

# **The role of hybrids in the process of speciation; a study of naturally occurring *Formica* wood ant hybrids**

**JACK BERESFORD**

Evolution, Sociality and Behaviour  
Organismal and Evolutionary Biology Research Program  
Faculty of Biological and Environmental Sciences  
University of Helsinki  
Finland

Academic Dissertation

To be presented with the permission of the Faculty of Biological and Environmental Sciences,  
The University of Helsinki, for public examination in Viikki, Biocentre 2, on the 22nd of October  
2021 at 12 o'clock noon

Helsinki 2021

- Supervised by: **Dr. Jonna Kulmuni**  
Organismal and Evolutionary Biology, Faculty of Biological and Environmental Sciences, University of Helsinki, Finland
- Prof. Liselotte Sundström**  
Organismal and Evolutionary Biology, Faculty of Biological and Environmental Sciences, University of Helsinki, Finland
- Prof. Roger K. Butlin**  
Department of Animal and Plant Sciences, University of Sheffield, and, Department of Marine Sciences, University of Gothenburg, Sweden
- Reviewed by: **Dr. Philine Feulner**  
Center for Ecology, Evolution and Biogeochemistry, Swiss Federal Institute of Aquatic Science and Technology, Switzerland
- Prof. Jürgen Gadau**  
Institute for Evolution and Biodiversity, University of Münster, Germany
- Examined by: **Prof. Dieter Ebert**  
Zoological Institute, University of Basel, Switzerland
- Custos: **Prof. Johanna Mappes**  
Department of Biological and Environmental Sciences, University of Helsinki, Finland
- Members of the thesis advisory committee: **Prof. Craig Primmer**  
Organismal and Evolutionary Biology, Faculty of Biological and Environmental Sciences, and, Institute of Biotechnology, Helsinki Institute of Life Science (HiLIFE), University of Helsinki
- Dr. Ari Löytynoja**  
Institute of Biotechnology, Helsinki Institute of Life Science (HiLIFE), University of Helsinki

ISBN 978-951-51-7555-7 (paperback)

ISBN 978-951-51-7556-4 (PDF)

<http://ethesis.helsinki.fi>

Printed by Unigrafia

Helsinki 2021

The Faculty of Biological and Environmental Sciences uses the Urkund system (plagiarism recognition software) to examine all doctoral dissertations.

*'We used to make fun of Edgar Anderson by saying that he was finding hybrids under every bush. Then we realized that even the bushes were hybrids.'*

*- Warren H. Wagner*

## **TABLE OF CONTENTS**

<b>ABSTRACT</b>	<b>6</b>
<b>TIIVISTELMÄ</b>	<b>7</b>
<b>SUMMARY</b>	<b>8</b>
<b>1 INTRODUCTION</b>	<b>8</b>
1.1 Speciation, hybridization and the role of hybrids	8
1.2 The fitness consequences of hybridization in nature	9
1.3 Mate choice and its consequences for hybridization and gene flow	11
1.4 Looking for genes underlying selection	12
1.5 <i>Formica aquilonia</i> x <i>F. polyctena</i> hybrids in southern Finland	13
<b>2 AIMS OF THIS THESIS</b>	<b>16</b>
<b>3 METHODS</b>	<b>17</b>
3.1 Ant husbandry: sampling, housing and egg rearing	17
3.2 Marker loci under selection and their utility	19
3.3 Population genetics	20
3.4 Mate choice assays	21
3.5 Realized mating in nature & its consequences	21
3.6 RNA sequencing, transcriptome assembly and SNP calling	22
3.7 Genomic assembly using both long and short reads	22
<b>4 MAIN RESULTS AND THEIR INTERPRETATION</b>	<b>23</b>
4.1 Hybridization and its potential role in the future evolution of populations of wood ants	23
4.2 Developing resources for the study of admixture and selection in wood ants	30
<b>5 CONCLUSIONS</b>	<b>34</b>
<b>6 ACKNOWLEDGEMENTS</b>	<b>35</b>
<b>7 REFERENCES</b>	<b>35</b>

**Chapter I**

**Chapter II**

**Chapter III**

**Chapter IV**

## LIST OF THESIS CHAPTERS

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

**I Beresford J**, Elias M, Pluckrose L, Sundström L, Butlin RK, Pamilo P, Kulmuni J. (2017) *Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown*. *Molecular Ecology*. 26: 4013–4026.

**II Beresford J**, Ferkingstad B, Nygård E, Thorogood R, Kulmuni J. *Mate choice and its consequences for hybrid populations and their persistence*. Manuscript.

**III Beresford J**, Morandin C, Nouhaud P, Butlin RK & Kulmuni J. *Introgression of loci with signatures of selection is associated with differentially expressed genes in hybrid wood ants*. Manuscript.

**IV** Nouhaud P, **Beresford J**, Kulmuni J. *Cost-effective long-read assembly of a hybrid *Formica aquilonia* × *Formica polyctena* wood ant genome from a single haploid individual*. *BioRxiv*. Doi: <https://doi.org/10.1101/2021.03.09.434597>

## CONTRIBUTIONS

	Original Idea	Study Design	Data Collection	Performing analyses	First Manuscript draft	Final Manuscript Preparation
<b>I</b>	JK	JK	ME, PP, JK	<b>JB</b> , JK	<b>JB</b> , JK	<b>JB</b> , ME, LS, RB, PP, JK
<b>II</b>	<b>JB</b> , RT, JK	<b>JB</b> , BF, RT, JK	<b>JB</b> , BF, EN	<b>JB</b> , RT*	<b>JB</b> , JK	<b>JB</b> , RT, JK
<b>III</b>	JK	<b>JB</b> , RKB, JK	<b>JB</b> , JK	<b>JB</b> , CM**	<b>JB</b> , JK	<b>JB</b> , CM, PN, RKB, JK
<b>IV</b>	JK, PN	PN	<b>JB</b> , PN, JK	PN	PN	<b>JB</b> , PN, JK

\*) JB analyzed the realized mating data from the natural population and fitness consequences of realized matings. RT analyzed data from the mate choice experiment.

\*\*) CM led the transcriptome assembly, JB performed all other analyses including differential expression, SNP calling and GO analyses.

<b>JK:</b> Jonna Kulmuni	<b>CM:</b> Claire Morandin	<b>RT:</b> Rose Thorogood
<b>RKB:</b> Roger Butlin	<b>PN:</b> Pierre Nouhaud	<b>PP:</b> Pekka Pamilo
<b>LS:</b> Liselotte Sundström	<b>ME:</b> Marianne Elias	<b>EN:</b> Elisa Nygård
<b>JB:</b> Jack Beresford	<b>LP:</b> Lucy Pluckrose	<b>BF:</b> Bendick Ferkingstad

© Jack Beresford (Summary & Layout)

© The authors, *Molecular Ecology* (I)

© The authors, (II, III, IV)

© Jack Beresford (Illustrations)

# ABSTRACT

Hybridization and gene flow between diverging populations is widespread among taxa, it can homogenize previously diverging populations and is often viewed as a counteractive force to speciation. While much work has focused on how populations accumulate barriers to gene flow in the gradual process of speciation, hybrids themselves have often received less attention. In this thesis I investigate fundamental questions about hybridization between two wood ant species, *Formica aquilonia* and *F. polyctena*. I utilized multiple different methods, including bioinformatics, transcriptomics and gene expression, population genetics, and mating behavioural assays. In Chapter I, I found evidence of widespread hybridization between these two species in southern Finland. Over half of the populations were hybrid, and my data indicated that they have had many independent origins. Furthermore, hybrid populations had variable levels of admixture, suggesting that hybridization has been historic and is ongoing in this system. In Chapter II, I characterised mate choice in virgin queens of *F. aquilonia* and hybrids, and compared this to the realized mate-choice of old queens in nature. I also recorded egg laying and hatching rates of old queens to study the consequences of hybridization. This study revealed asymmetric preferences of *F. aquilonia* and hybrid queens, which could lead to the persistence of hybrid populations in nature. Egg laying rates did not differ, however the eggs of hybrid queen hatched 83% less than those of *F. aquilonia* revealing the costs of hybridization. In Chapter III, I studied gene expression patterns associated with introgressed alleles at candidate barrier loci known to be under selection during development. This was achieved by developing two *de novo* transcriptomes of distinct hybrid lineages from a long-term study population of hybrid wood ants. I found that only a fraction of the transcriptomes were differentially expressed in association with introgression. In Chapter IV, I contributed towards the *de novo* genome assembly from a single hybrid male, which leveraged the large body size of these ants and recent advances in sequencing technology to create a cost-effective reference genome. Collectively the chapters of this thesis contribute fundamental knowledge towards our understanding of this emerging model system, as well as reference genomic and transcriptomic resources for future studies to investigate admixture and selection. My results add to the growing evidence of widespread hybridization between diverging taxa, advance our understanding of the genetic basis of selection in this system, and open new avenues of research into the persistence of *Formica* hybrids and their role in the future evolutionary trajectories of these species.

# TIIVISTELMÄ

Eriytymässä olevien populaatioiden risteytyminen ja geenivirta niiden välillä on yleistä eri eliöryhmissä. Se voi yhtenäistää jo eriytyneitä populaatioita, ja sen on usein katsottu toimivan vastavoimana lajiutumislle. Vaikka lajiutumisen aikana populaatioiden välille kertyviä geenivirran esteitä on tutkittu paljon, risteymiin kiinnitetään usein vähemmän huomiota. Tässä tutkielmassa selvitän kahden kekomuurahaislajin risteytymiseen liittyviä peruskysymyksiä tupsukekomuurahaisella (*Formica aquilonia*) ja kaljukekomuurahaisella (*F. polyctena*). Käytän useita eri menetelmiä, mukaan lukien bioinformatiikkaa, transkriptomiikkaa ja geenien ilmentymistä, populaatiogenetiikkaa sekä parittelykäyttämisen analysointia.

Ensimmäisessä luvussa löysin todisteita näiden kahden lajin laajasta risteytymisestä Etelä-Suomessa. Yli puolet populaatioista oli risteymiä, ja aineistoni mukaan risteytymistä on tapahtunut itsenäisesti useita kertoja. Lisäksi risteymäpopulaatioiden genomeissa oli vaihteleva määrä sekoittumista, mikä viittaa siihen, että risteytyminen on sekä vanhaa alkuperää että edelleen tapahtumassa. Toisessa luvussa kuvasin nuorten neitsytkuningatarten parittelukumppanin valintaa sekä tupsuloviniskalla että risteymillä, ja vertasin sitä vanhojen kuningatarten luonnossa tapahtuneeseen parinvalintaan. Mittasin myös vanhojen kuningatarten muninta- ja kuoriutumismääriä tutkiakseni risteytymisen seurauksia. Tämä tutkimus paljasti, että tupsukekomuurahaisen ja risteymien kuningattarilla on epäsymmetriset parinvalintamieltymykset, mikä saattaa johtaa risteymäpopulaatioiden säilymiseen luonnossa. Munintamäärissä ei ollut eroa, mutta risteymäkuningatarten munista kuoriutui 83% vähemmän kuin tupsukekomuurahaisen munista, paljastaen risteytymisen kustannukset. Kolmannessa luvussa tutkin geenien ilmentymiskuvioita liittyen toisesta lajista siirtyneisiin alleeleihin niissä estelokuksissa, joihin valinnan tiedetään vaikuttavan kehityksen aikana. Tein tämän kehittämällä *de novo* -transkriptomit kahdesta erillisestä risteymälinjasta pitkäaikaisessa seurantapopulaatiossamme. Sain selville, että ainoastaan murto-osaa transkriptomeista ilmennettiin eri tavalla toisen lajin alleeleihin liittyen. Neljännessä luvussa osallistuin yhden risteymäkoiraan *de novo* -genomin kokoamiseen, missä käytettiin hyväksi näiden muurahaisten suurta kokoa ja sekvensointiteknologian viimeisintä kehitystä kustannustehokkaan viitegenomin luomiseksi.

Kokonaisuutena tämän tutkielman osiot antavat perustietoa, joka auttaa meitä ymmärtämään tätä uutta mallisysteemiä, sekä luovat viitegenomi- ja transkriptomiresursseja tulevien sekoittumis- ja valintatutkimusten käytettäväksi. Tulokseni lisäävät todisteita siitä, että eriytyessä olevien eliöryhmien risteytyminen on yleistä, vievät eteenpäin ymmärrystämme valinnan geneettisestä taustasta tässä systeemissä ja avaavat uusia mahdollisuuksia kekomuurahaisten risteymien säilyvyydestä ja niiden roolista näiden lajien tulevassa evoluutiossa.

# SUMMARY

## 1 INTRODUCTION

### 1.1 *Speciation, hybridization and the role of hybrids*

The often gradual and continuous process of speciation gives rise to discontinuous biodiversity, i.e. groups of phenotypically similar individuals, whose similarity is based upon genetic variation. We humans habitually enjoy the categorization of all things, including phenotypically similar groups of organisms into species, to give structure and order to the world we inhabit. Biologists too rely upon Linnean concepts of species, for example, the ever popular Biological Species Concept: that groups of organisms that can successfully reproduce and create fertile, viable offspring, into the units we know as species ([Mayr 1963](#)). However, the reality of speciation does not fit our neat categorization of taxa, because hybridization between apparently divergent species or populations is commonplace. Indeed, speciation is often defined as: the process of evolving reproductive isolation through the accumulation of barriers to gene flow, such that sympatrically located populations can retain their distinctiveness ([Seehausen et al. 2014](#)). Hence, hybridization denotes places in the speciation continuum where such barriers to gene flow are not yet complete. What's more, modern genomics has revealed that hybridization is a powerful force that brings together novel combinations of genetic material. How hybrids experience selection, and whether they persist and promote further gene flow, can shape the evolutionary trajectory of species, and of hybrids themselves.

Evidence of hybridization between closely related species has been growing ([Mallet et al. 2005](#)), demonstrating that our described species are not so conveniently delineated by reproductive barriers. Further evidence from recent genome sequencing has revealed widespread hybridization throughout the tree of life ([Mallet et al. 2016](#)). In our modern anthropocene era, human activity such as anthropogenic climate change that causes species range shifts, and the introduction of invasive species, increasingly results in the secondary contact of once geographically isolated (allopatric) species, thus promoting hybridization ([Todesco et al. 2016](#), [Taylor & Larson 2019](#)). Hybridization can have dramatic and unpredictable consequences for natural populations. Sometimes hybridization results in the creation of novel species (e.g. [Rieseberg et al. 1995](#)), homogenizes divergence between previously distinct parental species ([Todesco et al. 2016](#)), or

even promotes reinforcement and selection for associative mate-choice mechanisms ([Servedio & Noor 2003](#), [Coughlan & Matute 2020](#)). From a broader perspective, the consequences of hybridization can be detrimental for the persistence of species, and is often cited as a cause for the reduction of biodiversity ([Rhymer & Simberloff 1996](#), [Todesco et al. 2016](#)), but hybridization can also add to or reshape existing biodiversity - a view that is steadily gaining traction ([Abbott et al. 2013](#), [Mallet 2007](#)). For example, hybrid populations may persist, acting as reservoirs for genetic variation ([Grant & Grant 2019](#)), and sometimes they can facilitate gene flow of adaptive genetic variants between parental gene pools ([The Heliconius Genome Consortium 2012](#), [Fontain et al. 2015](#)). Hybrid populations can even be fitter than populations of parental species ([Griebel et al. 2015](#)). Given these broad and dramatically different evolutionary outcomes of hybridization for which we have only just scratched the surface, new studies of hybridization are both important and timely, especially in the face of decreasing global biodiversity ([Wagner et al. 2021](#)). We especially lack studies from a broad array of taxa that consider the perspective of natural hybrid populations, their role in gene flow and their impact on the evolutionary trajectories of species. Moreover, how will species evolve and persist in the face of increasing hybridization?

## 1.2 *The fitness consequences of hybridization in nature*

Hybridization generates novel combinations of alleles from previously diverging populations. These novel combinations are then tested within the genomes and the environment of hybrids, and may have adaptive, neutral or deleterious fitness consequences ([Abbott et al. 2013](#)), though it's likely that hybrids experience a combination of these effects. The fitness of individuals, be they hybrid or not, is typically defined by their ability to survive, successfully find a mate, and contribute offspring to the next generation. Thus in hybrids their fitness is often measured relative to that of their parental species.

Deleterious combinations of alleles that result in hybrids of low fitness are grouped together as post-zygotic reproductive barriers to gene flow. These barriers can be split into extrinsic and intrinsic barriers ([Coyn & Orr 2004](#), [Seehausen et al. 2014](#)). Extrinsic barriers result from deleterious environment x genome interactions, for example, hybrids may be maladapted to their environment, or be unable to entice a mate in order to reproduce. Barriers resulting from intrinsic factors are caused by incompatibilities between the allelic combinations themselves, and can result in full or partial hybrid sterility, or hybrid breakdown and death. Incompatibilities that result in hybrid breakdown and sterility are often caused by epistatic interactions between

genes, and are known as Bateson-Dobzhansky-Muller incompatibilities (BDMIs; [Orr 1995](#)). In natural populations, we expect that F1 generation hybrids experience selection against dominant alleles (i.e. alleles that cannot be masked by heterozygosity). However, F1 hybrids may also experience heterosis as a result of outbreeding and increased genetic diversity throughout their genomes. F2 and later generations reveal recessive incompatibilities that become homozygous in some individuals, those primarily predicted to arise by the BDMI model ([Orr 1996](#), [Orr & Turrelli 2001](#)), and potentially severely impacting the fitness of hybrids. However, while the initial burden of hybridization may be severe, in later generations selection is expected to remove such deleterious combinations of alleles from hybrid populations, potentially leaving novel and adaptive combinations of alleles.

Given the frequency of hybridization in nature, particularly between recently diverged populations ([Mallet et al. 2016](#), [Grant & Grant 2019](#)), the cost of hybridization may be relatively low compared to the adaptive fitness benefits. For example, in *Heliconius* butterflies, mimicry wing patterns have historically introgressed between geographically sympatric populations of now divergent species ([The Heliconius Genome Consortium 2012](#)). Interestingly in the case of *Heliconius*, it's been shown that the introgression of loci controlling wing patternation occurred at multiple, physically unlinked loci in their genomes, meaning that these loci have introgressed (moved between populations through hybridization and selection) and been weeded out from their surrounding genomic material ([The Heliconius Genome Consortium 2012](#)). Introgressed alleles may confer adaptive benefits in other ways, for example, it's likely that modern Tibetan human populations have inherited high-altitude adapted genes from ancient *Denisovan*-humans ([Huerta-Sánchez et al. 2014](#)), that beneficially affect the structure and oxygen binding affinity of their hemoglobin.

Further to the points mentioned above, novel allelic combinations that arise in hybrids may also act differently in males and females. This is typically seen as differential fitness consequences for males and females, often associated with sex chromosomes ([Presgraves 2003](#)) or ploidy number in haplodiploids ([Koevoets & Beukeboom 2009](#)). This remarkably consistent phenomena manifests (usually, [Bendall et al. 2020](#)) as deleterious selection against in the heterogametic sex (i.e. the sex with one chromosome as opposed to two). Known as Haldanes rule ([Haldane 1992](#)), one major cause of this phenomena is recessive incompatibilities being "revealed" on heterogametic chromosomes, since heterogametic individuals lack compensatory alleles on their haploid sex chromosomes, or genome-wide in the case of haplodiploids ([Koevoets & Beukeboom 2009](#)). As such the fitness consequences of hybridization may be different in males and females.

### 1.3 Mate choice and its consequences for hybridization and gene flow

Mate choice is a key component of speciation that can promote or restrict hybridization and gene flow between diverging species, and is a key pre-mating barrier to gene flow during secondary contact ([Abbott et al. 2013](#)). We know that mate choice is driven by sexual selection, often acting on genetically determined traits ([Chenoweth & Blows 2006](#)). The strength of mate choice barriers to gene flow is expected to be dependent on the strength of sexual selection acting upon mate choice preferences ([Ritchie et al. 2007](#)). Importantly however, mate choice may not always evolve to favour conspecifics ([Pfenning 2007](#)), though where it does this can drive divergence and speciation by promoting assortative mating in areas of sympatry. When hybrids are strongly selected against, selection promoting mate choice and assortative mating may arise through reinforcement ([Servedio & Noor 2003](#), [Matute et al. 2010](#)). Crucially, where natural hybrid populations persist, mate choice behaviour of parental and hybrid individuals can shape the future evolutionary direction of gene flow in sympatric populations. However, curiously, it is the mate choice behaviour of hybrids that is often overlooked in studies of hybridization.

Mate choice can be based on a combination of visual, auditory and chemical/chemosensory cues ([Greenspan & Feveur 2000](#)), and depends upon the successful interaction of potential mates, often with males giving cues to females. Both mate choice and cues can diverge rapidly between even closely related species ([Colyott et al. 2016](#)), which can result in assortative mating during secondary contact, and forming barriers to gene-flow ([Abbott et al. 2013](#)). However, mate-choice is not a fixed barrier to gene flow, and in some species this barrier can break down due to changing or disturbed environments. For example, assortative mating in swordtail fishes breaks down under manipulated laboratory conditions, or in rivers that are highly disturbed, such that pollutants reduce water visibility ([Schumer et al. 2017](#)). Mate choice may also not evolve to favour conspecifics, as exemplified in the spadefoot toad, *Spea bombifrons*. Female toads have evolved facultative preferences that depend on the environment. They prefer conspecific males in deep, long-lasting pools, but prefer heterospecific males in shallow pools ([Pfenning 2007](#)). Hybrid tadpoles develop faster than pure *S. bombifrons* tadpoles, meaning that they are more likely to survive in the ephemeral shallow ponds that are prone to drying out. Hence selection favours hybridization in this context, though a proportion of the resultant hybrid males are typically sterile. In another example of heterospecific preferences, the copepod *Tigriopus californicus* has evolved inbreeding avoidance, preferring mates from different populations, likely because sibling matings result in inbreeding depression ([Palmer & Edmands 2000](#)). Copepod

hybrids are readily formed in nature, but while F1 hybrids experience heterosis, F2 hybrids experience severe dysfunction. In spite of this, these copepods show no signatures of outbreeding avoidance ([Palmer & Edmands 2000](#)). In summary, mate choice can result in assortative mating, however preferences can be context specific, not always adaptive, nor always act as a permanent barrier to gene flow.

However, as mentioned above, the mate-choice preferences of hybrids can also play a role. Hybrids can have novel mating phenotypes (reviewed in [Rosenthal 2013](#)), often resulting from novel epistatic interactions between the genetic basis of traits and preferences, which can have unexpected outcomes. For example, in the aforementioned hybrid swordtail fish, once water clarity was restored, hybrids began to assortatively mate, forming distinct clusters with their own mating phenotypes ([Schumer et al. 2017](#)). As such, the mating preferences of swordtail fishes maintain two distinct hybrid lineages, at least temporarily. However, whether these preferences and the resultant hybrid clusters reflect the bimodal distribution of the original parental lineages is unclear. Nevertheless the swordtail fish system demonstrates how genetically determined mate choice preferences can result in barriers to gene flow, and promote the persistence hybrid populations. On the other hand, *S. bombifron* hybrids exhibit no preferences, except in deep ponds, where they prefer the calls of (mostly) sterile hybrid males ([Schmidt & Pfenning 2016](#)). Maladaptive hybrid preferences such as these, have been suggested to act as reproductive isolating mechanisms in a number of systems ([Noor et al. 1997](#), [Russell & Magurran 2006](#), [Svedin et al. 2008](#), [Clark et al. 2010](#), [Lemmon & Lemmon 2010](#), [Latour et al. 2014](#)). Thus the resultant preferences of hybrids can be unpredictable, they may be adaptive or maladaptive depending on the context. Hybrid mate choice may itself result in a form of reproductive isolation, or it may facilitate the persistence of hybrid populations and gene flow by erecting permeable barriers to gene flow.

#### 1.4 Looking for genes underlying selection

Characterizing the genetic variants under selection in divergent and admixed populations is a central part of the study of hybrids, hybridization and speciation. In particular, identifying the genetic regions under selection, the genes and loci residing within, and how their function affects fitness is of primary interest and a major challenge. Earlier studies have historically relied upon genomic markers to study genetic variation and selection. Modern studies now utilize genome scans to compare hybrid, or divergent populations, scanning for genomic regions which are fixed for different alleles (so called islands of

speciation or divergence)([Turner et al. 2005](#), [Harr 2006](#), [Nosil et al. 2009](#)), regions of introgression, and regions of admixture in hybrids. Studies of recently diverged species pairs have revealed heterogeneous patterns of divergence across their genomes (e.g. [Nadeau et al. 2012](#)) highlighting the significance of developing genomic resources that facilitate the study of genome-wide patterns of admixture, as opposed to genetic markers which describe only a few regions. The knowledge, skills and tools required to build such resources, while complex, are becoming cheaper and more accessible. In spite of this, studies of admixture and gene flow are often restricted to model organisms with well curated genomic resources, though this is gradually changing.

The aforementioned divergence between populations occurs as mutations and variants are sorted and become fixed in the genomes of isolated populations. Such mutations can occur in both coding and regulatory regions of genomes. A significant proportion of genetic differentiation is expected to accumulate in regulatory regions ([Mack & Nachman 2017](#)), affecting the expression of genes, especially early in divergence ([Pavey et al. 2010](#), [Jones et al. 2012](#), [Wittkopp & Kalay 2011](#)). Novel combinations of genes and regulatory variants in hybrids can result in changes in gene expression, thus the study of hybrid gene expression is a way to study the genes under selection in hybrids. Studies of gene expression in hybrids have predominantly focused on characterizing gene expression differences in early generation (i.e. F1, F2) hybrids compared to parental individuals (e.g. [Tirosch et al. 2009](#)), or across hybrid zones (e.g. [Mack et al. 2016](#)). Gene expression in F1 hybrids is expected to be intermediate between the parents' overall expression patterns, but may also be influenced by both heterosis and incompatibilities that impact gene regulation ([Landry et al. 2007](#), [Mack & Nachman 2017](#)). As explained above, we expect incompatibilities caused by dominant alleles to be revealed in F1 hybrids, while recessive incompatibilities are revealed only in F2 and later generation hybrids where segregation and recombination create homozygotes. Gene expression in these generations may depart further from the parents, or from the intermediate expectation, but the variance in expression among individuals is also expected to increase because of the many different genotype combinations present in later generation hybrids. This might make detection of differential gene expression associated with introgression more challenging in later generation hybrids, but this is unclear from present studies.

### 1.5 *Formica aquilonia* x *F. polycтена* hybrids in southern Finland

In my thesis I study two species of mound-building wood ants of the *Formica rufa*-group; *F. polycтена* and *F. aquilonia*, and their hybrids, in southern Finland.

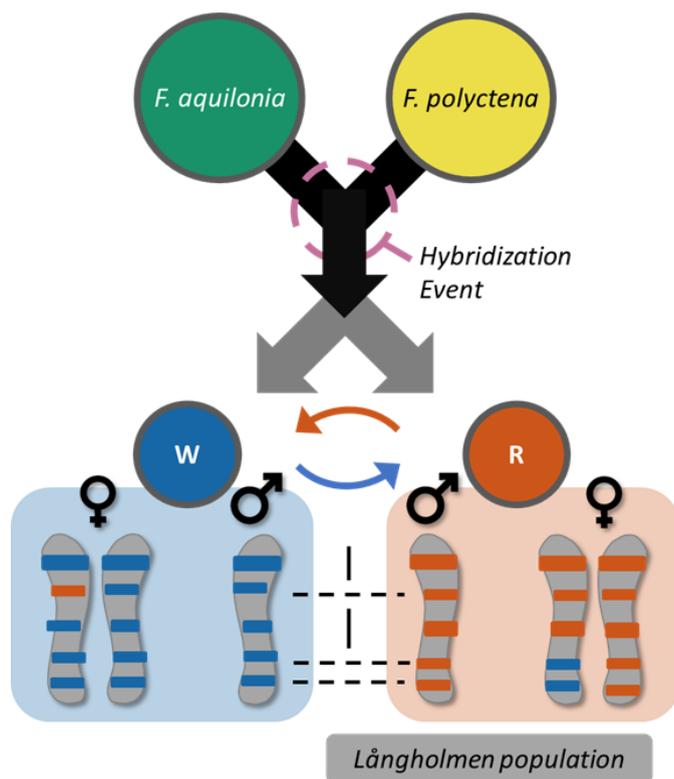
*F. polyctena* is predominantly distributed across central Europe, with the Southernmost extent of their range abutting the European mountain ranges (the Pyrennes, Alps and the Balkans), and their Northernmost range extending into central Sweden and Southern Finland ([Stockan et al. 2016](#)). *F. aquilonia* resides at higher latitudes; throughout Scandinavia and the Baltic states (Estonia, Latvia and Lithuania), as well as the Northern regions of the UK and an enclave in the Alps ([Stockan et al. 2016](#)). Hybrids form naturally between *F. aquilonia* and *F. polyctena* in the forests of southern Finland ([Kulmuni et al. 2010](#)), where the ranges of both species overlap. Our suspicion, though yet to be confirmed, is that *F. polyctena* has migrated to Finland from central Europe or from otherwise more southerly latitudes, resulting in secondary contact and hybridization. *F. aquilonia* and *F. polyctena* are part of a recently diverged clade, they are estimated to be 500K years diverged from one another, and they are not sister species ([Goropashnaya et al. 2012](#)).

Previous work has demonstrated that two genetically distinct and stable lineages of hybrid wood ants (named W and R) live within a single population called Långholmen ([Kulmuni et al. 2010](#), [Ghenu et al. 2018](#)). These lineages share a mitochondrial haplotype and thus likely originate from a single hybridization event ([Kulmuni et al. 2010](#)), though their genomes are now a makeup of alleles from both parentals. These two lineages appear to be maintained within the population with little gene flow between them, or into the population as a whole ([Kulmuni et al. 2020](#)). Their population sizes are likely different as well: W occupies some 20 or more nests, while R is a minority, inhabiting only 3 nests. Genome-wide estimates of  $F_{ST}$  between the lineages is 0.14 between males, and 0.07 between females ([Kulmuni et al. 2020](#)), revealing clear differences between the W and R lineages, though the exact contribution from each of the parental species to either lineage is unknown. Their genomes are composed of a mixture of loci with fixed differences between the lineages, while other loci are polymorphic for shared alleles ([Kulmuni & Pamilo 2014](#)).

Ants, like other hymenopteran species, have a haplodiploid sex determination system, meaning that males are haploid and develop from unfertilized eggs, while females are diploid, and develop from fertilized eggs. The haploid males of W and R are fixed for alternative alleles at loci that are not fixed in females, but that are often heterozygous (Figure 1). This pattern is indicative of negative selection in hybrid males (i.e. against introgressed alleles), and is reflected by allele frequency changes between early and late developmental stages of males and queens ([Kulmuni & Pamilo 2014](#), [Kulmuni et al. 2020](#)). Thus these loci are referred to as candidate male barrier loci, and their study allows us to capture evidence of selection “in action” in a natural hybrid population. Currently it is not known what genes and or factors are involved in the selection

acting in this population (though I went looking, see **Chapter III**). However, one recent study has found that introgressed allele frequencies in males correlate with mean spring temperatures in the area (i.e. temperatures that might affect male development) ([Martin-Roy et al. 2021](#)), suggesting that extrinsic factors could play a role.

We know that the Långholmen population has existed for decades already ([Pamilo 1993](#)), and the expectation is that the most deleterious combinations of alleles should be removed by selection in late-generation hybrids ([Runemark et al. 2018](#)). Yet we still see selection acting during the development of males, while females are often heterozygous for the same alleles. This phenomena has been observed in hybrid zones/clines, where such alleles might be maintained within a hybrid population through a balance of migration and selection ([Barton & Hewitt 1985](#), [Bazykin 1967](#)). Yet, the absence of gene flow into this population ([Kulmuni et al. 2020](#)) likely means that polymorphic loci under selection in these hybrids are maintained by a balance of positive and negative selection. This suggests that heterozygous females experience positive selection.



**Figure 1** The hybrid population in Långholmen are formed from a single hybridization event between *F. aquilonia* and *F. polycytena*, which resulted in two independent hybrid lineages, named W and R. At the bottom of the figure, a schematic representation of the adult genotypes from W and R is depicted by the grey chromosomes. These lineages have loci that were previously fixed differences between males of the two lineages (indicated in the figure with a dashed line) that are sometimes heterozygous in females of their lineage, though some of these are no longer fixed ([Kulmuni et al. 2020](#)).

## 2 AIMS OF THIS THESIS

My thesis consists of four chapters which examine diverse but related aspects of hybrids and hybridization between *F. aquilonia* and *F. polyctena*, ranging from gene expression and bioinformatics, populations genetics and behavioural analyses (Figure 2). Together, the chapters of my thesis aim to contribute fundamental knowledge and resources for this emerging hybrid model system. In **Chapter I** I surveyed populations of *F. polyctena* and *F. aquilonia* using microsatellite and mitochondrial DNA markers, in search of both hybrid and parental populations. In **Chapter II** I delved into the world of premating barriers to gene flow by studying the mate choice behaviour of hybrid and parental *F. aquilonia* queens, as well as their realized matings in nature, and the consequences of mate choice for hybrids. The third and fourth chapters build genomic and transcriptomic resources for the future study of this system. In **Chapter III** I studied the genes underlying selection in developing hybrid larvae by examining their gene expression patterns. In **Chapter IV** I contributed data towards the first assembled and annotated reference genome of a single *F. aquilonia* x *F. polyctena* hybrid male.

	Aims & Questions	Key results
I	<p><b>AIM: To identify hybrid and parental populations</b></p> <ol style="list-style-type: none"> <li>1. Are there other hybrid populations in southern Finland?</li> <li>2. Do hybrid populations have a single or multiple origins?</li> <li>3. Do other hybrid populations exhibit patterns of selection akin to those seen in Långholmen?</li> </ol>	<ol style="list-style-type: none"> <li>1. There is widespread signatures of hybridization across southern Finland, with no systematic geographical pattern.</li> <li>2. Evidence of multiple hybrid origins.</li> <li>3. At least one other hybrid population shows the systematic allele differences associated with selection in the Långholmen population.</li> </ol>
II	<p><b>AIM: To examine the mate choice preferences of hybrids and parentals, and to ascertain the impact of this choice</b></p> <ol style="list-style-type: none"> <li>1. Do hybrids and parental <i>F. aquilonia</i> queens exhibit patterns of mate choice?</li> <li>2. What are the realized matings in nature?</li> <li>3. What are the consequences of hybridization for egg laying and hatching rates?</li> </ol>	<ol style="list-style-type: none"> <li>1. Hybrids and parental <i>F. aquilonia</i> queens can discriminate</li> <li>2. Realised matings appear to have all occurring within the respective hybrid and <i>F. aquilonia</i> populations</li> <li>3. Egg laying rates did not differ between hybrid and <i>F. aquilonia</i> queens, but 83% fewer eggs hatched from hybrid queens, a strong signal of viability selection</li> </ol>
III	<p><b>AIM: To characterize gene expression patterns underlying selection in developing hybrid larvae</b></p> <ol style="list-style-type: none"> <li>1. What are the gene expression patterns associated with introgressed loci under selection?</li> <li>2. Are genes found near candidate male barrier loci differentially expressed?</li> </ol>	<ol style="list-style-type: none"> <li>1. Only a small number of all genes were differentially expressed genes in association with introgression at candidate male barrier loci.</li> <li>2. Genes near candidate male barrier loci under selection were not themselves differentially expressed.</li> </ol>
IV	<p><b>AIM: To develop a high quality reference genome</b></p> <ol style="list-style-type: none"> <li>1. Can we assemble a high quality reference genome from a single haploid male?</li> <li>2. To what extent does the genome of <i>Formica</i> hybrids represent the conserved gene set of other related species?</li> </ol>	<ol style="list-style-type: none"> <li>1. A cost-effective genome was assembled using long-sequence reads from a single haploid male.</li> <li>2. 97% of conserved single copy orthologs (BUSCOs) were recovered from the cleaned genome.</li> </ol>

**Figure 2** Thesis aims and main results

## 3 METHODS

### 3.1 *Ant husbandry: sampling, housing and egg rearing*

Field samples used in the chapters of my thesis were collected across multiple years (2001–2019). My supervisor, our group and I collected samples of different life stages and castes, including workers, old queens (>1 year), virgin queens and males (alates), as well as pupae which we reared until alates emerged. In every chapter we collected samples from the long-term hybrid population, Långholmen (Figure 3). We collected samples from other hybrid populations, as well as parental populations of *F. aquilonia* and *F. polycтена* from across southern Finland and the Åland archipelago (see the maps of **Chapters I & II**). In addition to these field samples, I also obtained mitochondrial sequences (**Chapter I**) from GenBank ([Benson et al. 2013](#)).

In **Chapter I** my coauthors and I survey the breadth of hybridization throughout southern Finland using 647 workers from 96 nests collected from 16 localities. In order to look for allele frequency differences between sexes, akin to those found by Kulmuni & Pamilo ([2014](#)), we also sampled alates (winged queens and males) at these localities.

In **Chapter II**, to study mate choice, virgin males and females from hybrid and parental *F. aquilonia* populations, and male alates of *F. polycтена* populations (we could not find queens) from across southern Finland and the Åland archipelago (see Figure 1. **Chapter II**) were sampled during 2018. To study realized mate-choice, old queens were collected from the Långholmen hybrid population and one parental population (Pusula) during 2019. All samples were kept in the laboratory with nest fragments (a handful of workers and nest material), which were fed and watered daily. Old queens were isolated on petri-dishes after acclimatizing to laboratory conditions (Figure 4), to track the number of eggs they laid and the proportion that hatched. I later dissected the queens to obtain their spermathecae for genotyping (Figure 5).

In **Chapter III**, I used hybrid larvae (n=173) sampled during 2014, from which we extracted both DNA and RNA for sequencing. Larvae were flash-frozen using dry ice within 24 hours of sampling to preserve their gene expression patterns.

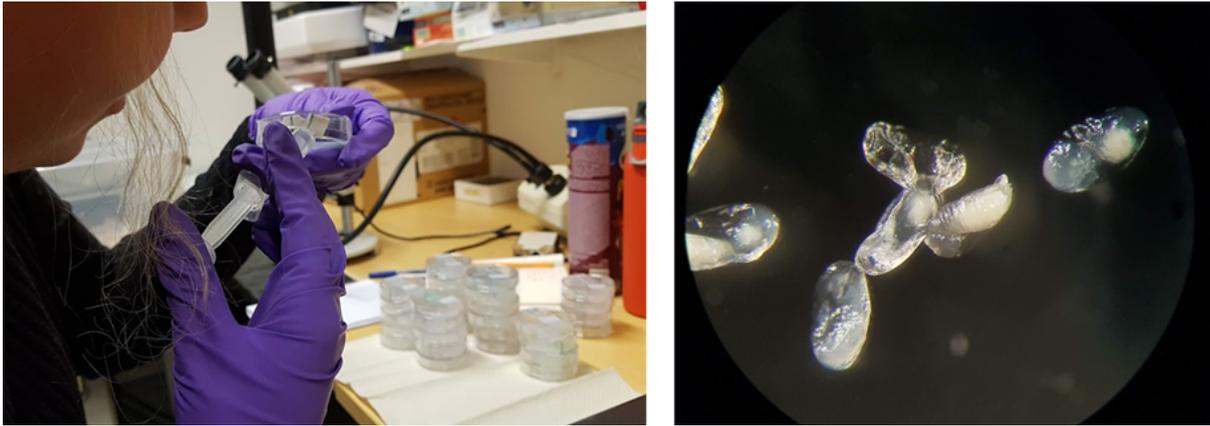
In **Chapter IV**, a single hybrid male was sampled from Långholmen for long-read (PacBio) sequencing, along with two unmated hybrid gynes (queens, also from Långholmen) for short read sequencing. All three samples were sequenced in association with the Global Ant Genomics Alliance (GAGA, [Boomsma et al. 2017](#)). Samples were flash-frozen upon collection. Female short-reads were used to refine the pac-bio genome assembly, and RNAseq samples from **Chapter III** were

used for annotation (assigning putative gene-coding regions) (see below, 3.7 Genomic assembly using both long and short reads).

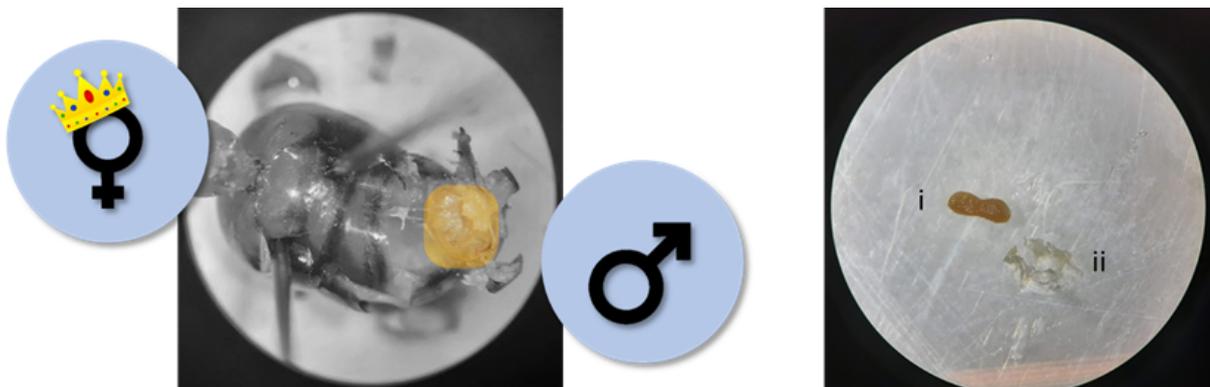
**Ethical Note:** There are no rules for the treatment of *Formica* ant species. Nevertheless these are living creatures. Furthermore, *Formica* may well be negatively affected by habitat disturbances caused by anthropogenic change and landscape management, and both parental species are listed as near-threatened (IUCN Red List of Threatened species: [Formica aquilonia 1996](#), [Formica polyctena 1996](#)). We adhered to ASAB guidelines for the ethical treatment of animals while collecting and handling these ants. We collected only a fraction of the queens, workers and brood from nests in order to minimize our impact upon nests. We also fed and watered our ants in the laboratory daily, and handled them with care. Old queens were euthanized before dissection.



**Figure 3** Collecting old *Formica* hybrid queens from the long-term study population in Långholmen, next to the Tvärminne Zoological Station (samples used in **Chapter II**). Left, Jonna Kulmuni gets stuck into a nest. Centre, a wash basin we used to help isolate queens from their relentless workers. Right, Heikki Helanterä showing Jonna & I the ropes. Photos © Jack Beresford.



**Figure 4** Photos of egg rearing used in **Chapter II**. **Left:** Old queens were isolated on petri-dishes while they laid eggs. Cloth strips were placed into the lips of the petri-dishes and wetted daily to prevent queens or eggs drying out without risking fungal infections in the queens or their eggs. **E. Right:** Eggs were then maintained on petri-dishes until they hatched. I put cloth strips into the lips of the petri-dishes which were moistened daily to keep queens hydrated. The left picture shows E. Nygård wetting the cloth strips. Photos © Jack Beresford.



**Figure 5** Photos showing spermathecae dissection from old queens collected from the field (**Chapter II**). **Left:** Under a light microscope I dissected out spermathecae, the sperm storage organ inside of queens. In the picture, the spermatheca is the “jelly-bean” shaped organ highlighted with colour, while the queen abdomen is in black and white. **Right:** Spermathecae (i) are contained within a transparent sack (ii) of the queen’s own tissue, which I removed prior to genotyping. Photos © Jack Beresford.

### 3.2 Marker loci under selection and their utility

Prior to and during my thesis, studies have characterised both microsatellite (hereafter referred to as msat) ([Kulmuni et al. 2010](#)) and SNP marker loci ([Kulmuni et al. 2020](#)), which have been shown to be associated with signals of selection in hybridizing *Formica* ants found in the Långholmen population ([Kulmuni & Pamilo 2014](#), [Kulmuni et al. 2020](#), [Martin-Roy et al. 2021](#)). These marker loci are an integral part of **Chapters I, II & III** of this thesis.

In **Chapters I & II** of this thesis I used the msat marker loci to genotype and study the genetic structure of my samples. These msat markers are powerful tools since not only are there introgressed alleles associated with selection, but

they also have alleles that are diagnostic for either the W or R lineage in the Långholmen population ([Kulmuni et al. 2010](#)). In total there are 9 established msat loci, and importantly, it is unlikely to be the msat alleles themselves that are directly under selection, but within genomic regions containing loci that are. These loci also (hint at the results) separate the hybrid and parental species remarkably well (**Chapter I**). In **Chapter III** I utilized the power of these msat loci to choose samples for RNA sequencing. By genotyping my larvae samples (n=173), not only could I choose samples from the W and R lineages, I also chose samples with and without introgressed microsatellite alleles known to be under selection. In the year that we sampled these larvae, we also know that selection occurred between early and late stages of development ([Kulmuni et al. 2020](#)). I also used the msat markers to sex the larvae, which are morphologically indistinguishable at the larvae stage. This is possible because males are haploid and females are diploid, meaning males are homozygous at all 9 loci.

In **Chapter III** I also used the candidate male barrier loci (the loci with fixed differences between W and R males that are indicative of viability selection, n=637) to quantify introgression in my larvae. This chapter utilized RNAseq (described below), which I mapped to the *F. exsecta* genome ([Dhaygude et al. 2019](#)), as was done by Kulmuni et al. ([2020](#)). Using their map to the *F. exsecta* genome I identified and counted the amount of introgressed alleles at candidate male barrier loci in my larvae samples. In addition to this, Kulmuni et al. ([2020](#)) characterized putative candidate genes, which were genes within 2500+/- bp of candidate male barrier loci, based on the *F. exsecta* genome annotations ([Dhaygude et al. 2019](#)). I examined whether these genes were differentially expressed (see below).

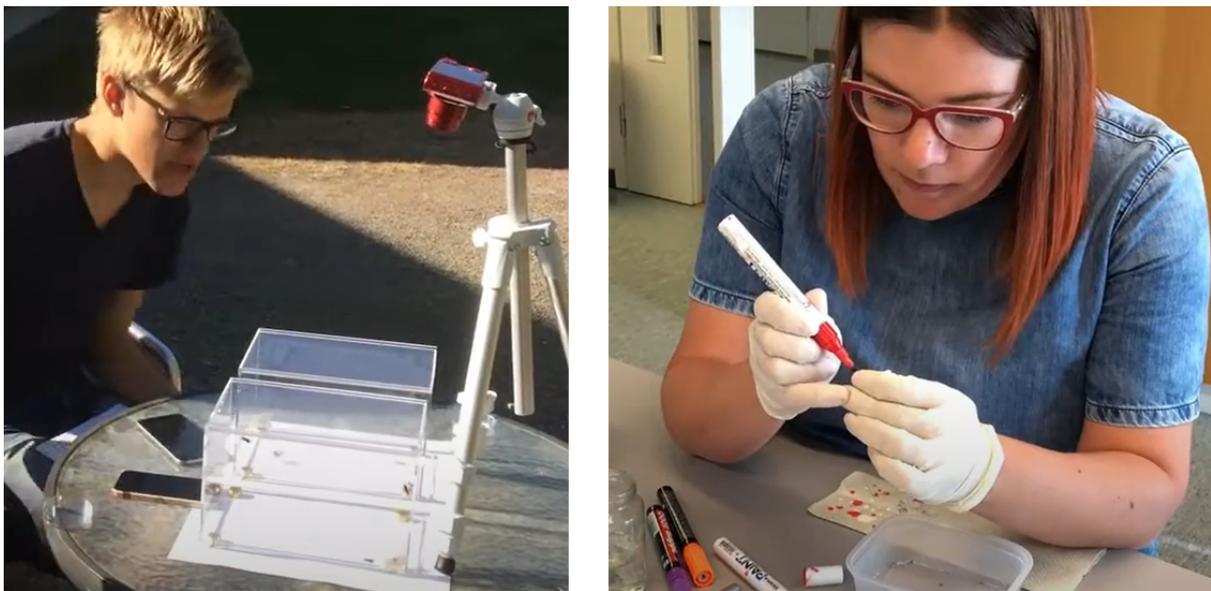
### 3.3 Population genetics

Population genetics methodologies are an integral part of **Chapters I & II**. Using the aforementioned microsatellite marker loci established by Kulmuni et al. ([2010](#)), I surveyed the population genetic structure of populations across Southern Finland (**Chapter I**). I also used these markers to examine realized mate choice (**Chapter II**). For this latter project, I used population genetic analysis to examine the combination of female genotypes and the males they mated with (by genotyping their spermathecae), in one hybrid and one parental population of *F. aquilonia*. Overall admixture and population genetic structure was analyzed using the program Structure ([Pritchard et al. 2000](#)) along with a method called  $\Delta K$  ([Evanno et al. 2005](#)) to examine the optimal number of predicted clusters among samples. I also used the R package Hierfstat ([Goudet 2005](#)). In both

chapters of my thesis Fisher's exact test ([Fisher 1922](#)) was used to test for differences in allele frequencies between males and females.

### 3.4 Mate choice assays

In **Chapter II** we used mate choice assays to assess the female mate choice of hybrid and *F. aquilonia* queens. Females were given different pairwise combinations of males to choose from, and their behaviour was recorded using a video camera. Mate choice assays were conducted during bright sunny days in clear acrylic boxes, to mimic the natural conditions of *Formica* spp. mating, i.e. a mating swarm, where females or males can use pheromones to attract mates of the opposite sex (see reference Hölldobler & Bartz 1985, in [Fortelius 2005](#)). Of note, *F. aquilonia* and *F. polyctena* lack a clear mating system description, though it has generally been described in two ways; i) alates gathering at well-defined places, or ii) as females mating on or near their natal nests (see [Fortelius 2005](#)). Female willingness to mate, the number of mating attempts each male made, the number that resulted in copulations and their duration were recorded. More details of the models and statistical analyses can be found from **Chapter II**.



**Figure 6** Mate choice experimental set up (**Chapter II**). **Left:** B. Ferkingstad setting up mate-choice experiments, trapping females in acrylic boxes with pairwise combinations of males. **Right:** J. Kulmuni marking queens and males with colour identifiers.

### 3.5 Realized mating in nature & its consequences

To study realized mating choice in nature I examined old queen (>1 year) genotypes and that of their spermathecae (Figure 5) using 9 previously mentioned microsatellite loci ([Kulmuni et al. 2010](#)). By studying the combination of female and male genotypes (from the spermathecae) I could determine with

whom each female had mated, and the extent of gene flow into their respective populations. For each queen I also recorded the egg laying and hatching rates (Figure 4), to examine the consequences of their mate choice.

### 3.6 RNA sequencing, transcriptome assembly and SNP calling

RNA sequencing (or RNA-seq) is a way to profile the transcriptomes (i.e. the gene expression levels) of organisms without a prior knowledge of the genes of an organism ([Wang et al. 2009](#)). Because this method relies upon the sequencing of messenger RNA, it is possible to assemble the sequenced reads into *de novo* transcriptomes (a representation of the expressed genes among samples), thus making it a useful method for studying the genes of non-model organisms. RNA-seq data is also useful for annotating genes within genomes, providing concrete evidence of expressed genes.

In **Chapter III** I characterized gene expression patterns associated with introgression at candidate male barrier loci in developing hybrid larvae from the Långholmen population. To do this, I sequenced paired-end reads from the RNA extracted from whole-body larvae. These were hybrid larvae representing both the W and R lineages found at Långholmen. From this data, the transcriptomes of both lineages were constructed using the *de novo* transcriptome assembler: Trinity ([Grabherr et al. 2011](#)). To obtain read counts, I mapped each sample's paired-end reads to their respective transcriptomes (i.e. W samples to the W transcriptome) using Bowtie2 ([Langmead & Salzberg 2012](#)), and read counts were estimated using RSEM ([Li & Dewey 2011](#)). Differential gene expression analysis was performed using DESeq2 ([Love et al. 2014](#)), to test for patterns of expression associated with the amount of introgression at candidate male barrier loci (see section 3.3 marker loci and their utility). I also examined differential expression associated with the actual size of the larvae, as a proxy for their age.

### 3.7 Genomic assembly using both long and short reads

The development of a modified SMRTbell library construction protocol, used to build a reference genome from a single *Anopheles* mosquito ([Kinghan et al. 2019](#)), has recently opened the door for constructing high-quality genomes from single individuals of small organisms. Using only a standard DNA extraction protocol of a single hybrid male, and illumina sequence reads from two hybrid gynes, we could create a cost-effective genomic reference for our study system. After constructing multiple assemblies, the initial assembly was constructed using the long-read assembler wtdbg2 ([Ruan & Li 2020](#)) from reads greater than 10kb in length. The completeness of the assembly was assessed using BUSCO ([Seppey et](#)

[al. 2019](#)) with the Hymenoptera ODB gene set v10. This initial assembly was selected based on its comparable length (280 Mb) to other recently assembled *Formica* ant genomes, those of *F. exsecta* (278Mb, [Dhaygude et al. 2019](#)) and *F. selysi* (290Mb, [Brelsford et al. 2020](#)). The assembly was then polished of sequencing errors using Racoon ([Vaser et al. 2017](#)), using the PacBio and short read Illumina reads. After contaminant and mitochondrial removal, TE annotation and anchoring of contigs to the *F. selysi* genome (see **Chapter IV** for details), the genome was annotated using both RNAseq reads (**Chapter III**) and protein hints.

## 4 MAIN RESULTS AND THEIR INTERPRETATION

My thesis adds fundamental knowledge towards hybridization between *F. aquilonia* and *F. polycтена* in southern Finland, with the broader aim of developing foundational resources and knowledge for the future study of this emerging model system. I showed that, although hybrids and parental *F. aquilonia* exhibit mate choice behaviour in experimental conditions (**Chapter II**), this is not a complete barrier to gene flow, and is likely to facilitate hybridization. Evidence of this is displayed by the existing widespread hybridization between *F. aquilonia* and *F. polycтена* across Southern Finland (**Chapter I**). I found that eggs of hybrid queens hatch 83% less compared to parental *F. aquilonia*, though both lay the same number of eggs (**Chapter II**) which indicates the costs of hybridization. I also found that introgressed marker loci, which exhibit signatures of selection, are associated with gene expression differences in developing hybrid larvae (**Chapter III**). **Chapters III & IV** also establish annotated transcriptomic and genomic resources for the future study of admixture and hybridization in this emerging model system. Below I discuss these results in the broader context of studying hybrid populations in nature, and their significance for the evolution of these species going forwards.

### 4.1 Hybridization and its potential role in the future evolution of populations of wood ants

#### **Hybridization is widespread with no systematic distribution across southern Finland**

In **Chapter I**, I surveyed populations of *F. aquilonia* and *F. polycтена* wood ants to examine the extent of hybridization throughout southern Finland, using

established mitochondrial (COI) and microsatellite markers. We previously knew of one hybrid population in Långholmen, but we were interested in finding out just how common hybridization is between these parental species. I showed that there were signatures of introgression and admixture (i.e. hybridization) of varying degrees in more than half of all the populations surveyed. The range of admixture at the genomic level, and the lack of many individuals with approximately 50% ancestry of both species, suggested that there has been a long history of hybridization between these species, and that it is likely ongoing. Additionally, the geographical distribution of hybrid and parental species populations lacked any systematic pattern, i.e. there was no geographical gradient/cline between parental lineages with hybrids in between (Figure 7), thus we characterized the pattern as a hybrid swarm.

This work adds to the growing body of literature describing widespread hybridization between diverging lineages. Hybridization among eusocial species in particular is widely documented (reviewed in [Feldhaar et al. 2008](#)), including hybridization between *F. polycтена* and *F. rufa*, another closely related species, in central Europe ([Seifert et al. 2010](#)). The outcomes of hybridization in ants can be unique, for example, the mind boggling phenomena of hybridogenesis occurs as the result of heterospecific matings between genetically distinct lineages in *Pogonomyrmex harvester* ants ([Sirviö et al. 2011](#)), *Solenopsis* fire ants ([Helms-Cahan & Vinson 2003](#)), and desert ants *Cataglyphis hispanica* ([Leniaud L et al. 2012](#)). Hybridization in ants can result in unidirectional introgression, for example the mitochondrial haplotype of two hybridizing *Serviformica* species, *F. selysi* and *F. cinerea*, is predominantly that of *F. cinerea* ([Purcell et al. 2016](#)).

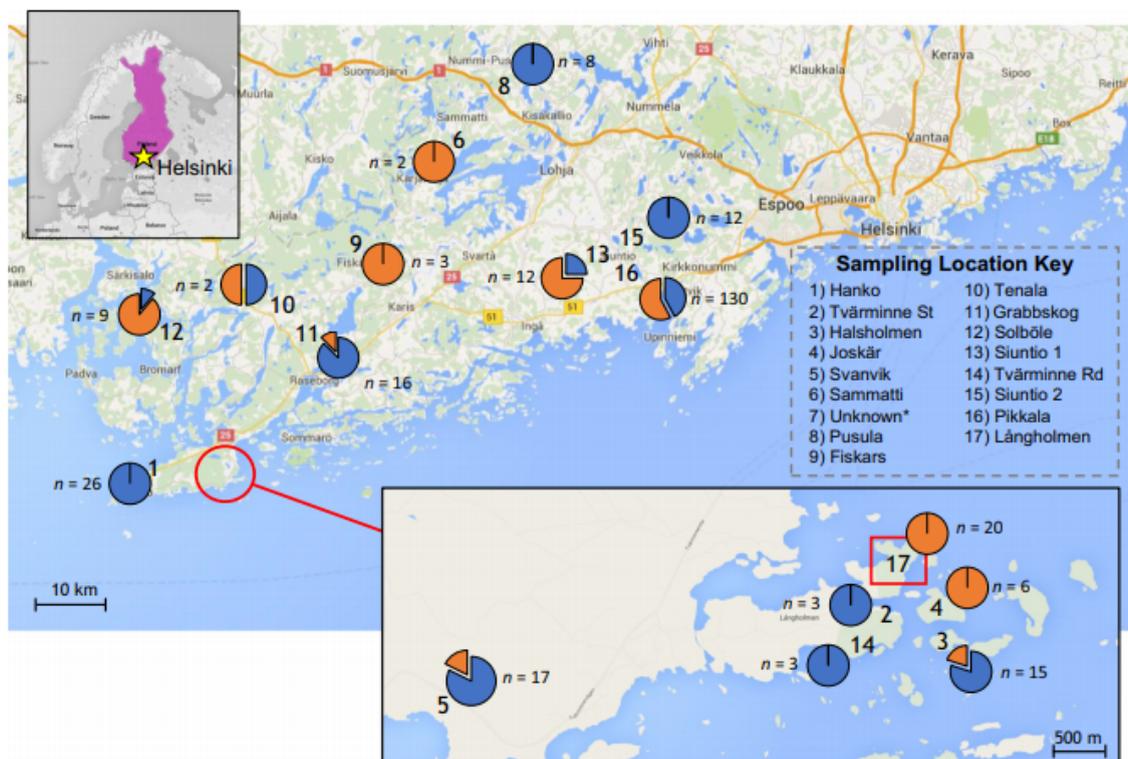
I also discovered a diversity of mitochondrial haplotypes from both parental species (**Chapter 1**), indicating that there have been several independent hybrid origins and that gene flow is bidirectional between these species (i.e. that multiple independent queens of both lineages have hybridized). Indeed, 6 populations possessed the mitochondrial haplotypes of both parental species. There was also a striking association between mitochondrial haplotypes and genomic marker loci (referred to as cytonuclear mismatch in **Chapter I**) (Figure 8), indicating that independent hybridization events have resulted in consistent combinations of genomic and mitochondrial DNA. Cytonuclear mismatch could be an indication that there is adaptive cytonuclear mismatch in these species. This pattern is likely the result of selection and or mate-choice with backcrossing to parental lineages.

Interestingly, cytonuclear mismatch has been cited as the cause for hybrid breakdown in other systems (reviewed by [Burton et al. 2013](#)), though the systematic pattern we observed does not, at least on the surface, support that prediction. Cytonuclear mismatch could be driven by female preference for

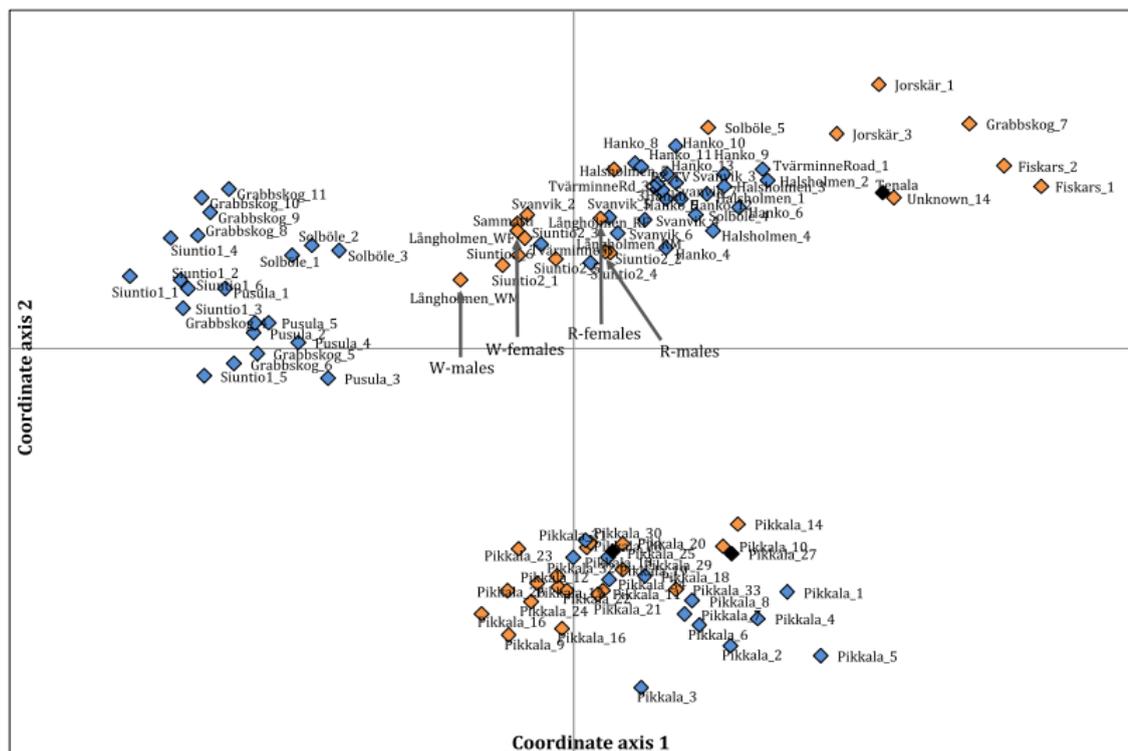
heterospecific mates, which would result in biased nuclear introgression. While heterospecific mating is commonplace in eusocial insects ([Feldhaar et al. 2008](#)), there is little to no evidence of heterospecific mating preferences in ants, save for those species intertwined by hybridogenesis. Preference for dissimilar mates could be a response to inbreeding, and inbreeding problems are known in ants ([Vitikainen et al. 2011](#)).

The last key result of this chapter was the discovery of another hybrid population that likely experiences different selection in males and females. We could see this because females of this population possessed alleles at a frequency higher than 10%, but that weren't observed in males, and was similar to the pattern observed by other studies in the Långholmen population ([Kulmuni et al. 2010](#), [Kulmuni & Pamilo 2014](#)). This also supports the idea that the independent hybrid populations suffer from hybrid breakdown, particularly of males, in line with predictions from Haldane's rule ([1922](#)). In spite of this, hybridization is widespread between the parental species.

By describing both hybrid and parental species populations, this survey contributes foundational knowledge towards the *Formica* hybrid system, which can be used to compare individuals of different genotypes from different populations (e.g. **Chapter II**). We also know that both *F. aquilonia* and *F. polyctena* are distributed throughout southern Finland ([Sorvari 2021](#)), though as yet it is unknown to what extent hybridization is prevalent throughout the rest of their sympatric ranges, or with other wood ant species in Finland.



**Figure 7** Sampling locations across southern Finland, to the west of Helsinki. Locations are labelled with a number (1-17) which are named in the key. Location 17 (in the red box) is the long-term hybrid population, Långholmen, while localities 1-16 were newly sampled that year. The pie-charts show the proportion of mitochondrial haplotypes sampled (the n indicates the number of samples) in the locations belonging to each parental lineage (blue = *F. aquilonia*, orange = *F. polycтена*). \*The location marked as “Unknown” in the key lacked GPS coordinates, but possessed *F. polycтена* mitotypes. There was no consistent pattern of distribution between *F. aquilonia*, *F. polycтена* or their hybrids (**Chapter I**, Figure 1).



**Figure 8** Principal coordinate analysis of the nuclear microsatellite markers from 96 nests. The 1st axis explains 18.6% of the variation and the 2nd axis 15.6%. The data points are coloured according to the mitochondrial haplotype that was sampled from those nests (blue = *F. aquilonia*, orange = *F. polycтена*, black = a mix of both parental haplotypes). The previously studied hybrid groups (W & R) are indicated with black arrows. The 1st axis divides the samples according to parental gene pools, where (putative) *F. aquilonia* is situated on the left, putative *F. polycтена* is on the right and hybrids are between. Hybrids show a clear association between the nuclear and mitochondrial markers that is the reverse of the parental populations (i.e. cytonuclear mismatch) (**Chapter I**, Figure 3).

### **Mate choice in *Formica* may promote hybrid persistence in nature**

In **Chapter I** I observed a consistent pattern of cytonuclear mismatch, which could be driven by either mate-choice and/or selection. For example, this pattern could be driven by female preference for heterospecific males, which would result in preferential introgression of nuclear, but not the maternally inherited mitochondrial DNA. Hybridization is also clearly widespread between these species in southern Finland (**Chapter I**), raising natural questions about

barriers to gene flow in these species. i) Can queens discriminate between males of different species, and does mate choice exist? ii) Do females prefer heterospecific mates? iii) And do hybrids prefer their own type (to their own detriment)?

We found that hybrid and *F. aquilonia* females showed clear asymmetric patterns for mate choice (Figure 9)(**Chapter II**). Hybrid queens were equally likely to mate with males from hybrid or *F. aquilonia* nests, but discriminated against *F. polycтена*, while *F. aquilonia* females preferred males of their own species. *F. aquilonia* females were also less willing to copulate than hybrid queens. Thus, females can discriminate between males of different types, and they do have mating preferences. Neither hybrid nor *F. aquilonia* females examined here exhibited preferences for heterospecific males, though hybrids did not discriminate against their own type. From this it's hard to conclude whether mate choice is the cause for the cytonuclear pattern we previously observed (**Chapter I**), though it seems less likely. We could not test *F. polycтена* females here so their preferences remain an open question. However, based on demographic modelling work, there is asymmetric gene flow into the *F. polycтена* gene pool from *F. aquilonia* in Europe, and the effective population size of *F. polycтена* is much smaller ([Portinha et al. 2021](#)). We might then expect *F. polycтена* to prefer outbreeding and heterospecific mates.

Given the above results I wanted to ask the question, could mate choice promote the persistence of these wood ant hybrids in nature? Mate choice may restrict or promote gene flow between parental species, or between them and their hybrids ([Schmidt & Pfenning 2016](#)). Furthermore, hybrid populations may not persist should they become swamped by gene flow from one parental species or the other ([Todesco et al. 2016](#)). Equally, should hybrid populations experience such strong selection against deleterious combinations of alleles that they cannot persist in nature, they will also perish. Thus, both are important to study, in order to predict the persistence of hybrid populations in nature.

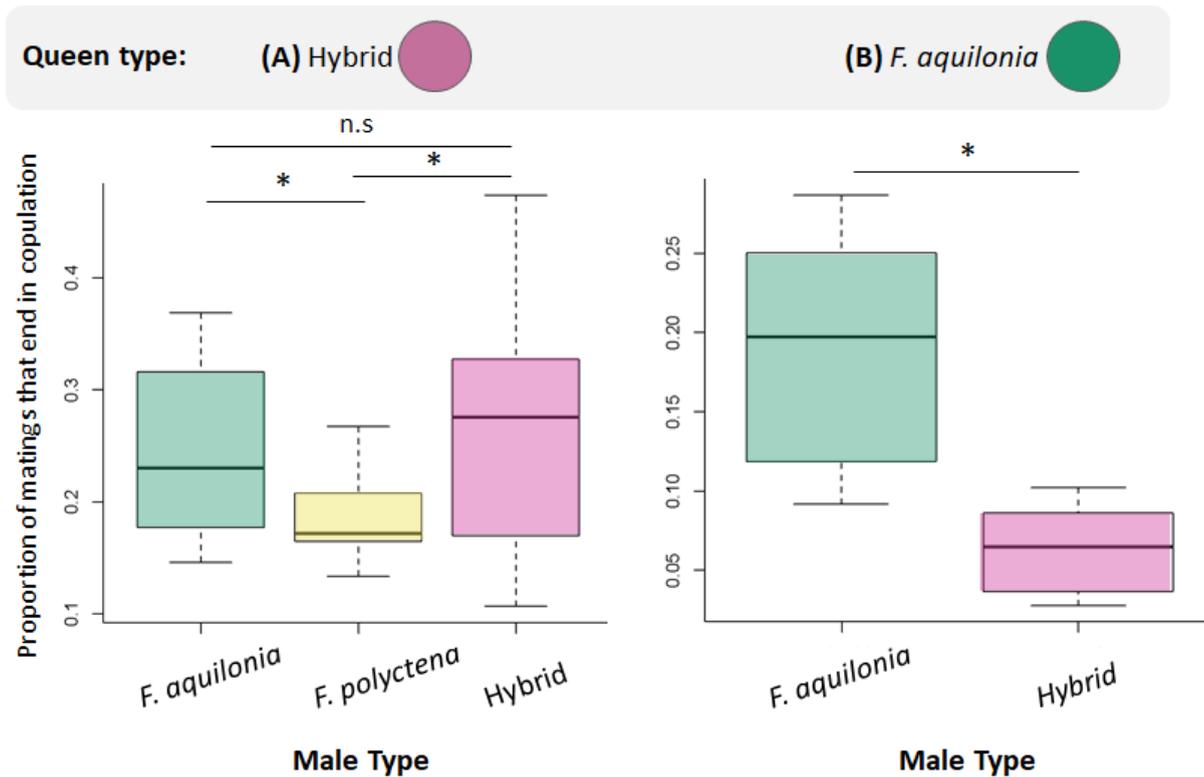
Based on my results from **Chapters I & II**, I expect that these hybrid populations are likely to persist in nature. Hybrids do exhibit mate choice, but crucially, they do not discriminate against their own type, as evidenced by both the experimental mate choice and realized mating results. We also found evidence of little to no gene flow into the hybrid or the parental *F. aquilonia* population (though we did find loci that were differentiated between males and females of both populations, see **Chapter II**). This result is not surprising, as previous studies of related *F. rufa* species ants have shown that they are poor dispersers ([Vitikainen & Sundström 2015](#)), and that large genetic differences between populations can be observed even over short geographical distances ([Mäki-Petäys et al. 2005](#)). Indeed in one of our previous studies, two hybrid

populations situated within 2 km of each other (Bunkkeri and Långholmen) have private alleles (**Chapter I**). Hence we do not expect hybrid populations to become swamped by parental genotypes.

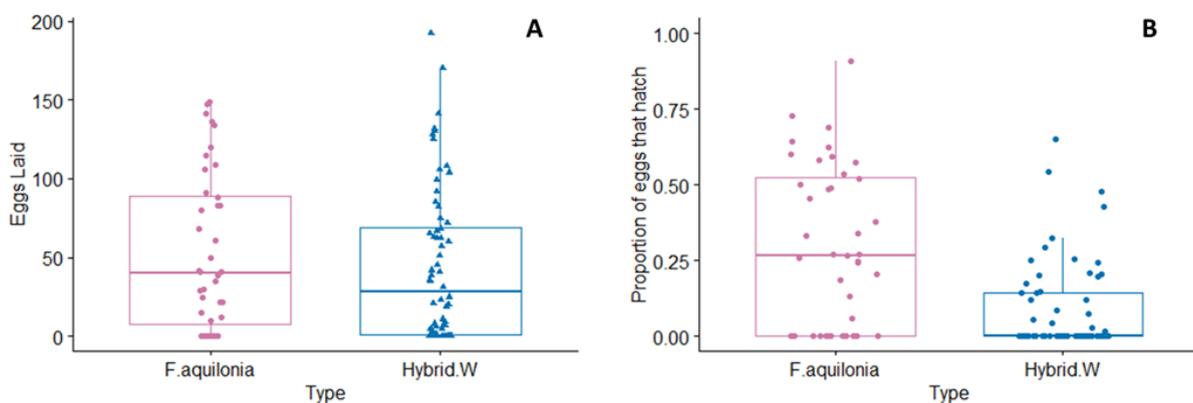
The second factor to consider for hybrid persistence is selection against deleterious combinations of alleles. In **Chapter II** I also documented a dramatic reduction in hatching success (-83%) for hybrid queen sexual brood, compared to that of parental *F. aquilonia*, under lab conditions, though their egg laying rates were the same (Figure 10). Our results indicate that eggs of hybrid queens suffered from viability selection and hybrid breakdown, or from post-mating prezygotic isolation mechanisms that prevent the formation of viable zygotes. This result is also in line with significant changes in allele frequencies between larval and adult stages in nature indicating natural selection acting on these hybrids ([Kulmuni & Pamilo 2014](#); [Kulmuni et. al. 2020](#)), and the allele frequency differences found in one of the newly documented hybrid populations (**Chapter I**).

Since the persistence of hybrid populations is dependent upon the survival of reproducing individuals, this reduction of survival in sexual brood may initially seem like a fitness burden. Yet, there are several other factors that may reduce the deleterious consequences of reduced egg hatching success. Queens from the focal species can live for 5 or more years ([Keller & Genoud 1997](#)) and each nest may contain hundreds of egg laying queens. Thus, even if only a fraction of hybrid eggs survive to maturity, there may still be enough to replace the queen caste and sustain the population. Furthermore, a large proportion of the developing individuals may not make it to maturity even in the parental populations due to cannibalism ([Schultner et al. 2014](#)) or worker policing ([Helanterä & Sundström 2007](#)), which may not be enforced in hybrid populations. In addition, we know that the Långholmen hybrid population has existed for decades ([Pamilo 1993](#)), a testament to the longevity of this population. Hence, the strong signal of viability selection may not be the death knell of hybrid populations, since it appears to have not reduced the likelihood of their persistence in nature.

Altogether, considering both the mate choice and selection documented in this thesis, I consider that it is likely that hybrid populations of *Formica* should persist in nature, meaning that they could act as reservoirs for genetic diversity and gene flow.



**Figure 9 Successful mating attempt data, measured using behavioural assays of virgin queens and males. A.** The proportion of male mating attempts that result in copulations with hybrid queens. Mating attempts by *F. polyctena* were less likely to lead to successful copulations compared to mating attempts by hybrid or *F. aquilonia* males (between which there was no difference). **B.** The proportion of successful copulations with *F. aquilonia* queens. Males had a significantly higher proportion of successful copulations with their own type. We could not test *F. polyctena* males with *F. aquilonia* females due to sampling issues. Note also, that the y-axis are different between the two graphs, and overall *F. aquilonia* queens were generally less likely to copulate, even with their own type. (Figure 4, Chapter II).



**Figure 10 Egg laying and hatching data from old (>1 year) field collected queens. A.** The egg laying rates did not differ between hybrid and *F. aquilonia* queens. **B.** The proportion of eggs that hatched from hybrid queens was 83% less than from *F. aquilonia* queens. (Figure 7, Chapter II).

**Hybrids may act as vehicles for gene flow in *Formica* species**

If hybrid populations persist, they could play a role in mediating gene flow between the parental populations, but this gene flow could be biased to one direction or the other. Results from **Chapter II** show that hybrid females discriminate against *F. polyctena* males, but not against *F. aquilonia* males. Furthermore, female willingness matters for successful copulations (see Figure 5, **Chapter II**). Thus, based on our results, female gene flow from hybrids is more likely to go towards *F. aquilonia* than *F. polyctena*. *F. aquilonia* females did discriminate against the hybrid males, but this barrier was incomplete, as evidenced by copulations between *F. aquilonia* females and hybrid males during the mate choice experiment. We have also previously documented both *F. aquilonia* and *F. polyctena* mitochondrial haplotypes are widespread in natural hybrid populations, even within the same populations (**Chapter I**), indicating that *F. aquilonia* females do hybridize in nature.

We did not examine *F. polyctena* female choice in this experiment due to sample limitations. However, based on demographic modeling *F. polyctena* has a smaller long-term effective population size compared to *F. aquilonia* ([Portinha et al. 2021](#)) and is the minority species in Finland ([Sorvari 2021](#)). Based on these observations we would predict *F. polyctena* females to be less choosy than *F. aquilonia*, they could also lack a preference altogether, or even prefer to out-breed. This would predict gene flow entering *F. polyctena*, either via hybrid or *F. aquilonia* males. Indeed, demographic modeling suggests the two parental species diverged with unidirectional gene flow from *F. aquilonia* into *F. polyctena*. Interestingly though, models based on Finnish populations also showed that there was more bidirectional gene flow ([Portinha et al. 2021](#)), suggesting that either Finnish populations, or the Finnish environment promote hybridization.

With our current expectations of gene-flow outlined above, we predict that gene flow should occur in both directions, but to a greater extent into *F. polyctena* rather than *F. aquilonia*. We also expect that hybridization and hybrid populations could be increasingly common given the incomplete premating barriers to gene flow, thus promoting their role as vehicles for gene flow between both parental species in nature.

## 4.2 Developing resources for the study of admixture and selection in wood ants

### **Genes underlying selection in hybrids**

Previous studies have established clear evidence of selection acting on introgressed alleles at marker loci during development in this wood ant hybrid system ([Kulmuni & Pamilo 2014](#), [Kulmuni et al. 2020](#), [Martin-Roy et al. 2021](#)). Based on work in other systems, divergent gene expression of hybrids can reveal

the loci under selection in hybrids ([Landry et al. 2007](#), [Mack & Nachman 2017](#)). In **Chapter III** my collaborators and I successfully assembled and annotated two *de novo* transcriptomes from both hybrid lineages (W and R) present in Långholmen (Table 1). I also characterized SNPs at 105 out of 637 candidate male barrier loci from the RNAseq data in all samples. I used the transcriptomes and SNP information to study the genes expression patterns associated with introgression at candidate male barrier loci. My aim was to capture genes that potentially underlie selection. However, contrary to my expectations, the results revealed that only a small proportion of the transcriptomes were differentially expressed in association with introgression at candidate male barrier loci (Figure 11). Furthermore, of the genes that were differentially expressed, few overlapped with potential candidate genes that were close to with loci under selection, and which could putatively underlie fitness differences between males and females.

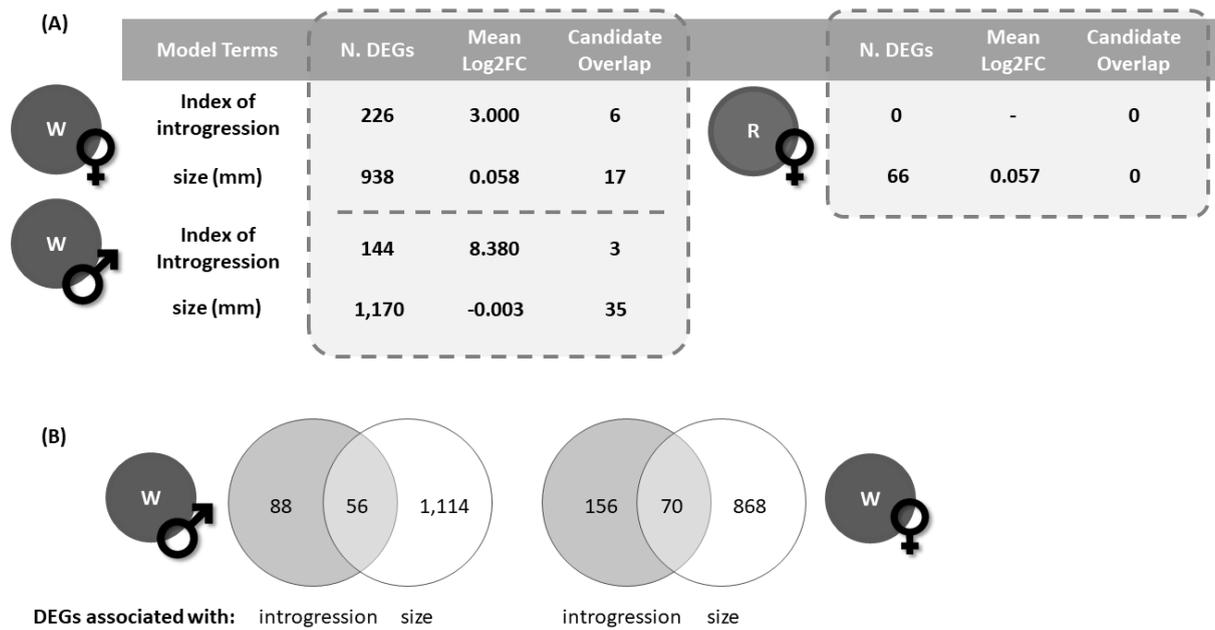
Incorporating data on survival with hybrid expression is likely to reveal the genetic basis of BDMIs ([Landry et al. 2007](#)), as I attempted to do with the introgression at loci known to be associated with survival and fitness during development in these hybrids. In the year we sampled, we know selection favoured introgressed alleles in W males ([Kulmuni et al. 2020](#)). As such these differentially expressed genes could underlie hybrid fitness in these developing ants. Other studies of hybridization have typically focused on comparisons between hybrids and parents to look for divergent expression patterns, but these studies often lack fitness information ([Alvarez et al. 2015](#)).

The moderate numbers of differentially expressed genes is perhaps surprising considering that hybrids are supposedly highly admixed individuals. Yet they are late-generation hybrids, having existed for nearly three decades ([Pamilo 1993](#)), and as such many of the genes or loci that cause large differences in expression (a pattern associated with BDMIs, e.g. [Tirosh et al. 2009](#)) may have already been removed by sorting and selection in these hybrids. Interestingly, hybrids between fire ant species, *Solenopsis invicta* x *S. richteri* revealed minimal “mis-expression” (n = 32 genes) compared to parental lineages (i.e. expression outside the range of either parental lineage) ([Ometto et al. 2011](#)). Ometto *et al.* suggested this was caused by expression being dominated by alleles from one of the lineages. If the alleles of one of our hybrid ant parental lineages (*F. aquilonia* or *F. polycтена*) is dominant, this might suggest why we see so little differential expression associated with introgression. Importantly, the signatures of hybrid breakdown are thought to be revealed by mis-expression, caused by deleterious epistasis ([Landry et al. 2007](#), [Ortiz-Barrientos et al. 2007](#)). However, this only accounts for diploid females, and does not account for the lack of differential expression that I found in males, where we would expect recessive incompatibilities to be revealed on their haploid genomes.

I also expected males to have larger differences in gene expression compared to females because they experience stronger selection ([Kulmuni & Pamilo 2014](#), [Kulmuni et al. 2020](#), [Martin-Roy et al. 2021](#)), yet this was not the case. It is possible that only a small number of genes can underlie severe dysfunction and even death in hybrids ([Brideau & Barbash 2011](#)), particularly if such genes are highly connected or pleiotropic ([Joeng et al. 2001](#), [Taylor et al. 2019](#), [Josephs et al. 2017](#)), or this may depend upon the networks in which they reside (reviewed by [Satokangas et al. 2020](#)). However, perhaps the lesson here comes by comparing these results to the genomic loci under selection. Both the candidate male barrier loci and the microsatellites are scattered throughout the genome of these hybrids, and at the time of writing, we lack a clear signal of large-effect loci, which could also reflect the small numbers of differentially expressed genes that I found. As such, the candidate loci may underlie small-effect yet genome-wide loci that are collectively experiencing selection, perhaps as part of the same gene networks ([Palmer & Feldman 2009](#)). Such networks could underlie a single phenotypic trait, or functional pathway. For example, a mutation study (i.e. where they generated mutates) in the yeast species *Saccharomyces cerevisiae* found that pathways enriched for functional mutations resulted in negative epistasis and lethality ([Constanzo et al. 2016](#)).

With the above in mind, that there are only a handful of differentially expressed candidate genes is perhaps not surprising: currently the candidate male barrier loci have yet to be directly tied to the function of any genes. Alternatively, these genes may be differentially expressed in another life stage. We think for example, that heterozygous hybrid females are selected for throughout their lifetimes ([Kulmuni & Pamilo 2014](#)). We also have yet to characterize precisely when during development that selection acts upon these hybrids, as previous studies only compared individuals at early (eggs) and late (adults) developmental stages ([Kulmuni & Pamilo 2014](#), [Kulmuni et al. 2020](#), [Martin-Roy et al. 2021](#)). A hint about this comes from **Chapter II**, as a significantly smaller proportion of eggs hatch from hybrid queens compared to parental *F. aquilonia*, and may suggest that selection occurs even before they have hatched. Thus we may not be capturing selection in the larvae studied here.

In summary I did find gene expression patterns associated with introgression at candidate male barrier loci, though only relatively small numbers of genes from individuals in the W lineage. These genes may nevertheless be proven to play important roles in the fitness consequences of developing hybrids. Furthermore this study contributes valuable transcriptomic and RNAseq data towards this system, which have already been used to annotate a genome of this system.



**Figure 11 Differential Expression Results.** **A)** Differentially Expressed Genes (DEGs) associated with index of introgression at candidate male barrier loci and size from models independently fitted for W males, W females, and R females, their mean log Fold Change per unit change in the independent variable (e.g. in the Index of Introgression) and the overlap between differentially expressed unigenes and genes annotated near candidate male barrier loci. **B)** Overlap between DEGs found associated with the index of introgression at candidate male barrier loci and size (mm). (Figure 3, **Chapter III**).

Genetic group	W transcriptome	R transcriptome
No PE-high quality reads	580.9 M	290.2 M
<b>Number of unigenes</b>	<b>16,481</b>	<b>14,897</b>
GC (%)	40.1	39.82
N50	3,031	3,864
Median Contig Length	1,091	1,532
Average Contig Length	1,786.84	2,324.95
Total Assembled Bases	29,448,851	34,634,723
No unigenes annotated in NR	10,617	11,088
No unigenes annotated in Swissprot	7,725	7,533
No unigenes annotated in GO	7,625	8,318

**Table 1** Transcriptome assembly statistics.

### ***Cost-effective genomic assembly from a single hybrid male***

The final, polished, reference genome was 276Mb with an overall BUSCO score for complete single-copy orthologs was 98.5%. Almost 92% of the ~5,098 contigs mapped to Arthropoda, while a total of 82 contigs mapped to Proteobacteria, 6 of which were found to be horizontal gene transfers (HGTs) into the hybrid genome. This is unsurprising as ants are known to harbour *Wolbachia* endosymbionts, which has previously been shown to have undergone HGTs into the genomes in *F. exsecta* ([Dhaygude et al. 2019](#)). A total of 17,426 protein-coding genes were annotated based on RNAseq reads (generated from samples in **Chapter III**), as well as protein hints. These protein-coding genes had a BUSCO completeness score of 97.4%. Other statistics, including N50 numbers, can be seen in **Chapter IV**. Finally, 78.2% of the contigs, accounting for 213Mb, could be anchored to 27 of the chromosomes assembled in the *F. selysi* genome ([Brelsford et al. 2020](#)), which we might assume to be contiguous chromosomes within our own hybrids, hence this is a “pseudo” chromosome-level genome assembly. Based on these metrics this reference genome is of comparable quality to other insect genomes assembled from single individuals. These results demonstrate the cost-effectiveness of modern sequencing technology, and provide reference material for future studies of admixture in this emerging model system.

## **5 CONCLUSIONS**

I have demonstrated that hybridization is widespread and ongoing between *F. aquilonia* and *F. polycтена* in southern Finland (**Chapter I**), and that, in spite of deleterious consequences to hybridizing (**Chapter II**), mate choice against heterospecific matings has not evolved to completely stop gene flow between the parental species (**Chapter II**). Hence I expect hybridization to continue, and for hybrid populations themselves to persist in nature. These populations may act as reservoirs of genetic diversity, and may facilitate gene flow between parental lineages.

I described gene expression patterns associated with introgression at candidate loci (**Chapter III**). This study aimed to take the first steps towards describing the molecular mechanisms and genomic architecture underlying selection in these hybrids. Towards this end my thesis also provides transcriptome (**Chapter III**) and genome references (**Chapter IV**) for the future study of hybridization and selection in this emerging and fascinating system.

## 6 ACKNOWLEDGEMENTS

First I must thank **Jonna**. I'll always remember hearing about your fascinating hybrid ants during my undergraduate in Sheffield, nothing had captured my interest so completely. What started as a summer project looking at mitochondrial DNA sequences eventually developed into my PhD and the projects contained here in this thesis. You introduced me to the fascinating world of ants, speciation and hybridization genetics - and without you this thesis would not have happened. You've always known what to do, and what the next step is. You have been dedicated and focused, and you taught me to never give up. You always strive to improve and you are a constant source of inspiration. I am so grateful that you took a chance on me, and believed in me. Thank you for everything.

To the rest of the “emerging” **Speciant**, your support and dedication as a group has been more than I could have asked for. Old and new members, I thank you all; **Ina, Pierre, Patrick, Beatriz, Raphael, Elisa** and the rest. I have loved our book groups, journal clubs, science retreats, lunches during the pandemic, and the odd behavioural experiment in Tvärminne. Our meetings were a source of normality during strange times. I'm proud and grateful to have been a part of it. You all helped me get through the pandemic, and the end of my PhD.

Thanks to my collaborators, **Claire, Rose** and **Pierre**, without whom this thesis would not have been possible. **Claire**, without you I might still have been lost in the transcriptomics - you helped push me towards the end, and helped me to focus on a goal. **Rose**, your advice and kind support have been invaluable, and you have a knack for steering me to the right course. **Pierre**, thank you for your insight into all things speciation, but also for your patience, and for the hilarious banter. Also thanks to **Elisa** for all of your reliable lab work and the fantastic company. Thanks too to **Bendick** for his hard work on the behavioural data. And thanks to **Heikki** for showing me how to hunt for *Formica* queens.

Thanks to **Lotta** and **Roger**, my invaluable supervisors. **Lotta**, you helped to get me started in my PhD, and funded me as I got started. Throughout my PhD **Jonna** and I have come to you with questions, and you have been there to guide us. Thank you. And thanks to **Roger**, I don't know how many times you have helped me to see the wood for the trees. Several times I came to you with, what I thought was a complex problem. I'd plan out an hour meeting, and you'd sort me out in 15 minutes.

To the other members of **Team Antzz**, now turned **Team ESB**; to the many people that remain, and those before me. In particular **Perttu, Dalial, Ed, Katja, Siiri, Dimitri, Jana, Stafva, Kishor, Franzi, Purabi** and **Nick** (who could never find the rum). Thank you for all of the wonderful company over the years. Together we made a playground for science, where anyone can ask questions and learn. It is your collective input that helped to shape and prepare me for interviews, and helped me hone my presentations. I've loved being a part of group discussions, talking about each other's science and asking questions. I feel privileged to have been a part of it for so long.

**Unni** and **Sanja**, the nail-painting-lamentation gang, warrant special thanks. We put the world to rights (sort of) in our morning coffee sessions. Watching you both prepare for your defenses gave me hope that I could too. **Unni**, you were the big sister I never had. I miss working with you every day. **Sanja**, you were our beacon of honesty and reality. You were never afraid to tell us what you thought. I will never forget our science slam poetry. Life as a PhD hasn't felt the same since you both left.

To my "people" who loved bubbles, **Jo**, **Michelle** and **Andrew**, I would not have made it through without you. Your company, love and support are more than I could have asked for. Thank you for the wonderful evenings, crazy gifts that made me laugh, the sad moments and the joyous chaos that we are together. Thanks of course also to the moon.

Thanks to **Antti** for being a great friend. You taught me all the best Finnish phrases, and always had a plan to get up to mischief. Though maybe "plan" is too strong a word. I had no idea that one awkward meeting in Prisma (not twice but thrice in the same evening!), would turn into such an awesome friendship.

**Sedeer**, thank you for being a good friend, and an ear for my worries. Thanks for the boxes, board games, fantastic food and for checking on me when I needed a friend.

**Ana** and **Tad**, I miss you both. I learnt a lot about myself from spending time with you. Thank you for deep conversations, for the games and the laughter. You made life joyous.

To **Ville** and **Noora**, thank you for including me in your antics. Your friendships were invaluable to me over the years. Thanks for the foxes, festivals, and the good times. Thanks too to the **KF folk** for always making me feel welcome over the years.

Thanks to the many other people over the years who have helped me and been fantastic company during my PhD. **Miia**, **Tiuski**, **Marion**, **Aina**, **Miquel**, **Emma**, **Carly**, **Henry**, **Eirik**, **Pauliina**, **Nuno**, **Jessica** to name a few - there are too many to thank, and I never had time for you all. Thanks too to the organising committee of Spring Symposium 2018 who made it possible, and making me believe that I could organise something big. Thanks to **Aapo** and **Marjukka** for the capoeira and friendship.

Thanks to all the staff on the Viikki campus, PhD boards, Tvarminne staff and MES lab staff. You have all been the glue that makes research possible.

For my family - this thesis is dedicated to you all. Together you are my rock, and my home to come back to. Even from another country, I know I can always depend on you. Thank you for the unwavering love and support. Thank you **mum**, for calling and thinking about me every day. Thank you **dad**, for being the kind and gentle person that you are. Thank you **Chloe**, for our shared bond. Thank you **Nango**, for your enthusiasm and love. In loving memory of my **grandfather**, whom I miss every day.

Thank you **Jenny**, my best friend. Thank you for all the love, the mad moments and all the rest. You helped me to believe in myself, and gave me confidence. You have been dedicated and supportive, and I am proud to call you my partner.

Countless incredible people have helped me throughout my PhD. I appreciated every friend and helping hand along the way. My thesis has been funded by a variety of sources over the years, including initial funding thanks to Lotta from the Centre of Excellence in Biological Interactions, the Pro Fauna et Flora Fennica foundation, and finally the LUOVA PhD school.

I thank **Philine Fulner** and **Jürgen Gadau** for their quick and efficient pre-examination of my thesis, and **Dieter Ebert** for being my willing opponent.

### *An ode to my ants*

What do ants do,  
When no one's around?  
Do you think that they frolic,  
Carefree on a mound?

In the heart of the woods,  
At the dawn of the spring,  
Little ants wake and stretch  
And queens don their bling.

The queens come to the surface,  
Snatch moments of sun,  
They take a deep breath,  
Now there's work to be done.

In the bowels of their city,  
They lay a new clutch,  
Among them is Dory,  
She's not special as such.

This is one story,  
Of an ant called Dory,  
Right now, she's an egg,  
But she's from the top peg.

She'll be a queen,  
Though she'll not rule supreme,  
She'll have no fanfare, no,

She's a baby making machine,

Once Dory is fully grown,  
She's got to find a mate,  
But ah, there's so many!  
Now there's no time for a date...

She's found a good guy,  
Whom she quite fancies,  
But oh no it's all wrong,  
He's from the wrong species!

Oh well never mind,  
Now she'll have to make do,  
She'll find a cozy nest,  
And start laying eggs, look, one... two...

Now, here's my moment,  
As David Attenborough,  
Leering at nests,  
I'm a nerdy scientist, Bro!

I snatch up the queens,  
Cart them back to my lair,  
It's time to lay eggs,  
No need for despair...

I wait for the brood,

Counting eggs, look! One... two...

But don't hold your breath,  
Or else you'll turn blue

I wait till they hatch,  
And who should emerge?  
But cute little larvae,  
Only... their DNA is merged!

Their parents were,  
From two different species,  
They're hybrid ants,  
But wait a minute, what even are species?  
*No time for that now!*

From what we can tell,  
Each ant has a distinct smell,  
Which tells a picky queen everything she  
needs to know,  
So these ants are like connoisseurs of...  
snow?

An eau de toilette,  
With all the information,  
Certainly handy when,  
There's no time for explanations,

But what should it matter,  
Who mates with whom?  
It's the 21st century, mum,  
Let's go, get a room!  
In these uncertain times,  
It's time for some change!

It's just that when you get hybridization,  
Things might go wrong...  
Dory... I have some bad news,  
Some of your offspring, they won't live so  
long,

Some of your daughters,  
Their DNA just doesn't work,  
And the ones who survive,  
They might have some quirks,

So, what do ants do,  
When no one's around?  
Do you think that they frolic,  
Carefree on a mound?

I think I might just know....

## 7 REFERENCES

- Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., ... Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26(2), 229–246. <https://doi.org/10.1111/j.1420-9101.2012.02599.x>
- Alvarez, M., Schrey, A. W., & Richards, C. L. (2015). Ten years of transcriptomics in wild populations: What have we learned about their ecology and evolution? *Molecular Ecology*, 24(4), 710–725. <https://doi.org/10.1111/mec.13055>
- Barton, N. H., & Hewitt, G. M. (1985). Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics*, 16(1), 113–148. <https://doi.org/10.1146/annurev.es.16.110185.000553>
- Bazykin, A. D. (1969). Hypothetical Mechanism of Speciation. *Evolution*, 23(4), 685–687. JSTOR. <https://doi.org/10.2307/2406862>
- Bendall, E., Mattingly, K., Moehring, A., & Linnen, C. (2020). Lack of intrinsic postzygotic isolation in haplodiploid male hybrids despite high genetic distance. <https://doi.org/10.1101/2020.01.08.898957>

- Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2013). GenBank. *Nucleic Acids Research*, 41(Database issue), D36–42. <https://doi.org/10.1093/nar/gks1195>
- Beresford, J., Elias, M., Pluckrose, L., Sundström, L., Butlin, R. K., Pamilo, P., & Kulmuni, J. (2017). Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown. *Molecular Ecology*, 26(15), 4013–4026. <https://doi.org/10.1111/mec.14183>
- Boomsma, J. J., Brady, S. G., Dunn, R. R., Gadau, J., Heinze, J., Keller, L., Moreau, C. S., Sanders, N. J., Schrader, L., Schultz, T. R., Sundstrom, L., Ward, P. S., Wcislo, W. T., Zhang, G., & GAGA Consortium. (2017). The Global Ant Genomics Alliance (GAGA). *Myrmecological News*, 25, 61–66.
- Brelsford, A., Purcell, J., Avril, A., Tran Van, P., Zhang, J., Brütsch, T., Sundström, L., Helanterä, H., & Chapuisat, M. (2020). An Ancient and Eroded Social Supergene Is Widespread across *Formica* Ants. *Current Biology*, 30(2), 304–311.e4. <https://doi.org/10.1016/j.cub.2019.11.032>
- Brideau, N. J., & Barbash, D. A. (2011). Functional conservation of the *Drosophila* hybrid incompatibility gene *Lhr*. *BMC Evolutionary Biology*, 11, 57. <https://doi.org/10.1186/1471-2148-11-57>
- Burton, R. S., Pereira, R. J., & Barreto, F. S. (2013). Cytonuclear genomic interactions and hybrid breakdown. *Annual Review of Ecology, Evolution, and Systematics*, 44, 281–302.
- Cahan, S. H., & Vinson, S. B. (2003). Reproductive Division of Labor between Hybrid and Nonhybrid Offspring in a Fire Ant Hybrid Zone. *Evolution*, 57(7), 1562–1570. JSTOR.
- Chenoweth, S. F., & Blows, M. W. (2006). Dissecting the complex genetic basis of mate choice. *Nature Reviews Genetics*, 7(9), 681–692. <https://doi.org/10.1038/nrg1924>
- Clark, M. E., O'Hara, F. P., Chawla, A., & Werren, J. H. (2010). Behavioral and spermatogenic hybrid male breakdown in *Nasonia*. *Heredity*, 104(3), 289–301. <https://doi.org/10.1038/hdy.2009.152>
- Colyott, K., Odu, C., & Gleason, J. M. (2016). Dissection of signalling modalities and courtship timing reveals a novel signal in *Drosophila saltans* courtship. *Animal Behaviour*, 120, 93–101. <https://doi.org/10.1016/j.anbehav.2016.07.015>
- Costanzo, M., VanderSluis, B., Koch, E. N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., & Lee, S. D. (2016). A global genetic interaction network maps a wiring diagram of cellular function. *Science*, 353(6306).
- Coughlan, J. M., & Matute, D. R. (2020). The importance of intrinsic postzygotic barriers throughout the speciation process. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1806), 20190533. <https://doi.org/10.1098/rstb.2019.0533>
- Coyne, J., & Orr, H. (2004). *Speciation*. (Sinauer Associates, 2004).
- Dasmahapatra, K. K., Walters, J. R., Briscoe, A. D., Davey, J. W., Whibley, A., Nadeau, N. J., Zimin, A. V., Hughes, D. S. T., Ferguson, L. C., Martin, S. H., Salazar, C., Lewis, J. J., Adler, S., Ahn, S.-J., Baker, D. A., Baxter, S. W., Chamberlain, N. L., Chauhan, R., Counterman, B. A., ... The *Heliconius* Genome Consortium. (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, 487(7405), 94–98. <https://doi.org/10.1038/nature11041>
- Dhaygude, K., Nair, A., Johansson, H., Wurm, Y., & Sundström, L. (2019). The first draft genomes of the ant *Formica exsecta*, and its *Wolbachia* endosymbiont reveal extensive gene transfer from endosymbiont to host. *BMC Genomics*, 20(1), 301. <https://doi.org/10.1186/s12864-019-5665-6>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Feldhaar, H., Foitzik, S., & Heinze, J. (2008). Lifelong commitment to the wrong partner: Hybridization in ants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1505), 2891–2899. <https://doi.org/10.1098/rstb.2008.0022>

- Fisher, R. A. (1922). On the Interpretation of  $\chi^2$  from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society*, 85(1), 87–94. JSTOR.  
<https://doi.org/10.2307/2340521>
- Fontaine, M. C., Pease, J. B., Steele, A., Waterhouse, R. M., Neafsey, D. E., Sharakhov, I. V., Jiang, X., Hall, A. B., Catteruccia, F., Kakani, E., Mitchell, S. N., Wu, Y.-C., Smith, H. A., Love, R. R., Lawniczak, M. K., Slotman, M. A., Emrich, S. J., Hahn, M. W., & Besansky, N. J. (2015). Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science*, 347(6217). <https://doi.org/10.1126/science.1258524>
- Fortelius, W. (2005). Mating behaviour in the polygynous/polydomous wood ant *Formica aquilonia*. *Annales Zoologici Fennici*, 42(3), 213–224. JSTOR.
- Ghenu, A.-H., Blanckaert, A., Butlin, R. K., Kulmuni, J., & Bank, C. (2018). Conflict between heterozygote advantage and hybrid incompatibility in haplodiploids (and sex chromosomes). *Molecular Ecology*, 27(19), 3935–3949. <https://doi.org/10.1111/mec.14482>
- Goropashnaya, A. V., Fedorov, V. B., Seifert, B., & Pamilo, P. (2012). Phylogenetic relationships of Palaearctic *Formica* species (Hymenoptera, Formicidae) based on mitochondrial cytochrome B sequences. *PloS One*, 7(7), e41697.
- Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5(1), 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., ... Regev, A. (2011). Trinity: Reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, 29(7), 644–652. <https://doi.org/10.1038/nbt.1883>
- Grant, P. R., & Grant, B. R. (2019). Hybridization increases population variation during adaptive radiation. *Proceedings of the National Academy of Sciences*, 116(46), 23216. <https://doi.org/10.1073/pnas.1913534116>
- Greenspan, R. J., & Ferveur, J.-F. (2000). Courtship in *Drosophila*. *Annual Review of Genetics*, 34(1), 205–232. <https://doi.org/10.1146/annurev.genet.34.1.205>
- Griebel, J., Gießler, S., Poxleitner, M., Faria, A. N., Yin, M., & Wolinska, J. (2015). Extreme environments facilitate hybrid superiority—The story of a successful *Daphnia galeata* × *longispina* hybrid clone. *PloS One*, 10(10), e0140275.
- Haldane, J. B. S. (1922). *Sex ratio and unisexual sterility in hybrid animals*. <https://doi.org/10.1007/bf02983075>
- Harr, B. (2006). Genomic islands of differentiation between house mouse subspecies. *Genome Research*, 16(6), 730–737.
- Helanterä, H., & Sundström, L. (2007). Worker policing and nest mate recognition in the ant *Formica fusca*. *Behavioral Ecology and Sociobiology*, 61(8), 1143–1149. <https://doi.org/10.1007/s00265-006-0327-5>
- Huerta-Sánchez, E., Jin, X., Asan, Bianba, Z., Peter, B. M., Vinckenbosch, N., Liang, Y., Yi, X., He, M., Somel, M., Ni, P., Wang, B., Ou, X., Huasang, Luosang, J., Cuo, Z. X. P., Li, K., Gao, G., Yin, Y., ... Nielsen, R. (2014). Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*, 512(7513), 194–197. <https://doi.org/10.1038/nature13408>
- Jeong, H., Mason, S. P., Barabási, A.-L., & Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature*, 411(6833), 41–42.
- Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., ... Kingsley, D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, 484(7392), 55–61. <https://doi.org/10.1038/nature10944>
- Josephs, E. B., Wright, S. I., Stinchcombe, J. R., & Schoen, D. J. (2017). The relationship between selection, network connectivity, and regulatory variation within a population of *Capsella grandiflora*. *Genome Biology and Evolution*, 9(4), 1099–1109.
- Keller, L., & Genoud, M. (1997). Extraordinary lifespans in ants: A test of evolutionary theories of ageing. *Nature*, 389(6654), 958–960. <https://doi.org/10.1038/40130>

- Kingan, S. B., Heaton, H., Cudini, J., Lambert, C. C., Baybayan, P., Galvin, B. D., Durbin, R., Korlach, J., & Lawniczak, M. K. N. (2019). A High-Quality De novo Genome Assembly from a Single Mosquito Using PacBio Sequencing. *Genes*, 10(1), 62. PubMed. <https://doi.org/10.3390/genes10010062>
- Koevoets, T., & Beukeboom, L. (2009). Koevoets T, Beukeboom LW. Genetics of postzygotic isolation and Haldane's Rule in haplodiploids. *Heredity* 102: 16-23. *Heredity*, 102, 16–23. <https://doi.org/10.1038/hdy.2008.44>
- Kulmuni, J., Nouhaud, P., Pluckrose, L., Satokangas, I., Dhaygude, K., & Butlin, R. K. (2020). Instability of natural selection at candidate barrier loci underlying speciation in wood ants. *Molecular Ecology*, 29(20), 3988–3999. <https://doi.org/10.1111/mec.15606>
- Kulmuni, J., & Pamilo, P. (2014). Introgression in hybrid ants is favored in females but selected against in males. *Proceedings of the National Academy of Sciences*, 111(35), 12805–12810. <https://doi.org/10.1073/pnas.1323045111>
- Kulmuni, J., Seifert, B., & Pamilo, P. (2010). Segregation distortion causes large-scale differences between male and female genomes in hybrid ants. *Proceedings of the National Academy of Sciences*, 107(16), 7371–7376. <https://doi.org/10.1073/pnas.0912409107>
- Landry, C. R., Lemos, B., Rifkin, S. A., Dickinson, W. J., & Hartl, D. L. (2007). Genetic properties influencing the evolvability of gene expression. *Science (New York, N.Y.)*, 317(5834), 118–121. <https://doi.org/10.1126/science.1140247>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Latour, Y., Perriat-Sanguinet, M., Caminade, P., Boursot, P., Smadja, C. M., & Ganem, G. (2014). Sexual selection against natural hybrids may contribute to reinforcement in a house mouse hybrid zone. *Proceedings of the Royal Society B: Biological Sciences*, 281(1776), 20132733. <https://doi.org/10.1098/rspb.2013.2733>
- Lemmon, E. M., & Lemmon, A. R. (2010). Reinforcement in chorus frogs: Lifetime fitness estimates including intrinsic natural selection and sexual selection against hybrids. *Evolution*, 64(6), 1748–1761. <https://doi.org/10.1111/j.1558-5646.2010.00955.x>
- Leniaud, L., Darras, H., Boulay, R., & Aron, S. (2012). Social hybridogenesis in the clonal ant *Cataglyphis hispanica*. *Current Biology*, 22(13), 1188–1193.
- Li, B., & Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12(1), 323. <https://doi.org/10.1186/1471-2105-12-323>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Mack, K. L., Campbell, P., & Nachman, M. W. (2016). Gene regulation and speciation in house mice. *Genome Research*, 26(4), 451–461. <https://doi.org/10.1101/gr.195743.115>
- Mack, K. L., & Nachman, M. W. (2017). Gene Regulation and Speciation. *Trends in Genetics: TIG*, 33(1), 68–80. <https://doi.org/10.1016/j.tig.2016.11.003>
- Mäki-Petäys, H., Zakharov, A., Viljakainen, L., Corander, J., & Pamilo, P. (2005). Genetic changes associated to declining populations of *Formica* ants in fragmented forest landscape. *Molecular Ecology*, 14(3), 733–742. <https://doi.org/10.1111/j.1365-294X.2005.02444.x>
- Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, 20, 229–237. <https://doi.org/10.1016/j.tree.2005.02.010>
- Mallet, J. (2007). Hybrid speciation. *Nature*, 446(7133), 279–283. <https://doi.org/10.1038/nature05706>
- Mallet, J., Besansky, N., & Hahn, M. W. (2016). How reticulated are species? *BioEssays*, 38(2), 140–149. <https://doi.org/10.1002/bies.201500149>

- Martin-Roy, R., Nygård, E., Nouhaud, P., & Kulmuni, J. (2021). Differences in thermal tolerance between parental species could fuel thermal adaptation in hybrid wood ants. *The American Naturalist*. <https://doi.org/10.1086/715012>
- Matute, D. R. (2010). Reinforcement Can Overcome Gene Flow during Speciation in *Drosophila*. *Current Biology*, 20(24), 2229–2233. <https://doi.org/10.1016/j.cub.2010.11.036>
- Mayr, E. (2013). Animal Species and Evolution. In *Animal Species and Evolution*. Harvard University Press.
- Nadeau, N. J., Whibley, A., Jones, R. T., Davey, J. W., Dasmahapatra, K. K., Baxter, S. W., Quail, M. A., Joron, M., French-Constant, R. H., Blaxter, M. L., Mallet, J., & Jiggins, C. D. (2012). Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1587), 343–353. <https://doi.org/10.1098/rstb.2011.0198>
- Noor, M. A. F. (1997). Genetics of sexual isolation and courtship in male hybrids of *Drosophila pseudoobscura* and *Drosophila persimilis*. *Evolution; International Journal of Organic Evolution*, 51(3), 809–815. <https://doi.org/10.1111/j.1558-5646.1997.tb03663.x>
- Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, 18(3), 375–402. <https://doi.org/10.1111/j.1365-294X.2008.03946.x>
- Nouhaud, P., Beresford, J., & Kulmuni, J. (2021). Cost-effective long-read assembly of a hybrid *Formica aquilonia* × *Formica polyctena* wood ant genome from a single haploid individual. *BioRxiv*, 2021.03.09.434597. <https://doi.org/10.1101/2021.03.09.434597>
- Ometto, L., Ross, K. G., Shoemaker, D., & Keller, L. (2012). Disruption of gene expression in hybrids of the fire ants *Solenopsis invicta* and *Solenopsis richteri*. *Molecular Ecology*, 21(10), 2488–2501.
- Orr, H. A. (1995). The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, 139(4), 1805–1813.
- Orr, H. A. (1996). Dobzhansky, Bateson, and the genetics of speciation. *Genetics*, 144(4), 1331–1335.
- Orr, H. A., & Turelli, M. (2001). The Evolution of Postzygotic Isolation: Accumulating Dobzhansky-Muller Incompatibilities. *Evolution*, 55(6), 1085–1094. <https://doi.org/10.1111/j.0014-3820.2001.tb00628.x>
- Ortiz-Barrientos, D., Counterman, B. A., & Noor, M. A. (2007). Gene expression divergence and the origin of hybrid dysfunctions. *Genetica*, 129(1), 71–81.
- Palmer, C. A., & Edmands, S. (2000). Mate choice in the face of both inbreeding and outbreeding depression in the intertidal copepod *Tigriopus californicus*. *Marine Biology*, 136(4), 693–698. <https://doi.org/10.1007/s002270050729>
- Palmer, M. E., & Feldman, M. W. (2009). Dynamics of hybrid incompatibility in gene networks in a constant environment. *Evolution: International Journal of Organic Evolution*, 63(2), 418–431.
- Pamilo, P. (1993). Polyandry and allele frequency differences between the sexes in the ant *Formica aquilonia*. *Heredity*, 70(5), 472–480. <https://doi.org/10.1038/hdy.1993.69>
- Pavey, S., Collin, H., Nosil, P., & Rogers, S. (2010). The role of gene expression in ecological speciation. *Annals of the New York Academy of Sciences*, 1206, 110–129. <https://doi.org/10.1111/j.1749-6632.2010.05765.x>
- Pfennig, K. S. (2007). Facultative Mate Choice Drives Adaptive Hybridization. *Science*, 318(5852), 965. <https://doi.org/10.1126/science.1146035>
- Portinha, B., Avril, A., Bernasconi, C., Helanterä, H., Monaghan, J., Seifert, B., Sousa, V. C., Kulmuni, J., & Nouhaud, P. (2021). Whole-genome analysis of multiple wood ant population pairs supports similar speciation histories, but different degrees of gene flow, across their European range. *BioRxiv*, 2021.03.10.434741. <https://doi.org/10.1101/2021.03.10.434741>

- Presgraves, D. C. (2003). A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. *Genetics*, 163(3), 955–972.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155(2), 945–959.  
<https://doi.org/10.1093/genetics/155.2.945>
- Purcell, J., Zahnd, S., Athanasiades, A., Türlér, R., Chapuisat, M., & Brelsford, A. (2016). Ants exhibit asymmetric hybridization in a mosaic hybrid zone. *Molecular Ecology*, 25(19), 4866–4874. <https://doi.org/10.1111/mec.13799>
- Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, 27(1), 83–109.  
<https://doi.org/10.1146/annurev.ecolsys.27.1.83>
- Rieseberg, L. H., Van Fossen, C., & Desrochers, A. M. (1995). Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature*, 375(6529), 313–316.  
<https://doi.org/10.1038/375313a0>
- Ritchie, M. G. (2007). Sexual selection and speciation. *Annu. Rev. Ecol. Evol. Syst.*, 38, 79–102.
- Rosenthal, G. G. (2013). Individual mating decisions and hybridization. *Journal of Evolutionary Biology*, 26(2), 252–255. <https://doi.org/10.1111/jeb.12004>
- Ross, K. G., & Shoemaker, D. D. (2005). Species delimitation in native South American fire ants. *Molecular Ecology*, 14(11), 3419–3438.
- Ruan, J., & Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. *Nature Methods*, 17(2), 155–158. <https://doi.org/10.1038/s41592-019-0669-3>
- Runemark, A., Trier, C. N., Eroukhmanoff, F., Hermansen, J. S., Matschiner, M., Ravinet, M., Elgvin, T. O., & Sætre, G.-P. (2018). Variation and constraints in hybrid genome formation. *Nature Ecology & Evolution*, 2(3), 549–556.  
<https://doi.org/10.1038/s41559-017-0437-7>
- Russell, S. T., & Magurran, A. E. (2006). Intrinsic reproductive isolation between Trinidadian populations of the guppy, *Poecilia reticulata*. *Journal of Evolutionary Biology*, 19(4), 1294–1303. <https://doi.org/10.1111/j.1420-9101.2005.01069.x>
- Satokangas, I., Martin, S. H., Helanterä, H., Saramäki, J., & Kulmuni, J. (2020). Multi-locus interactions and the build-up of reproductive isolation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1806), 20190543.  
<https://doi.org/10.1098/rstb.2019.0543>
- Schmidt, E. M., & Pfennig, K. S. (2016). Hybrid female mate choice as a species isolating mechanism: Environment matters. *Journal of Evolutionary Biology*, 29(4), 865–869.  
<https://doi.org/10.1111/jeb.12818>
- Schultner, E., Gardner, A., Karhunen, M., & Helanterä, H. (2014). Ant larvae as players in social conflict: Relatedness and individual identity mediate cannibalism intensity. *The American Naturalist*, 184(6), E161–E174.
- Schumer, M., Powell, D. L., Delclós, P. J., Squire, M., Cui, R., Andolfatto, P., & Rosenthal, G. G. (2017). Assortative mating and persistent reproductive isolation in hybrids. *Proceedings of the National Academy of Sciences*, 114(41), 10936–10941.  
<https://doi.org/10.1073/pnas.1711238114>
- Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., Peichel, C. L., Saetre, G.-P., Bank, C., Brännström, Å., Brelsford, A., Clarkson, C. S., Eroukhmanoff, F., Feder, J. L., Fischer, M. C., Foote, A. D., Franchini, P., Jiggins, C. D., Jones, F. C., ... Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, 15(3), 176–192. <https://doi.org/10.1038/nrg3644>
- Seifert, B., Kulmuni, J., & Pamilo, P. (2010). Independent hybrid populations of *Formica polyctena* Xrufa wood ants (Hymenoptera: Formicidae) abound under conditions of forest fragmentation. *Evolutionary Ecology*, 24(5), 1219–1237.  
<https://doi.org/10.1007/s10682-010-9371-8>
- Seppy, M., Manni, M., & Zdobnov, E. M. (2019). BUSCO: assessing genome assembly and annotation completeness. In *Gene prediction* (pp. 227–245). Springer.

- Servedio, M. R., & Noor, M. A. F. (2003). The Role of Reinforcement in Speciation: Theory and Data. *Annual Review of Ecology, Evolution, and Systematics*, 34(1), 339–364. <https://doi.org/10.1146/annurev.ecolsys.34.011802.132412>
- Sirviö, A., Pamilo, P., Johnson, R. A., Jr, R. E. P., & Gadau, J. (2011). Origin and Evolution of the Dependent Lineages in the Genetic Caste Determination System of Pogonomyrmex Ants. *Evolution*, 65(3), 869–884. <https://doi.org/10.1111/j.1558-5646.2010.01170.x>
- Sorvari, J. (2021). Distribution of Finnish mound-building Formica ants (Hymenoptera: Formicidae) based on using a citizen science approach. *European Journal of Entomology*, 118, 57–62.
- Stockan, J. A., Robinson, E. J. H., Trager, J. C., Yao, I., & Seifert, B. (2016). *Wood Ant Ecology and Conservation* (E. J. H. Robinson & J. A. Stockan, Eds.). Cambridge University Press; Cambridge Core. <https://doi.org/10.1017/CBO9781107261402.002>
- Svedin, N., Wiley, C., Veen, T., Gustafsson, L., & Qvarnström, A. (2008). Natural and sexual selection against hybrid flycatchers. *Proceedings of the Royal Society B: Biological Sciences*, 275(1635), 735–744. <https://doi.org/10.1098/rspb.2007.0967>
- Taylor, M. A., Wilczek, A. M., Roe, J. L., Welch, S. M., Runcie, D. E., Cooper, M. D., & Schmitt, J. (2019). Large-effect flowering time mutations reveal conditionally adaptive paths through fitness landscapes in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 116(36), 17890–17899.
- Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology & Evolution*, 3(2), 170–177. <https://doi.org/10.1038/s41559-018-0777-y>
- Tirosh, I., Reikhav, S., Levy, A., & Barkai, N. (2009). A Yeast Hybrid Provides Insight into the Evolution of Gene Expression Regulation. *Science (New York, N.Y.)*, 324, 659–662. <https://doi.org/10.1126/science.1169766>
- Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G., & Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications*, 9(7), 892–908. <https://doi.org/10.1111/eva.12367>
- Turner, T. L., Hahn, M. W., & Nuzhdin, S. V. (2005). Genomic islands of speciation in Anopheles gambiae. *PLoS Biol*, 3(9), e285.
- Vaser, R., Sović, I., Nagarajan, N., & Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Research*, 27(5), 737–746.
- Vitikainen, E., Haag-Liautard, C., & Sundström, L. (2011). Inbreeding and reproductive investment in the ant Formica exsecta. *Evolution*, 65(7), 2026–2037. <https://doi.org/10.1111/j.1558-5646.2011.01273.x>
- Vitikainen, E. I. K., Haag-Liautard, C., & Sundström, L. (2015). Natal Dispersal, Mating Patterns, and Inbreeding in the Ant Formica exsecta. *The American Naturalist*, 186(6), 716–727. <https://doi.org/10.1086/683799>
- Wagner, D. L., Grames, E. M., Forister, M. L., Berenbaum, M. R., & Stopak, D. (2021). Insect decline in the Anthropocene: Death by a thousand cuts. *Proceedings of the National Academy of Sciences*, 118(2). <https://doi.org/10.1073/pnas.2023989118>
- Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1), 57–63. <https://doi.org/10.1038/nrg2484>
- Wittkopp, P. J., & Kalay, G. (2011). Cis-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews. Genetics*, 13(1), 59–69. <https://doi.org/10.1038/nrg3095>