ant *Formica rufa*. Biochimie 80(431): 343–346.

Taylor MJ, Bandi C & Hoerauf A (2005) *Wolbachia* bacterial endosymbionts of filarial nematodes. Adv Parasitol 60: 245–284.

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R & Leunissen JA (2007) Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 35 (Web Server issue): W71–4.

Van Borm S, Wenseleers T, Billen J & Boomsma JJ (2001) *Wolbachia* in leafcutter ants: a widespread symbiont that may induce male killing or incompatible matings. J Evol Biol 14(5): 805–814.

Vavre F, Fleury F, Lepetit D, Fouillet P & Bouletreau M (1999) Phylogenetic evidence for horizontal transmission of *Wolbachia* in host-parasitoid associations. Mol Biol Evol 16(12): 1711–1723.

Wang L, Weber AN, Atilano ML, Filipe SR, Gay NJ & Ligoxygakis P (2006) Sensing of Gram-positive bacteria in *Drosophila*: GNBP1 is needed to process and present peptidoglycan to PGRP-SA. EMBO J 25(20): 5005–5014.

Wenseleers T (2001) Conflict from cell to colony. Unpublished PhD thesis. University of Leuven.

Wenseleers T, Sundstrom L & Billen J (2002) Deleterious *Wolbachia* in the ant *Formica truncorum*. Proc Biol Sci 269(1491): 623–629.

Werner T, Liu G, Kang D, Ekengren S, Steiner H & Hultmark D (2000) A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. Proc Natl Acad Sci USA 97(25): 13772–13777.

Werren JH (1997) Biology of *Wolbachia*. Annu Rev Entomol 42(1): 587–609.

Worobey M, Bjork A & Wertheim JO (2007) Point, counterpoint: The evolution of pathogenic viruses and their human hosts. Annu Rev Ecol Evol Syst 38: 515–540.

Wu CI & Li WH (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci USA 82(6): 1741–1745.

Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24(8): 1586–1591.

Original papers

I Viljakainen L & Pamilo P (2005) Identification and molecular characterization of *defensin* gene from the ant *Formica aquilonia*. Insect Mol Biol 14: 335–338.

II Viljakainen L & Pamilo P (2008) Selection on an antimicrobial peptide defensin in ants. J Mol Evol. In press.

III Viljakainen L, Evans JD, Hasselmann M, Rueppell O, Tingek S & Pamilo P (2008) Rapid evolution in immune proteins of social insects. Manuscript.

IV Viljakainen L, Reuter M & Pamilo P (2008) *Wolbachia* transmission dynamics in *Formica* wood ants. BMC Evol Biol 8: 55.

Original publications are not included in the electronic version of the dissertation.

497. Alaviuhkola, Terhi (2007) Aromatic borate anions and thiophene derivatives for sensor applications

498. Törn, Anne (2007) Sustainability of nature-based tourism

499. Autio, Kaija (2007) Characterization of 3-hydroxyacyl-ACP dehydratase of mitochondrial fatty acid synthesis in yeast, humans and trypanosomes

500. Raunio, Janne (2008) The use of Chironomid Pupal Exuvial Technique (CPET) in freshwater biomonitoring: applications for boreal rivers and lakes

501. Paasivaara, Antti (2008) Space use, habitat selection and reproductive output of breeding common goldeneye (*Bucephala clangula*)

502. Asikkala, Janne (2008) Application of ionic liquids and microwave activation in selected organic reactions

503. Rinta-aho, Marko (2008) On monomial exponential sums in certain index 2 cases and their connections to coding theory

504. Sallinen, Pirkko (2008) Myocardial infarction. Aspects relating to endogenous and exogenous melatonin and cardiac contractility

505. Xin, Weidong (2008) Continuum electrostatics of biomolecular systems

506. Pyhäjärvi, Tanja (2008) Roles of demography and natural selection in molecular evolution of trees, focus on *Pinus sylvestris*

507. Kinnunen, Aulis (2008) A Palaeoproterozoic high-sulphidation epithermal gold deposit at Orivesi, southern Finland

508. Alahuhta, Markus (2008) Protein crystallography of triosephosphate isomerases: functional and protein engineering studies

509. Reif, Vitali (2008) Birds of prey and grouse in Finland. Do avian predators limit or regulate their prey numbers?

510. Saravesi, Karita (2008) Mycorrhizal responses to defoliation of woody hosts

511. Risteli, Maija (2008) Substrate specificity of lysyl hydroxylase isoforms and multifunctionality of lysyl hydroxylase 3

512. Korsu, Kai (2008) Ecology and impacts of nonnative salmonids with special reference to brook trout (*Salvelinus fontinalis* Mitchill) in North Europe

513. Laitinen, Jarmo (2008) Vegetational and landscape level responses to water level fluctuations in Finnish, mid-boreal aapa mire – aro wetland environments

Book orders:
OULU UNIVERSITY PRESS
P.O. Box 8200, FI-90014
University of Oulu, Finland

Distributed by
OULU UNIVERSITY LIBRARY
P.O. Box 7500, FI-90014
University of Oulu, Finland

UNIVERSITY OF OULU P.O.B. 7500 FI-90014 UNIVERSITY OF OULU FINLAND

ACTA UNIVERSITATIS OULUENSIS

ISBN 978-951-42-8927-9 (Paperback)
ISBN 978-951-42-8928-6 (PDF)
ISSN 0355-3191 (Print)
ISSN 1796-220X (Online)

UNIVERSITY of OULU
OULUN YLIOPISTO