Sexual Conflict and Gene Expression in *Drosophila melanogaster*

PAOLO INNOCENTI
Dissertation presented at Uppsala University to be publicly examined in Zootissalen, EBC, Villavägen 9, Uppsala. Saturday, October 1, 2011 at 09:30 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Sexual conflict is broadly defined as a conflict between the evolutionary interests of the two sexes. Depending on the genetic architecture of the traits involved, it can occur at the level of male-female interactions or take the form of selection acting to change the mean of a shared trait against the sign of its genetic correlation. The aim of my thesis was to use genome-wide expression profiles in the model organism Drosophila melanogaster to provide novel insights in the study of sexual conflict.

First, we studied the female post-mating response to partition transcriptional changes associated with reproduction from male-induced effects, which are known to be harmful to females. We found substantial changes in expression of metabolic pathways associated with the activation of reproduction, while male-specific effects were dominated by the onset of an immune response. Changes in female response under different mating strategies was studied using experimental evolution: we found that monogamous females suffered decreased fecundity and their gene expression profiles suggested an overall weaker response to mating. To identify sexually antagonistic genes, we used hemiclonal lines and associated their sex-specific fitness with genome-wide transcript abundance. We confirmed the presence of a negative covariance for fitness and identified a group of candidate genes experiencing sexually antagonistic selection. We then focused on mitochondria, which can enable the accumulation of deleterious mutations with sex-specific effects due to their maternal inheritance, and found few effects on nuclear gene expression in females but major effects in males, predominantly in male-specific tissues. Finally, we used published data to compare intraspecific and interspecific genetic variation for a set of transcripts, to test whether speciation occurs along lines of maximum genetic variance.

In conclusion, gene expression techniques can generate useful results in the study of sexual conflict, particularly in association with phenotypic data or when integrated with published datasets.

Keywords: Sexual conflict, sexual selection, male-female coevolution, gene expression, transcriptome, microarrays, sexual dimorphism, Drosophila


© Paolo Innocenti 2011

ISSN 1651-6214 0346-5462
ISBN 978-91-554-8130-8
urn:nbn:se:uu:diva-156567 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-156567)
List of Papers

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


The following papers were written during the course of my doctoral studies but are not part of the present dissertation:


Morrow E.H. and Innocenti P. Female postmating immune responses, immune system evolution and immunogenic males. *Submitted manuscript.*


Cover illustration by Lauren Redman (♀ fly) and Paolo Innocenti (♂ fly).
# Contents

## Introduction
- My Thesis ................................................. 7
- Sexual conflict ........................................... 7
- Gene expression .......................................... 9
- Aims ....................................................... 10

## Material and Methods
- Study species: *Drosophila melanogaster* ................. 11
- Experimental designs ...................................... 12
- Microarrays ............................................... 14
- Statistical analysis ....................................... 15

## Results and discussion
- Female post-mating response .............................. 16
- Sexually dimorphic and antagonistic genes ............... 19
- Sex-specific selective sieve on mtDNA evolution ........ 21
- Evolvability and sexual constraints in gene networks ... 22

## Conclusions
- .............................................................. 26

## Sammanfattning på svenska
- .............................................................. 28

## Riassunto in italiano
- .............................................................. 33

## Acknowledgements
- .............................................................. 38
Introduction

My Thesis

Gene expression data can provide novel insights in the study of sexual conflict.

Sexual conflict

Males and females are defined by the size of the gametes they produce. Males produce small and mobile gametes (spermatozoa) and females produce large and non-mobile gametes (ova). This unbalanced primary parental investment leads to different strategies in which the two sexes can maximize their fitness (see Kokko et al., 2006). As a result, males and females may differ in their optima for morphological (e.g. body size), physiological (e.g. metabolic rate) and behavioural traits (e.g. mating frequency, parental care). Whenever this difference exists, there is potential for sexual conflict, broadly defined as “a conflict between the evolutionary interests of the two sexes” (Parker, 1979).

Consider a simple trait, coded by the same single autosomal locus in both sexes, and assume that selection acts, on this trait, in opposite directions in males and females (i.e. sexually antagonistic selection). If two alleles exist conferring a higher or a lower value to the phenotype, one of them will be positively selected in one sex, and the other will be positively selected in the other sex. This simple scenario is an example of intralocus sexual conflict (Rice and Chippindale, 2001). In reality, most traits show a more complex genetic architecture, and in a quantitative genetic framework it can be defined quoting Lande (1979), as occurring when selection acts to change the mean of a trait in the two sexes against the sign of its genetic correlation.

This concept is not new in evolutionary biology. Kottler (1980) examines the correspondence between Darwin and Wallace on the topic of conspicuous and ‘protective’ coloration of animal species. They discuss the evolution of sexual dimorphism (usually conspicuous males opposed to ‘protective’ females) in terms of sexual vs. natural selection. Wallace’s thoughts are summarized by the following passage:
The combined results of mutual [sexual] selection and equal inheritance of the sexually selected color variations would be equally conspicuous sexes. But whenever one sex was endangered more than the other by conspicuous coloration, natural selection prevented sexual selection from acting on the sex in greater danger, and it converted the equal inheritance of the variations sexually selected in the more conspicuous sex into sex-limited inheritance, so the sex in greater danger did not acquire, by inheritance, conspicuous coloration.

This passage contains all the ‘ingredients’ of intralocus sexual conflict: an initial situation of perfect intersexual genetic correlation and standing genetic variation, selection acting antagonistically in the two sexes, and a decrease in between-sex covariance to allow the evolution of sexual dimorphism. This framework would be analytically developed and extended more than a hundred years later (Lande, 1980). The challenge today is to determine how important this process is in extant populations; if, how and how fast this conflict can be resolved through, for example, the evolution of sexual dimorphism, and which traits tend to be antagonistically selected or highly constrained (Bonduriansky and Chenoweth, 2009).

A second consequence that arises from the very existence of sexual reproduction concerns male-female interactions, in particular in the virtually ubiquitous case of absence of reciprocal lifelong monogamy (when the correlation between the fitness of a male and his mate is less than 1). Let’s assume that a given trait appears in an individual of one sex (e.g. a male) and gives that individual an advantage in male-female interactions, by increasing the “efficiency” of that trait or any other trait mediated by it, for example by reducing his mate’s propensity/ability to mate with other males. This trait will likely increase the individual male’s fitness – even if it causes a decrease in female fitness – if the male does not have to ‘share’ the paternity of the offspring with other males (Parker, 1979). From the female perspective, however, this means a fitness loss (a net fitness loss if the indirect genetic benefits to the male offspring are not greater than the direct costs), and females will be selected to evolve counteradaptations to minimize the cost imposed by the male trait, in a cycle of adaptation-counteradaptations that has been described as an “arms race”, or sexually antagonistic coevolution (see Arnqvist and Rowe, 2005, for a thorough review), while the process, as a whole, has been termed interlocus sexual conflict.
The fruit fly *Drosophila melanogaster* provides an example of this type of conflict: multiple mating is known to be costly for females, as it reduces their lifespan (Kuijper *et al.*, 2006), and some component of the male ejaculate, in particular, are known to manipulate her reproductive behaviour, inducing a refractory period and possibly a suboptimal egg-laying activity (Avila *et al.*, 2011). Although the recent activity in this field of study has led to fast advancement in knowledge of specific molecular interactions, it is not clear what processes are involved in determining this fitness cost of mating.

In sum, the different ecological roles played by males and females, the different selective pressures to which they are subject, and intrinsic differences in maternal and paternal contribution to the offspring create the potential for an intricate network of interactions, many of which remain to be understood.

**Gene expression**

Despite the extensive phenotypic differences between males and females, they are genetically nearly identical. In most species, the genomes of the two sexes differ by only a few genes located on sex-specific chromosomes (Ellegren and Parsch, 2007). What produces the sexual dimorphism that we observe must therefore be a different use that the sexes make of the genetic material, or, in other words, the differential expression of the respective genomes (Rinn and Snyder, 2005).

Gene expression represents the lowest ‘phenotypic’ level where a large amount sexual differentiation (and arguably genetic variance) exists. I argue that these characteristics make gene expression a particularly useful trait to study sexual conflict, provided we can adequately link gene sequence with phenotypic traits (genotype-phenotype map). The main obstacle to this goal is represented by the manifold complexity of the epistatic interactions and the pleiotropic constraints of the transcriptome.

Over the last decade, the increased technical and logistical availability of genomic techniques, together with decrease in costs, has made it possible to obtain gene expression data on a scale useful for ecological and even quantitative genetics experiments (e.g. Ayroles *et al.*, 2009). So far, the study of sexual conflict has largely been confined to the analysis of phenotypic traits, with morphological, behavioural and life-history traits being the focus of the laboratory (Chippindale *et al.*, 2001) as
Sexual conflict and gene expression in *D. melanogaster*

well as natural (Foerster *et al.*, 2007) populations. High throughput genomic data might provide useful results, adding another dimension and contributing to the understanding of this complex discipline.

**Aims**

The aim of my thesis is two-fold. First, to probe the potential of transcriptome analyses in studying intra- and inter-locus sexual conflict. Second, to use these genomic tools, and combining them with more ‘classical’ experimental approaches, to contribute novel results in the study of male-female coevolution and sex-specific adaptations and constraints, and possibly to generate novel hypotheses.

Given the breadth of the field of study, I focused on a small and heterogeneous subset of topics: *small* because of time and resource limitations, and *heterogeneous* in an attempt to

i) provide a representative sample of the variety of the research in this discipline, ii) demonstrate the flexibility of gene expression analysis in a range of experimental designs, and iii) follow promising leads generating interesting hypotheses. In particular, I focused on:

- Female post-mating response (Paper I) and its evolution under different sexual selection regimes (Paper II).
- Sexual dimorphism and sex-specific selective pressures on the expression of the fruit fly genome (Paper III).
- Sex-specific effects mtDNA variation in their interaction with the nuclear genome (Paper IV).
- Evolvability and sexual constraints in interspecific divergence of gene networks along lines of genetic variance (Paper V).
Materials and methods

Study species: *Drosophila melanogaster*

*Drosophila melanogaster* (the fruit fly or vinegar fly) is an insect of the order Diptera, and a famous member of the large *Drosophila* genus. It was the first model organism to be used for genetic analysis (Morgan, 1910), and today it is arguably one of the most studied and best known animal species in the laboratory.

For several reasons, *D. melanogaster* is well suited to be the study species for the object of my thesis. First of all, as already mentioned, the genetic architecture has been studied in detail, and the genome has been fully sequenced (Adams *et al.*, 2000). The vast majority of the genes have been identified and functionally characterized and these information are available in curated public repositories, and lots of mutants, recombinant lines and other genetic tools (e.g. RNAi) are available. Second, the full genome sequence of 12 species within the *Drosophila* group is also available, and can be used (and has been used, see Paper V) to compare traits among species and potentially extend some results across species boundaries. Third, although surprisingly little information is available about the ecology of the species in the wild, laboratory adapted or field-collected populations are suitable and have been extensively used for population genetics, quantitative genetics and experimental evolution studies (Begun *et al.*, 2000; Rice *et al.*, 2005; Mackay, 2010; Chenoweth and McGuigan, 2010; Edward *et al.*, 2010).

The fly genome contains ≈14000 genes. The nucleus carries 4 pairs of chromosomes: sexual chromosomes, 2 autosomes and a ‘dot’ fourth chromosome, which contains only a handful of genes. The sexual chromosome system is X/Y, and the sex is determined by the ratio of Xs to autosomes (Ashburner *et al.*, 2004).

Male and females are sexually dimorphic: females are larger than males, show a different pigmentation pattern (females typically have a striped abdomen, while the distal half of male abdomen is completely pigmented) and males carry sex combs (modified bristles) on the tarsal segment of the legs. Other less apparent dimorphic traits include the
Sexual conflict and gene expression in *D. melanogaster*
cuticular hydrocarbons (CHCs), a set of sexually selected complex compounds which play an important role in determining individual reproductive success.

Mating in this species occurs close to oviposition sites, where males wait for females and court them following a complex but stereotypical ritual (Spieth, 1974). Mating lasts on average \( \approx 15 \) minutes, after which females start laying eggs and show decreased mating propensity. Both males and females mate multiply during their life, female can store sperm and sperm competition is known to occur (Parker, 1970).

**Populations and strains**

In my study I employed flies from two different sources. The LH\(_M\) population was the source for most of my experiments (Paper I, Paper II, Paper III). It is a large outbred wild-type population that has been maintained under the same rearing protocol for over 400 non-overlapping generations. The population is maintained in a set of 56 vials at a large size (1792 adults) under competitive conditions and at moderate larval density (for a detailed description, see Rice *et al.*, 2005).

For paper IV, we used lines originally created by Clancy (2008), derived from \( w^{118} \) isofemales lines in which mitochondria from different origin were introgressed.

Paper V is based on published and publicly available data, and in particular genomic data derived from the Raleigh population (Ayroles *et al.*, 2009) and from seven species of the *Drosophila* genus (Zhang *et al.*, 2007).

**Experimental designs**

Most of my experiments were aimed at collecting genomic data, but a large proportion of the experimental work was dedicated to obtaining flies with the desired characteristics. In some cases, phenotypic data were also collected to allow a comparison or an association between phenotypic and genomic traits.

In Paper I we were interested in assaying female response to mating and remating, and we used females from the base population (LH\(_M\)), mated to males produced by the crossing of two inbred lines obtained from the same population. The samples (virgin or mated females) were collected 6 h after mating.
Comprehensive Summary

Paper II describes the results of a relatively long-term study based on an experimental evolution approach, in which different selective pressures were applied to replicate populations derived from the base population (LH_M). In one of the treatments, we removed components of sexual selection for over 50 generations, and tested how female post-mating response changed compared to control populations, in a setup very similar to the one employed in the previous experiment (Paper I).

For Paper III we took advantage of the hemiclone system, a technique developed by Bill Rice (Rice, 1996; Chippindale et al., 2001; Rice et al., 2005), which involves a series of genetic crosses to produce groups of individuals genetically identical for half of the diploid genome. In the first step, a single wild-type male is crossed to several ‘clone-generator’ [C(1)DX, y; f; T(2;3) rdgC st in ri p^Pbw^P] females. These DX-CG females have a Y chromosome, a compound-X and a translocation between chromosomes 2 and 3. Male progeny will therefore inherit the X chromosome and the wild-type autosomic haplotype from the father and the Y and T(2;3) from the mother. A single heterozygous male from F1 is then mated again to a group of DX-CG females, to ‘amplify’ the haplotype. This step can be repeated ad libitum, allowing the maintenance of large numbers of (male) flies all with the intact haploid genome. In the last step, these males are mated either to a second type of females, ‘DX-LHM’ [C(1)DX, y; f], to produce focal hemiclonal males, or to wild-type females, to produce hemiclonal females (Abbott and Morrow, 2011). The resulting lines of flies will share (within-line) a near-complete haplotype. This method has three main advantages over inbred lines. First, the lines are “wild-type” in respect to gene expression, since they are not expected to show a different level of homozygosity than the base population; in quantitative genetics terms, dominance is excluded from the genetic variance component. Second, indirect genetic effects (e.g., maternal effects) are absent, since parents of different lines do not statistically differ from each other in the degree of relatedness with the offspring of each line. Last, artificial reduction of the allelic pool due to loss of lines carrying highly deleterious recessive alleles does not occur (Innocenti, 2011). We used this technique to produce 100 hemiclonal lines, and measure sex-specific fitness before selecting a subset of them to measure gene expression.

As mentioned above, the lines used in Paper IV were originally created by Clancy (2008), and further manipulated by Damian Dowling to ensure successful introgression of the different mitochondria strains.

Finally, Paper V did not require any experimental work, since we

**Microarrays**

We used microarrays as tools to quantify the expression of genes in *D. melanogaster*. Briefly, a microarray is a surface to which thousands of synthetic oligonucleotide sequences (probes) are attached. Each probe represents the complimentary DNA to the target (e.g. a gene in the organism to be studied). The transcripts (RNA) are extracted from the organism, prepared, labeled with fluorescent dye and hybridized to the microarrays, and the measure of abundance of each transcript is inferred from the level of fluorescence of each probe.

This technique has become very popular over the last decade, as it allowed the gathering of a large amount of information about an organism’s transcriptome in a single event. The development of whole-genome arrays for many model organisms, completed by annotation of the probes to genes with known functions, has made it possible for researchers to extend the use of these tools to the field of evolutionary biology.

Among the many manufacturers of microarray, we opted for Affymetrix GeneChip Drosophila genome 2.0, a widely used single color whole-genome oligonucleotide array. The use of a single chip type throughout all the experiments, and the amount of data already produced with the same microarray greatly facilitates the comparison between datasets.

Sample collection was carried out by flash freezing the flies in a bath of liquid nitrogen, subsequently stored at −80 °C. The protocol for total RNA extraction was developed in our laboratory by slightly modifying manufacturer’s instruction to allow Trizol (Invitrogen) extracted RNA pellets to be purified with RNeasy (Qiagen) Mini Kit columns. Quality control, labeling and hybridization of the RNA samples was carried out by the Uppsala Array Platform.

More recently, RNA-seq techniques are developing fast and show the potential to outcompete microarrays for transcriptome profiling, but the comparisons between the two show a general agreement in the results, proving that microarrays today are still a reliable (and cheaper) solution with the advantage of a solid and mature statistical framework (Malone and Oliver, 2011).

**Availability of the data.** All the gene expression data obtained in our experiments were deposited in the Gene Expression Omnibus database,
and are available with the following accession numbers:

**Paper I:** GSE12834  
**Paper II:** GSE30089  
**Paper III:** GSE17013  
**Paper IV:** GSE24729

**Statistical analysis**

Statistical analysis of microarray data has developed a lot during the last few years. Among the wide range of solutions offered by microarray manufacturers or independent institutions, the BioConductor project (G Gentleman *et al.*, 2004) has been my tool of choice. It is a suite of open source bioinformatics software which has the advantage of running within the flexible R environment (R Development Core Team, 2009), enabling the user to perform the full stream of analysis in a uniform language.

Raw data were pre-processed using the well established and widely adopted algorithms RMA (Irizarry *et al.*, 2003) and VSN (Huber *et al.*, 2002), while for the downstream analysis we employed a different set of statistical packages, according to the specific goal of the experiments (see Material and Methods sections in each paper). In particular, in Paper V we used WOMBAT (Meyer, 2007), a stand-alone tool for REML multivariate mixed models analysis.
Results and discussion

Female post-mating response

Although in *D. melanogaster* the amount of sperm transferred from a male to a female in a single mating is sufficient to fertilize a large number of eggs, females mate multiple times and relatively frequently. This is surprising, giving that additional matings seem to cause a reduction in female reproductive success (Chapman *et al*., 1995; Kuijper *et al*., 2006).

The female response to mating in fruit flies has been the subject of earlier studies, which tested the differences in gene expression between virgin and mated females (Lawniczak and Begun, 2004; Mack *et al*., 2006; Kapelnikov *et al*., 2008) or their response to different components of the ejaculate (McGraw *et al*., 2004, 2008).

The specific aim of my study was to try to isolate the effect of mating from the vast change in female physiology due to activation of the reproductive process, as the female mating status shifts from virgin to mated, and eventually identify candidate processes generating the cost of mating. To do so, I investigated the transcriptional changes occurring between virgin, mated and re-mated female fruit flies.

Overall, we found that mating affects the expression of a large number of transcripts (4369 at a 5% false discovery rate) confirming mating has a dramatic effect on female phisiology. We grouped these transcripts into four categories according to their expression profile: groups of genes which respond to the first mating and then return towards the initial condition (group A, \( n = 271 \)) or maintain the same level of expression (B, \( n = 1389 \)), groups of genes which show cumulative response after mating (C, \( n = 1794 \)) or respond only to the second mating (D, \( n = 915 \)).

We used the Gene Ontology (GO) annotations to assign a biological function to the significant transcripts. The GO project provides structured controlled vocabularies (i.e. ontologies) that describe gene products in terms of their associated biological processes, cellular components and molecular functions in a species-independent manner (Ashburner *et al*., 2000). It has a hierarchical-like structure (technically as a directed acyclic graph) which allows to link genes to more or less specific
terms. While this database generally provides very useful information, some caveat should be discussed. The first is how to determine when a functional category is predominantly represented among a group of genes: any set of genes will tend to be annotated to large categories rather small categories just by chance. To control for this effect, a hypergeometric test for over-representation (equivalent to a Fisher’s exact test) is usually performed. A second problem can derive from annotation bias: in species where functional annotation is incomplete, categories representing well characterized functions or processes will contain a larger proportion of genes, while others will not appear simply because links between gene identity and function is unknown. Finally, genes can be annotated to more than one distinct biological function, because single genes can perform different functions in different stages of an organism life-cycle (e.g. larvae, pupae and adults in Drosophila), and different gene products can be produced from the same gene through mechanisms such as alternative splicing. A consequence of these issues is that obtaining over-represented categories is only the first step of an analysis, and the researcher should weight the results by the biological meaning given by the context of the experiment, introducing a level of discretion which must be taken into account when making more general inferences.

The functional analysis of our significant transcripts highlighted two types of responses to mating. The first is characterized by a substantial change in metabolism, linked to activation of reproductive processes including active biogenesis and organization of cellular components due to mitosis and oogenesis, metabolism, as well as targeting and moving of proteins and other compounds. This result, which relates to the switch between a non-reproductive to a reproductive state, is expected and in line with what is expected when production of mature eggs for oviposition is triggered. The second response is most apparent after the second mating and indicates the presence of an immune reaction to mating. A large number of immune-related genes is upregulated by mating, and two of the three components of this immune reaction show the strongest signal: elements for microbial recognition (peptoglycan recognition proteins, PGRPs) and effectors (mainly antimicrobial peptides, AMPs). Induction of Drosophila immune-related genes as a consequence of mating has been documented in previous studies (Lawniczak and Begun, 2004; McGraw et al., 2004; Mack et al., 2006; McGraw et al., 2008), and we provide evidence of the same reaction after subsequent matings. Since mounting an immune response is costly (Zerofsky et al., 2005; McKea...
et al., 2008), repeated activation of this system by multiple matings could account for (part of) the cost of mating (Fowler and Partridge, 1989; Rice et al., 2006).

Evolution of female post-mating response

In Paper II we aimed at investigating the changes in female post-mating response when populations were subject to a change in selective pressures due to an alteration of levels of sexual selection. We used an experimental evolution approach to enforce monogamy in replicate populations of LH_M: females were allowed to mate only once before males were removed, while in the promiscuous treatment, males and females were left to interact following the base population culturing protocol, during which period females mate multiple times.

We tested the female response to selection by measuring differences in both reproductive output and post-mating gene expression profile, after being mated to males from the same and the opposite treatment. The fitness assays showed that monogamous females had lower reproductive success than promiscuous females, regardless of the type of males they were mated with, and this outcome was confirmed under a different set of conditions. The observed reduction in monogamous female fecundity came somewhat as a surprise, because it seems to indicate that the levels of sexual selection in the promiscuous treatment are somehow adaptive for females; in fact, numerous hypothesis has been proposed to explain direct (see Arnqvist and Nilsson, 2000) or indirect (Yasui, 1998) benefits of polyandry. The lack of effects of treatment on males, on the other hand, suggests that monogamous males did not evolve a less ‘harmful’ ejaculate (Chapman et al., 1995), and as a consequence, no increase in fitness was expected in females mated to those males, which would be predicted by a reduced cost of mating. Our design did not remove pre-copulatory sexual selection in the monogamous treatment, but male removal after the first mating introduced relaxed selection for female resistance to male courtship/harassment (which is present on both promiscuous treatment and the base population). Such environmental differences can cause accumulation of slightly deleterious mutation or recombination of existing variation with suboptimal phenotypic effects, potentially contributing to explain the difference in reproductive success between treatments.

The comparison between female expression profiles confirmed a significant and widespread difference between monogamous and promiscu-
ous females, but no effect of the type of males they were mated with. After multiple testing correction, 1141 were significantly differentially expressed, and these genes were highly over-represented among genes found in the previous experiment (see Paper I) to be affected by mating. Remarkably, genes less expressed in monogamous females (compared to promiscuous females) tended to be up-regulated by mating, and vice-versa, indicating that the mating response in females from the monogamous treatment was over-all weaker (or more ‘virgin-like’), which is in line with the observed decreased fecundity. Also, the set of significant transcript is over-represented among the list of genes associated with female fitness, which was obtained from another experiment (Paper III).

To increase the level of detail and gather additional information about the set of transcripts we identified, we made use of a published database, FlyAtlas (Chintapalli et al., 2007). It consists of a collection of gene expression data from samples of several tissues in D. melanogaster, measured with the same microarray platform used in our experiments. This resource provides valuable information about tissue specificity of genes under standard conditions and overcomes some of the limitation of using whole-body samples for our assays. We used the subset of FlyAtlas data corresponding to the significant transcripts to build a square matrix of among-gene correlations, and used an algorithm (van der Laan and Pollard, 2003) to identify significant clusters of correlated genes (or ‘modules’), representing genes with similar expression across tissues. This method performed well and, when combined with GO terms analysis on each single cluster allowed us to extract a stronger ‘signal’ from the data, at a price of a higher volume of information. Details of the results by module are described in Paper II.

**Sexually dimorphic and antagonistic genes**

The aim of Paper III was to explore the genetic basis of intralocus sexual conflict, by associating a complex phenotypic trait such as sex-specific fitness to genome-wide transcriptome variation (Innocenti, 2011).

First, we generated one hundred hemiclonal lines using the technique described in Material and Methods, by sampling one hundred haplotypes from the LH\textsubscript{M} base population. We then measured sex-specific fitness for each line under competitive condition: male fitness, in our assay and in our base population, is determined by the proportion of offspring sired by a male during the 18 h oviposition period in which eggs laid
Sexual conflict and gene expression in *D. melanogaster*

by females are retained for the following generation. In our assay, focal males competed with outbred males from a LH$_M$ population with a genetic marker for eye color. Female fitness was measured as number of eggs laid in the oviposition period mentioned above. We found a rather large proportion of additive genetic variance for fitness (higher in females than males), although it can be expected that controlled laboratory condition would decrease the environmental component of such variation, therefore potentially inflating heritability estimates. We also found negative covariance for fitness between the sexes ($r_{MF} = -0.52$), which confirms a previous results on the same population (Chippindale et al., 2001) and represent the signature of intralocus sexual conflict, since it is the predicted outcome in presence number of traits showing sexually antagonistic selection and positive intersexual genetic covariance.

The following step was a first attempt to determine the identity and number of the traits (genomic loci in our experiment) underlying the observed realized antagonistic fitness variation. Fifteen lines were selected among those exhibiting extreme or minimal intersexual fitness reversal, and individuals of both sexes were collected to obtain the gene expression profiles. The large dataset (120 samples) gave us high power in detecting sex-bias in expression: our results are remarkably similar to a previously published dataset (Ayroles et al., 2009) of comparable size and quality. The similarity of the data, despite the differences in the population employed, the laboratory and the methods of creation of lines (inbreeding), is promising because it confirms the reproducibility of microarray data and indicate that our results reflects true differences in gene expression in adult male female fruit flies (see Innocenti, 2011, for details). After this preliminary analysis, we tested antagonistic selection on gene expression as a significant interaction between sex and fitness on gene expression. We found $\approx 8\%$ of the genes ($n = 1292$) to be significant after multiple testing correction, the vast majority showing opposite sign for sex-specific fitness. The analysis of the chromosomal distribution highlighted an enrichment of the X chromosome for these loci, confirming a longstanding theoretical prediction (Rice, 1984).

Despite the effort in producing high-quality data for a large number of lines yielding good results, some limitations still persist and will hopefully be addressed in the future. The first is a conceptual simplification: we estimated the association between fitness and gene expression, but antagonistic selective pressure represents the potential for intralocus sexual conflict. Such potential is realized whenever covariance be-
tween male and female expression exists. At the level of gene expression this is assumed to be a common occurrence, and preliminary data (not shown) supports this scenario, but the picture is still incomplete. The second limitation is that the sample size (15 hemiclonal lines) is still rather small to estimate efficiently some parameters (i.e. the intersexual genetic correlation), for many ‘problematic’ genes (e.g. transcripts with very low mean/variance in one or both sexes). Hopefully, as costs (the main limiting resource) decrease, such limitation could be overcome. Finally, a whole level of complexity, determined by gene-gene interaction/correlations, has been ignored. To date, the maximum dimensionality of a variance-covariance matrix to be estimated varies with the methods employed for the estimation, but can hardly handle more than a few hundreds of traits. There are important computational and theoretical limitation that are now receiving the attentions of many researchers, and some solutions start to be proposed (Kirkpatrick and Meyer, 2004; Meyer and Kirkpatrick, 2005; Hine and Blows, 2006).

The issues described here represent a major challenge for the field of evolutionary genomics, especially when investigating complex phenotypes with genomic data with a quantitative approach. Paper V represents an attempt in treating (sex-specific) gene expression data in a quantitative genetics multivariate framework.

**Sex-specific selective sieve on mtDNA evolution**

Mitochondria represent an example of unbalanced parental transmission of genetic material: they are interesting from a sexual conflict perspective because mtDNA can only make a direct and adaptive response to selection through females, creating potential for a ‘sex-specific selective sieve’ (Paper IV). We tested this hypothesis by comparing nuclear gene expression profiles in males and females of five *D. melanogaster* strains showing identical nuclear DNA but mitochondria of different origin (from five isolated geographical locations). We found that mtDNA variance does not affect female gene expression but has a large impact in males: the expression of more than 9% of the tested transcripts was found to be dependent on mtDNA. Moreover, the majority of these genes exhibit male-bias in expression. Modules of correlated expression revealed that the largest clusters of transcripts were expressed almost exclusively in the male-limited tissues, the testes and the accessory glands.

Taken together, these results are strong evidence for sex-specific effects.
of mtDNA variation: they can be interpreted as there being an accumulation of mutations on the mitochondrial genome with effects on male, but not female, phenotype. Assuming that most changes in gene expression are deleterious (Gilad et al., 2006), this indicates a much greater mutational load in males than females (Parsch, 2011). This assumption is backed up by indirect links with fitness, because the differentially expressed transcripts are associated with male – but not female – fitness, and one experimental mtDNA haplotype causes male sterility in the w^{1118} genetic background, whereas it does not have this effect in the nuclear background of origin (suggesting the evolution of counteradaptations in the nuclear genome that restore lost male function). Moreover, if widespread differences in gene expression are neutral to selection, this phenomenon is also expected to be observed in females, assuming equal sex-specific sensitivity to transcriptional levels.

As noted by Parsch (2011), a first open question pertains the molecular mechanisms causing mutations on a very small genome to have genome-wide effects on nuclear gene expression. To date, the knowledge of mito-nuclear interaction is somewhat limited, although some networks of anterograde and retrograde regulation have been described (Liu and Butow, 2006; Woodson and Chory, 2008). The discovery of an important regulatory role of non-coding sequences in gene-poor Y chromosome (Lemos et al., 2010) suggests that a similar mechanism might act in mtDNA. A corollary question is how these mechanisms can act differentially in the two sexes, and today it remains an open issue.

**Evolvability and sexual constraints in interspecific divergence of gene networks along lines of genetic variance**

In my last study (Paper V), I decided to address an additional level of complexity of the genome, which was not taken account in Paper III: gene-gene interactions, and in particular genetic networks, and explore variation in their expression as a whole. Genetic networks have been a focus of research for a relatively long time (e.g. Wagner, 1994), and a large body of literature has discussed theoretical aspects of their evolution (Barabasi and Oltvai, 2004). For historical reasons, common concepts in this field were developed somewhat independently from most of the classical evolutionary theory. Computational models mainly focus on topology of networks, and use parameters such as robustness, redun-
dancy, modularity and capacitance, borrowing concepts used in physics or engineering, e.g. for describing electronic circuits (Alon, 2003; Hasty et al., 2002). This approach has been criticized (e.g. Lynch, 2007) because it does not take into account simple concepts in population genetics, which could potentially explain the observed patterns without invoking ad hoc explanations, or because they assume an adaptive significance of all their components. For example, it can be shown how distributed robustness can evolve as by-product mutation and duplication, without a relevant role of natural selection in the process (Lynch, 2007).

We sought to explore the evolution of gene networks from a different perspective: using data on the expression levels of genes in well-characterized genetic pathways, we explored the existing variation within- and across-species, and between the sexes, in the Drosophila group. We choose three networks, the Map-Kinase (Widmann et al., 1999), the Toll (Valanne et al., 2011) and the insulin-receptor/Foxo pathways (Nuzhdin et al., 2009). To describe the patterns of variation within species, we used the dataset from Ayroles et al. (2009), which measure gene expression in inbred lines from a field-collected population of D. melanogaster, to build a \( G \) matrix of genetic variances and covariances for all the components of the network, both in males and females separately, and including the \( B \) matrix of across-sexes covariances (Lande, 1980). To assess divergence across species, we build a \( D \) matrix of phenotypic variances and covariances between species means for the same set of genes, using the datasets from Zhang et al. (2007), which includes seven species (D. melanogaster, D. simulans, D. yakuba, D. ananassae, D. pseudoobscura, D. virilis and D. mojavensis) in the Drosophila group.

We applied a number of different frameworks to compare these multidimensional spaces, with the aim of testing whether divergence across species occurred along lines of maximum genetic variation within species. In other words, if divergence, or speciation, is constrained by the available standing genetic variance in a population, and how between-traits constraints influence this process. We started using Schluter’s (1996) approach, which consists of calculating the angle between the two eigenvectors which explains most of the genetic variance \( (g_{max}) \) and most of the realized divergence \( (d_{max}) \). We also tried methods employed in previous studies: compare the subspaces by sequential addition of eigenvectors (Strang, 1998) or using portions of the \( G \) and \( D \) using the Krzanowski (1979) approach (both described in Blows et al., 2004). We eventually decided to employ Hansen and Houle (2008) approach (described in the
Material and Methods) because it is derived from the rigorous classical framework of multivariate response to selection (Lande, 1979). Very briefly, it is based on calculating respondability, the vector which describes the response to selection (the product of $\mathbf{G}$ and the divergence vector $\beta$). Evolvability, $e_\beta$, will be the projection of respondability on the direction of $\beta$, or the amount of response which lies in the direction of the observed divergence.

In general, we found that the effective dimensionality of the genetic networks (a measure of the complexity of the variance structure) is fairly large (Kirkpatrick, 2008, $n_D$ between 2 and 3), with a rank equal or greater than half the number of traits ($k$), for both the full or the sex specific matrices, indicating that genetic variance spans in a rather large number of dimensions of the $k$-space. The complexity of expression profiles of a network is therefore intermediate between that of a single gene and the sum of the $k$ genes in the network. In every pairwise species comparison, we found the observed evolvability (Hansen and Houle, 2008) to be close to maximum evolvability (i.e. $e_\beta$ calculated on $\mathbf{g}_{max}$), and always higher than the average evolvability, indicating some degree of constraint (see Results of paper V for an explanation). Conditional evolvability – $c_\beta$, evolvability in presence of stabilizing selection – is higher than average when $B$ is taken into account, while for males and females separately $c_\beta$ exhibit more variable patterns, and interestingly, in some cases (MapK and Toll pathways), it seems to be negatively correlated across species pair in the two sexes.

Admittedly, there is a step that naturally follows the estimation of such parameters, and is lacking in the present study: the estimation of the uncertainty around these parameters. We ran 200 replicates of each model, bootstrapping samples at the ‘line’ level, to obtain a confidence interval around our estimate of $\mathbf{G}$. Multivariate models are very computationally intensive, and obtaining a large number of bootstrap replicates for models with more than thirty traits can require months. The REML algorithm is also likely to fail to reach convergence for resampled datasets, which leads to the necessity of increasing the number of iterations. We eventually managed to obtain a minimum number of pseudo-replicates ($n > 150$), but faced a more serious theoretical problem. The key parameters estimated with Hansen and Houle (2008) approach, $e_\beta$ and $c_\beta$, can be interpreted only in relationship with their maximum, average and minimum values ($e_{max}$, $\bar{e}$, $\bar{c}$ and $e_{min}$), but these values change with $\mathbf{G}$ at each iteration, making it unclear how the intervals should be calculated. Moreover, since $e_{max}$ and $e_{min}$ are calculated
using \( g_{\text{max}} \) and \( g_{\text{min}} \), it is critical that in every replicate they refer to the ‘same’ eigenvector, which is highly likely for \( g_{\text{max}} \), but not for \( g_{\text{min}} \) (since eigenvectors are ordered by their eigenvalues, the last eigenvector can be unstable, \textit{i.e.} change order, when eigenvalues approach 0). For these reasons, at this stage we preferred not to include uncertainty estimates, until a stronger theory is developed.
Conclusions

In my thesis, I aimed to show that gene expression data can provide novel insights in the study of sexual conflict. To achieve this aim, I addressed several open questions in the field, and used microarrays to quantify and compare the expression profiles of *D. melanogaster* genomes, also integrating them with phenotypic data.

My experiments involving a simple comparison between expression profiles of different treatments produced a list of candidate genes which are putatively involved in the response to such treatments. When comparing gene expression in virgin, single mated and re-mated females I found substantial differences between them, and using functional annotation of the genes I tried to identify the biological processes underlying these changes. The comparison between flies with different mitochondrial variants yielded the unexpected results of an unbalanced response between the sexes: a negligible difference among females and a large difference among males. The analysis of sex-specific expression and tissue specificity of these transcripts highlighted how a large proportion of them are male-biased and predominantly active in male reproductive tissues.

More complex designs required the simultaneous measurement of phenotypic and genomic data. These types of experiments allowed a direct comparisons between the two, namely, to measure the mean difference or the variance of a phenotypic trait and investigate its underlying ‘genetic’ basis. I focused again on female post-mating response, comparing both female reproductive success and gene expression profile after over 30 generations of experimental evolution under two different sexual selection pressures. These two traits indicated a decreased reproductive success in monogamous females due to a tendency of the mating-responsive genes to show a sub-optimal expression level, suggesting an overall weaker post-mating response. In a different experiment, I measured male and female reproductive success in a set of hemiclonal lines, and associated their fitness with their expression profiles, measured in both sexes and as reproductively active adult individuals. Notably, this study allowed me to compare the expression of the same haplotype in both sexes and investigate the effects of the antagonistic genetic variation for fitness at
In the last experiment I decided to make use of two different sets of published data, to integrate them in an attempt to test a novel hypothesis (i.e., which was not addressed by the original studies). One of the strengths of microarray data lies also in its public availability and relatively standardized experimental designs and annotations, which allows their integration and their subsequent use in different contexts. In the future, an improved framework for data storage, handling and annotation of these type of data could facilitate meta-analysis and generalization of results, which are central for biological inferences. When technology resolves the technical issues related to the high per-sample costs and the statistical analysis of large and complex datasets, a particularly interesting subject of future research would be the study of gene expression variation within and between population and their relationship, and the unravelling of between-sexes and between-genes covariation, which both require large sample sizes, long and carefully planned experiments and exceptional computational resources.

It is clear that the type of microarray studies performed are mainly exploratory: it does not identify and characterize a single molecular mechanism in its entirety, but rather points the direction for future experiments, to be carried out with different techniques and more focused aims. On the other hand, it provides a relatively easy tool for screening a vast amount of traits underlying phenotypes of interest, with a two-fold benefit: a less laborious task of processing a single trait at a time, and provide information on interaction of multiple traits. The statistical price to pay for obtaining high volume of information is a certain number of false positives and false negatives, which depends on the background noise, on arbitrary choices on the experimental design or manipulation of the samples with unknown consequences, and the skills of the experimenter, which is faced with a large amount of data and has the task of extracting the true signal. In particular, it is responsibility of the researcher to single out the biologically meaningful information, choosing some types of analyses over others among the countless possible, and potentially discarding useful elements. Such elements, however, can be of special interests to other researchers. For this reasons, full availability and reproducibility of analyses and results in open and transparent formats is essential and should be prioritised.
Sammanfattning på svenska

Konflikter mellan kön och genuttryck hos bananflugan, Drosophila melanogaster

Definitionen av honor och hanar är storleken på den köns cell de producerar. Medan honor producerar stora och icke rörliga köns celler (ägg), producerar hanar små och rörliga köns celler (spermier). Denna obalans mellan könen i investering av att könen även har olika strategier för att maximera deras “fitness”. Resultatet är att honor och hanar skiljer sig i deras optimala egenskapsvärd för morforlogiska, fysiologiska egenskaper samt i beteendekarakterer. Dessa skillnader kan resultera i en konflikt mellan de evolutionära intressen hos de respektive könen – en så kallad könskonflikt. Könskonflikter kan förekomma både på nivån för interactioner mellan könen eller som en selektion som age-rar i motsatta riktningar på en egenskap båda könen har, beroende på den genetiska uppbyggnaden av egenskapen.

Trots de stora fenotypiska skillnader som finns mellan könen är de nästan identiska genetiskt sett. Hos de flesta arter skiljer sig köns genomen enbart på några få gener som sitter på de könsspecifika kromosomerna. Anledningen till att vi observerar skillnader mellan könen måste därför vara att könen använder sina arvsanlag på olika sätt, eller med andra ord, att könen uttrycker sina respektive genom olika. Genut-
tryck är den lägsta fenotypiska nivå där det finns stora skillnader mellan könen. Jag argumenterar för att genomuttryck därför är särkilt lämpliga för studier av konflikter mellan könen, givet att det går att etablera en länk mellan genotyp och fenotyp.

Jag har två mål med min avhandling: Först, att undersöka poten-

tialen för att använda genomuttryck i studier av sexuella konflikter inom och mellan locus. Mitt andra mål är att använda dessa genomiska verktyg kombinerat med mer “klassiska” experimentella tillvägagångs sätt för att vidareutveckla den kunskap som finns om hur honor och hanar samevolverar samt visa på könsspecifika anpassningar och begränsningar. Mer specifikt jag har jag fokuserat på:
Comprehensive Summary

- Honlig respons efter parning och dess evolution under olika former av sexuell selektion.
- Sexuell dimorfism och könsspecifika selektionstryck på uttrycket av bananflugans genom.
- Könsspecifik påverkan av interaktioner mellan mitokondrievariation och det nukleära genomet.
- Inom- och mellanarts orientering av genetisk variation i honor och hanar.


**Resultat**

Trots att mängden sperma som en hane av *D. melanogaster* överför till honan vid ett enda parningstillfälle är tillräckligt för att befrukta ett stort antal ägg, så parar sig honor flera gånger och relativt ofta. Detta är förvånande eftersom flera parningar kan sänka honans reproduktiva framgång. Det specifika målet med en av mina studier (uppsats I) var att försöka isolera effekten av själva parningen från de övriga förändringar som induceras i honors fysiologi när olika reproduktionprocesser aktiveras, och slutligen identifiera kandidatprocesser som genererar kostnaden av parningar. För att göra detta, undersökte jag förändringar i tran-skriptomet mellan oparade flugor, parade flugor samt återparade flugor. Vi fann att parning påverkar uttryck av ett stort antal transkript (4369 med en så kallad "false discovery rate" på 5%) vilket bekräftar att parning har en dramatisk påverkan på honors fysiologi.

Analyserna av de signifikant skilda uttryckta transkriptomen indikerade två typer av responsen till parning. Den första karakteriseras av stora förändringar i metabolismen, och är bunden till aktivering av reproduktiva processer som till exempel aktiv biogenesi samt organisation
Sexual conflict and gene expression in *D. melanogaster*

av cellulära komponenter på grund av mitosis, oogenesis, metabolism, samt inriktning och förflyttningar av proteiner och andra ämnen. Den andra responsen är tydligast efter den andra parningen och indikerar förekomsten av en immunrespons på parning. Ett stort antal immunrelaterade gener uttrycktes mer efter parningar, och två av de tre delarna av immunförsvarvare visar på den starkaste signalen: ämnen som används för att känna igen mikrober (peptoglycan igenkänningsproteiner, PGRPs) och ämnen som motarbetar infektioner (framförallt antimikrobiella peptider, AMPs). Induktion av *Drosophila* immunrelaterade gener efter parning har dokumenterats i tidigare studier, och jag visar här att samma reaktion förekommer även efter påföljande parningar. Eftersom aktivering av immunsystemet medför kostnader, kan upprepade aktiveringar av immunsystemet förklara delar av kostnaden honor har av flera parningar.

I nästa försök ville jag undersöka förändringar i honors postparnings-respons efter att populationerna hade blivit utsatta för olika selektionstryck, genom att nivåerna av sexuell selektion de olika replikatpopulationerna av LH₉M utsattes för förändrades experimentellt (uppsats II). I den monogama behandlingen tillåts honor enbart para sig en gång innan hanarna togs bort från experimentuppställningen. I den promiskuösa behandlingen hölls honor och hanar tillsammans på samma sätt som i experimentprotokollet för baspopulationen, under vilken tid honorna parar sig flera gånger.

Jag testade sedan honornas respons till dessa selektionstryck genom att mäta skillnader i både reproduktiv framgång och post-par/-nings/-genut/-tryck, efter att de hade parats med hanar både från sin egen behandling och den motsatta behandlingen. Resultaten visade att monogama honor hade lägre reproduktiv framgång än de promiskuösa honorna, oavsett vilken typ av hane de hade parat sig med. Dessa resultat bekräftades även under ett annat experiment utfört under andra förutsättningar.

Jämförelsen mellan honors genuttryck bekräftade en signifikant och vidspridd skillnad mellan monogama och promiskuösa honor, men jag fann ingen effekt av vilken typ av hane de parats med. Efter att ha korrigerat resultaten för multipla test, 1141 transkript var signifikant skilda mellan grupperna och dessa gener var mycket överrepresenterade bland de gener som i förra experimentet fanns vara påverkade av parning. Anmärkningsvärt var att gener som var mindre uttryckta i monogama honor (jämfört med promiskuösa honor) tenderade till att vara uppregulerade av parning, och vice versa. Detta indikerar att parnings/-responsen
hos honor från den monogama behandlingen var, överlag, svagare (mer lik de oparade honorna), vilket stämmer överens med den minskade fe-kunditet som observerades.


För att fastställa identiteten och antalet av de gener som låg bakom den antagonistiska fitnessvariationen mellan könen testade jag antagonistisk selektion på genuttryck som en signifikant interaktion mellan kön och fitness på genuttryck. Cirka 8% av generna var signifikanta efter korr.korrektion för multipla test, och utav dessa hade de absolut flesta motsatta tecken för könsspecifik fitness. Analysen för hur dessa generna var fördelade över kromosomerna visade på att de flesta sitter på X-kromosomen, vilket bekräftar de teoretiska förväntningarna.

Mitokondriellt DNA (mtDNA) är ett exempel på obalanserad parental nedärvning av genetiskt material: ur perspektiv av konflikter mellan kön är de intressanta eftersom mtDNA enbart kan visa på direkta, adaptiva, responser genom selektion på honor, och detta ger möjlighet för ett könsspecifikt selektivt såll (uppsats IV). Jag testade denna hypotes genom att jämföra nukleärt genuttryck i hanar och honor av fem olika D. melanogaster linjer med identiskt nukleärt DNA men med mitkondrier av olika ursprung (från fem isolerade geografiska områden). Jag fann att mtDNA variation inte påverkar honors genuttryck men har en stor effekt på hanars genuttryck: trycket av fler än 9% av de transkript som inkluderades i studien var beroende av mtDNA. Majoriteten av dessa gener visade dessutom på hanligt bias i deras uttryck. Moduler av korrelerat uttryckta gener visade på att de största klustren av transkript var uttryckta i vävnader som enbart förekommer i hanar, testiklarna och accessoriska körtlar.

I den sista uppsatsen (uppsats V), använder jag tidigare publicerat
data för att jämföra riktningen av inomarts genetisk variation för ett set av genomiska “pathways” mätta från en stor panel inavlade linjer av D. melanogaster med variation mellan andra arter i släktet Drosophila. Detta med syftet att testa om divergensen mellan arter/artbildning sker i riktningar av maximal inomarts genetisk variation. En av fördelarna med microarraydata ligger i att datan publiceras i depositioner och ofta kommer från relativt standardiserade experiment, vilket gör det möjligt att integrera olika studier och använda dem i olika sammanhang. I framtiden kommer förbättrade möjligheter till att spara och hantera data underlätta för metaanalyser och generaliseringar av resultat, vilket är centralt för en ökad biologisk förståelse.

Sammanfattningsvis, tekniker för att uppskatta genuttryck kan ge användbara och nya resultat när konflikter mellan könen studeras, och kan även kombineras med fenotypiskt data. Eftersom det insamlade datat dessutom är tillgängligt efter publicering, tillåts jämförande av olika undersökningar vilket ger stora möjligheter till fortsatta.
Riassunto in italiano

Conflitto sessuale ed espressione genica in *Drosophila melanogaster*

Maschi e femmine vengono definite dalla taglia dei gameti che essi producono. I maschi producono gameti piccoli e mobili (detti spermatozoi) e le femmine producono gameti grandi e immobili (uova). Questo sbilanciamento nell’investimento parentale primario porta a strategie diverse attraverso le quali i sessi possono massimizzare la loro *fitness*. Come risultato, i maschi e le femmine possono avere *optima* differenti per caratteri morfologici, fisiologici o comportamentali. Ogni qualvolta tale differenza si presenta, esiste potenziale per un conflitto sessuale, genericamente definito come “un conflitto tra gli interessi evolutivi dei due sessi”. A seconda dell’architettura genetica dei caratteri coinvolti, può manifestarsi a livello di interazioni maschio-femmina o prendere la forma di selezione che agisce per spostare la media di un carattere comune ai due sessi contro il segno della sua correlazione genetica.

Nonostante le vistose differenze fenotipiche tra maschi e femmine, essi condividono quasi interamente il loro patrimonio genetico. Nella maggior parte delle specie, i genomi dei due sessi differiscono solo per alcuni geni che risiedono sui cromosomi sessuali. Il dimorfismo sessuale che osserviamo è quindi prodotto dal diverso uso che i sessi fanno del materiale genetico, o, in altre parole, dell’espressione differenziale dei rispettivi genomi. L’espressione genica rappresenta il più basso livello ‘fenotipico’ per cui esiste una ingente differenza tra i due sessi. Io sostengo che questa caratteristica rende l’espressione genica un carattere particolarmente utile nello studio del conflitto sessuale, a condizione di stabilire un collegamento adeguato tra il materiale genetico e i caratteri fenotipici di interesse.

Lo scopo della mia tesi è duplice. Primo, sondare il potenziale dell’analisi del trascrittoma nello studio del conflitto sessuale intra- ed inter-locus. Secondo, usare questi strumenti genomici, anche combinandoli con approcci sperimentali più ‘classici’, per raggiungere nuovi risultati nello studio della coevoluzione tra il maschio e la femmina e dell’adat-
tamento e dei vincoli evolutivi sesso-specifici. In particolare, mi sono concentrato su:

- Risposta post-copulatoria nella femmina e la sua evoluzione a differenti regimi di selezione sessuale.
- Dimorfismo sessuale e pressioni selettive sesso-specifiche nell’espressione genica del genoma del moscerino della frutta.
- Effetti sesso-specifici della variazione mitocondriale nella loro interazione col genoma nucleare.
- Orientamento intra- e inter-specifico della varianza genetica nei maschi e nelle femmine.

In tutti i miei esperimenti ho utilizzato l’organismo modello *D. melanogaster* (Moscerino della frutta), un insetto dell’ordine dei Ditteri. È una specie cosmopolita sexualmente dimorfica, il cui genoma è stato completamente sequenziato e ben caratterizzato: contiene più di 14000 geni, distribuiti su quattro coppie di cromosomi con un sistema di determinazione del sesso di tipo X/Y. Nella maggior parte dei miei esperimenti ho utilizzato una popolazione di grandi dimensioni e adattata al laboratorio (LHM). Per quantificare l’espressione genica in questa specie, abbiamo utilizzato microarray oligonucleotidici monocolore (Affymetrix GENECHIP Drosophila genome 2.0).

**Risultati**

Nonostante in *D. melanogaster* la quantità di sperma trasferito da un maschio ad una femmina in un singolo accoppiamento è sufficiente a fertilizzare un grande numero di uova, le femmine si accoppiano ripetutamente e con frequenza relativamente alta. Questo è sorprendente, dal momento che un numero eccessivo di accoppiamenti sembra causare una riduzione del successo riproduttivo della femmina. Lo scopo specifico del mio studio (articolo I) è stato di provare ad isolare l’effetto dell’accoppiamento dal cambiamento fisiologico dovuto all’attivazione del processo riproduttivo, e possibilmente identificare i processi che generano il costo dell’accoppiamento. Per fare questo, ho investigato i cambiamenti trascrizionali che avvengono tra femmine vergini, accoppiate e accoppiate una seconda volta. In generale, abbiamo trovato che l’accoppiamento altera l’espressione di un gran numero di trascritti (4369 ad una soglia
di 5% di falsi positivi), confermando the l’accoppiamento ha un effetto drammatico sulla fisiologia della femmina.

L’analisi funzionale dei trascritti statisticamente significativi ha evidenziato due tipi di risposte all’accoppiamento. La prima è caratterizzata da un cambiamento sostanziale nel metabolismo, collegato all’attivazione di processi riproduttivi che includono biogenesi attiva e organizzazione di componenti cellulari conseguenti a mitosi e oogenesi, metabolismo, così come individuazione e spostamento di proteine e altre molecole. La seconda risposta è più evidente dopo il secondo accoppiamento e indica la presenza di una reazione immunitaria all’accoppiamento. Un grande numero di geni collegati al sistema immunitario viene stimolato dall’accoppiamento, e due dei tre componenti della risposta immunitaria mostrano il segnale più forte: elementi di riconoscimento microbico (peptoglycan recognition proteins, PGRPs) ed effettori (principalmente peptidi antimicrobici, AMP). L’induzione di geni legati al sistema immunitario come conseguenza dell’accoppiamento è stato documentato in studi precedenti, e noi contribuiamo la dimostrazione di questa reazione come conseguenza di accoppiamenti ripetuti. Poiché indurre una risposta immunitaria è costoso, una attivazione ripetuta di questo sistema tramite accoppiamenti multipli può spiegare parte del costo dell’accoppiamento.

In un secondo tempo, abbiamo studiato i cambiamenti nella risposta post-copulatoria della femmina quando la popolazione è soggetta a cambiamenti della pressione selettiva dovuta a un’alterazione dei livelli di selezione sessuale (articolo II). Abbiamo utilizzato un approccio di evoluzione sperimentale per imporre la monogamia in popolazione replicata di LH_{M}: alle femmine è stato permesso di accoppiai una sola volta prima di rimuovere i maschi, mentre nel trattamento ‘promiscuo’, maschi e femmine sono stati lasciati interagire seguendo il protocollo della popolazione ancestrale, periodo durante il quale le femmine si accoppiano ripetutamente.

Abbiamo testato la risposta della femmina alla selezione misurando la differenza nella loro fecondità e nel loro profilo di espressione genica post-copula, dopo essersi accoppiate con maschi appartenenti al loro stesso trattamento o al trattamento opposto. I saggi di fecondità hanno evidenziato un successo riproduttivo inferiore per le femmine monogene rispetto a quelle promiscue, indipendentemente dal tipo di maschio a cui venivano accoppiate, e questo risultato è stato confermato in un range di condizioni diverse.
Il confronto tra profili di espressione ha confermato una generale e significativa differenza tra le femmine monogame e quelle promiscue, ma nessun effetto del tipo di maschi a cui sono state accoppiate. A seguito di correzione per test multipli, 1141 trascritti sono risultati espressi in modo differenziale, e questi geni sono altamente sovra-rappresentati tra geni alterati nella loro espressione dall’accoppiamento, come descritto nello studio precedente. In particolare, geni meno espressi in femmine monogame tendono ad aumentare la loro espressione come conseguenza dell’accoppiamento, e viceversa, suggerendo che la risposta all’accoppiamento nelle femmine del trattamento monogamo è generalmente meno intensa, in accordo con la ridotta fecondità osservata.

Lo scopo dell’esperimento successivo (articolo III) è stato esplorare le basi genetiche del conflitto sessuale intralocus, associando un tratto fenotipico complesso come la fitness sesso-specifica alla varianza trascrittornica dell’intero genoma.

In primo luogo, abbiamo generato cento linee emiclonali campionando cento aplotipi dalla popolazione ancestrale, e successivamente abbiamo misurato la fitness sesso-specifica per ciascuna linea in condizioni di competizione. Abbiamo trovato una consistente porzione di varianza genetica additiva per la fitness, e una covarianza negativa per la fitness nei due sessi, che rappresenta il segno di conflitto sessuale intralocus.

Nel passo successivo, abbiamo provato a determinare l’identità e il numero dei geni che sottendono alla variazione antagonista per la fitness che è stata osservata. Abbiamo testato la selezione antagonista sull’espressione genica come interazione significativa tra il sesso e la fitness, e trovato $\approx 8\%$ dei geni significativi a seguito di correzione per test multipli, la gran parte dei quali mostra segno opposto per la fitness sesso-specifica. L’analisi della distribuzione cromosomica ha evidenziato una concentrazione di questi loci sul cromosoma X, confermando una predizione teorica precedente.

I mitocondri rappresentano un esempio di trasmissione parentale sbilanciata di materiale genetico: sono interessanti dal punto di vista del conflitto sessuale perché il mtDNA può rispondere in modo diretto e adattativo alla selezione solo attraverso le femmine, creando potenziale per un ‘setaccio selettivo sesso-specifico’ (articolo IV). Abbiamo testato questa ipotesi comparando il profilo di espressione genica nucleare nei maschi e nelle femmine in cinque linee di D. melanogaster con identico DNA nucleare ma mitocondri di origini differenti (da cinque località geografiche isolate). Abbiamo trovato che la varianza nel DNA mitocond-
diale non influenza l’espressione genica nelle femmine ma ha un impatto importante nei maschi: l’espressione di più del 9% dei trascritti testati è risultato essere dipendente dal mtDNA. Inoltre, la maggioranza di tali geni mostra espressione a livelli più elevati nel maschio. I moduli di espressione correlata hanno rivelato che i gruppi di trascritti più grandi sono espressi quasi esclusivamente nei tessuti solo maschili, quali i testicoli e le ghiandole accessorie.

Nell’ultimo esperimento (articolo V), ho usato dati pubblicati in database per comparare la direzione della varianza genetica intraspecifica per un gruppo di pathways genomici con la varianza tra specie sorelle nel gruppo *Drosophila*, allo scopo di testare se la differenziazione interspecifica (speciazione) avviene lungo linee di minima resistenza evolutiva (*i.e.* che massimizzano la varianza genetica intraspecifica). Uno dei punti di forza dei dati prodotti con microarray è la loro disponibilità pubblica, unita a un design e un’annotazione relativamente standardizzati, che permette la loro integrazione e il loro uso in contesti diversi. In futuro, un ambiente di lavoro disegnato per archiviare tali dati, maneggiarli e annotarli in modo efficiente potrebbe facilitare le meta-analisi e la generalizzazione dei risultati, che sono passi fondamentali nell’inferenza a carattere biologico.

In conclusione, le tecniche per l’espressione genica possono produrre nuovi e utili risultati nello studio del conflitto sessuale, anche in associazione a dati fenotipici. Inoltre, la disponibilità pubblica dei set di dati permette l’integrazione dei risultati, un migliore livello risolutivo dei dettagli funzionali e il test di nuove ipotesi.
Acknowledgements

First of all, in order of importance, I’d like to thank my supervisor, Ted. It has been an honour and a pleasure to work with you, and I could not have hoped for better guidance throughout these years: I always had the right advice, and all the decisions you made were in my best interest. I hope many other students will have the privilege of your supervision in the future.

I also want to thank Göran Arnqvist, my second supervisor, for helping me with relevant comments and advice throughout these years, and for several disputes about open software, statistical tools and a variety of other subjects. I am happy to say that neither of us has moved a millimeter from his own position. A special thanks also goes to Damian Dowling and Steve Chenoweth, who both supervised me for some time during my visit to Australia. It has been very interesting to collaborate with you and I have learned heaps of stuff.

Thanks to all the people in my department, and the neighbouring departments. The EBC building has been an interesting and entertaining place to spend long working - and not working - hours. I had several insightful conversations with many people, and received very useful advice from some people. I also spent a large amount of time with a few people, doing sensible and slightly less sensible (but equally rewarding) things. Most of the people in these categories overlap, so I will thank you in tightly sparse order: Mirjam, Björn, Urban & the fly lab, Simone and Alexei, Làra, Kasja, Jossan and Mårten, Amber, Richard, Erem, Alex Fish, Arild, Elena (and Jeff), Jessica, Niclas, Axel, Andreas, Hanna, Hanne and Alex, Magnus, Sandra, Pia, Phillip, Christian, Hannes, Maria and Omar, Yoshi, Matt, Fernando. Thanks to Reija for all the help in the molecular lab.

My life outside the EBC has been quite entertaining as well. A bunch of italians (or italian-oriented) friends here in Uppsala have been excellent company and support during those long winters, and helped me to remember how a good dinner has nothing to do with the quality of the food: thanks to Paolo and EllaKlara, Andrea and Rocio, Francesco and Maja, Moreno and Kristina. And Mele, of course. The poker night team, Andrea, Steve, Beau and Sophie are responsible for joyful evenings full
of good fun and bad poker. Thanks also to Reelika and Stefan, Misia and Samer, Nora and Loic for memorable events and evenings.

An ice cream without the cream, as a king without a crown: a special thanks goes to Piggelin. Perfect from its name to its taste, this humble popsicle is the hidden jewel of Swedish culture. When will you get the acknowledgement you deserve?

Before moving to Sweden, I had a brief and intense experience in London, which was so exciting thanks to many people, in particular Shaula who was my spiritual guide in the manifold complexity of the big city, Ilaria, and Sofia, who hired me at the coolest coffee shop in town.

During my undergraduate and master studies I managed to make a few friends. I’d like to thank Daniele and Giovanni, with whom I shared my first year at Physics, and made me realize that I was not smart enough for understanding simple physical systems, and I would have had better chances with complex biological systems. Mateo, Cristina, Stefano, Madda, Amina, Mazzo and Fabi made my days at BioDip in Milan a very pleasant experience.

It’s good not to be forgotten. I thank all my mates back in Italy: every time I had a chance to return to my home town, it felt like I never left. I have always been welcomed with a casual nod, as usual. I suspect most of them never actually realized I was not living in Bergamo anymore. Thanks to Barbo and Laura, Figù and Angela, Mateo and Preda, Gianlu, Panci, Panunzio and Chiara, Guido and Fresca, Pietro and Chiara, Elena and Fede, Moch, Erminio and Stefi, Curcu and Diana, Giulia and Andrea, Valeria and Nicola, Mancio and Sara, Arres, Rosso and Fede.

Thanks to the best family ever: my parents, Dario and Teresa, my siblings, Carlo and Delia, and also Zia Lisa, Zia Delia, Francesca and Dante. You have been all very supportive (except you, Dante, and being two years old is not an excuse).

Finally, in reversed order of importance, I want to thank Lauren.
Bibliography


Sexual conflict and gene expression in *D. melanogaster*


Sexual conflict and gene expression in *D. melanogaster*


Sexual conflict and gene expression in \textit{D. melanogaster}


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 843

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology.