Insect Molecular Biology

Insect Molecular Biology (2015) **24**(2), 139–154 doi: 10.1111/imb.12154

INVITED REVIEW

We can't all be supermodels: the value of comparative transcriptomics to the study of non-model insects

Sara J. Oppenheim,* Richard H. Baker,* Sabrina Simon† and Rob DeSalle*

**Department of Entomology, Division of Invertebrates, Sackler Institute for Comparative Genomics, American Museum of Natural History, New York, NY, USA;*
†Biosystematics Group, Wageningen University, Wageningen, The Netherlands

Abstract

Insects are the most diverse group of organisms on the planet. Variation in gene expression lies at the heart of this biodiversity and recent advances in sequencing technology have spawned a revolution in researchers' ability to survey tissue-specific transcriptional complexity across a wide range of insect taxa. Increasingly, studies are using a comparative approach (across species, sexes and life stages) that examines the transcriptional basis of phenotypic diversity within an evolutionary context. In the present review, we summarize much of this research, focusing in particular on three critical aspects of insect biology: morphological development and plasticity; physiological response to the environment; and sexual dimorphism. A common feature that is emerging from these investigations concerns the dynamic nature of transcriptome evolution as indicated by rapid changes in the overall pattern of gene expression, the differential expression of numerous genes with unknown function, and the incorporation of novel, lineage-specific genes into the transcriptional profile.

Keywords: RNA-Seq, non-model organism, differential expression, comparative transcriptomics, NGS, insects.

First published online 19 December 2014.

Correspondence: Dr. Robert DeSalle, Sackler Institute for Comparative Genomics, American Museum of Natural History, 79th Street at Central Park West, New York, NY 10024, USA. Tel.: 212-769-5670; e-mail: desalle@amnh.org

Introduction

How phenotypic diversity is generated is a major question in modern biology. The basis of such diversity in insects is an interesting question not only evolutionarily, but also in an applied context, because insects are major agricultural pests and vectors of many human diseases. As the genomics revolution has progressed, it has become clear that simply detecting the genetic differences between two phenotypically divergent forms often tells us little about the underlying basis of the divergent trait. This is particularly true in cases where highly similar (or even identical) genomes give rise to different phenotypes, as is seen in the polyphenism of eusocial species and in the radical sexual dimorphism observed in many insect orders. Above the species level, two taxa may exhibit minimal differences at the nucleotide level yet show extreme phenotypic divergence. King & Wilson (1975) first pointed out that the basis of many phenotypic changes might not be harboured in simple point mutations in structural proteins, but rather centred in regulatory gene expression. Recently this phenomenon has been invoked in insect evolution (Catalan *et al.*, 2012; Hardison & Taylor, 2012; Saminadin-Peter *et al.*, 2012; Glaser-Schmitt *et al.*, 2013), and a full understanding of the striking phenotypic evolution observed in insects will require both genomic and transcriptomic approaches.

In the very recent past comparative studies of gene expression were limited to a few organisms, but the advent of next-generation sequencing, combined with continuous progress in RNA sequencing (RNA-Seq) data quantity and quality (Picelli *et al.*, 2013), have democratized the technology. It is now possible for researchers outside the model organism canon to use comparative transcriptomics, and studies of a variety of insect traits have recently benefited from this approach. Because neither a reference genome nor extensive knowledge of the biology of the targeted insect are needed to accomplish a detailed transcriptome analysis, numerous and phylogenetically diverse insect species have been examined in the past decade. By tracking the expression

of genes across several species or under various conditions and developmental stages, entomologists have gained valuable insight into the evolution and molecular basis of insect development, physiology and behaviour.

Many of the earliest insect RNA-Seq studies involved simple annotation analysis to categorize the genes expressed in a single tissue in one species. As methods of assembly, annotation and expression quantitation have improved, researchers are increasingly using a comparative approach (among species, sexes, tissue types, life stages, etc.) to isolate the transcriptional basis of phenotypic diversity. In the present review, we examine the areas of insect biology employing this approach, with particular focus on studies highlighting the dynamic nature of transcriptome evolution through rapid changes in gene expression and incorporation of new genes into transcriptional profiles. We include all insect species examined to date, with the exception of the 'supermodel' species *Drosophila melanogaster*, whose transcriptomics are discussed in detail elsewhere (Graveley *et al*., 2011; and summary databases at http://flybase.org/static_pages/feature/previous/articles/2010_03/RNA-seq_data.html). Important technical aspects of the RNA-Seq method (Box 1) have been discussed in detail in several reviews (Wang *et al*., 2009; Martin and Wang 2011; Wolf, 2013; etc.), and there are some pitfalls to the approach that must be considered (Box 2). Overall, however, this approach has provided a broadening of opportunities for research at the level of gene expression in an increasingly diverse group of non-model insects.

RNA-Seq applied to insect biology

RNA-Seq (also known as whole-transcriptome shotgun sequencing) technology now permeates numerous fields in evolutionary biology and has been used to address questions concerning differential gene expression, phylogenetic relationships, gene creation and gene family evolution, rates of protein evolution, genotype–phenotype association, chromosome organization and the regulation of development. To quantify the escalation of studies examining insect transcriptomes we conducted a literature search using the terms 'expressed sequence tags' ('EST'), 'microarray', 'RNA-Seq' and 'transcriptome'. We excluded marker-based analyses such as phylogenomic and population genomic studies from our survey and instead focused on research that examines changes in transcriptional profiles associated with phenotypic differences. This search revealed that most branches of insect diversity research now use transcriptome analysis (Fig. 1). As shown in Fig. 2A, RNA-Seq technology has quickly supplanted previous genomic approaches (e.g. EST sequencing and microarray cDNA hybridization) as the preferred method to examine gene expression patterns.

The broad utility of RNA-Seq is illustrated by the diversity of insect groups and research questions addressed with this methodology (Fig. 2B). Much of the insect research using RNA-Seq reflects an applied focus on traits of importance to human health or agriculture (e.g. blood meal dynamics, vector-borne infection and resistance to pesticides), but other areas in which RNA-Seq is popular involve a more basic understanding of the evolutionary and behavioural biology of insects (e.g. caste differentiation, gregariousness, olfaction and diapause). We review recent work in these fields by focusing on three critical aspects of insects' evolutionary success: morphological development and plasticity, physiological response to the environment, and sexual dimorphism.

Development and plasticity

Morphological development

Even insects with apparently simple hemimetabolous development undergo striking changes as they mature from egg to adult, and recent studies have compared *de novo* assembled transcriptomes across different developmental stages to identify the genes and signalling pathways involved (Chen *et al*., 2010; Ewen-Campen *et al*., 2011; Zeng *et al*., 2013). Most studies to date have involved the Holometabola. Zheng *et al*. (2012) compared gene expression between larval and pulpal stages of the tephritid fruit fly *Bactocerca dorsalis* and identified a large number of differentially expressed genes, most involved in metabolic pathways (e.g. oxidative phosphorylation and glycolysis). Li *et al*. (2013) compared the transcriptomes of four developmental stages (embryos, larvae, pupae and adults) of the lepidopteran maize pest *Athetis lepigone* and found that most differentially expressed genes were downregulated in larvae and upregulated in pupae. These expression differences were consistent with characteristic features of each life stage; for example, cuticular protein genes were upregulated in larvae and downregulated in pupae. In the only study we are aware of involving hemimetabolous insect development, Bao *et al*. (2013) conducted sex- and stage-specific transcriptome-wide analysis of immune function genes in the brown planthopper, *Nilaparvata lugens*, and found that protein families that sense foreign microbial infection (e.g. Peptidoglycan recognition proteins), mediate innate immunity (e.g. clip-domain serine proteases), and immune response effector genes, had relatively high expression levels in nymphs and extremely low levels in eggs. In contrast, some genes contributing to antibacterial defence (e.g. i-type lysozymes) and regulation of immune-response effectors (e.g. Toll genes) had their

Box 1. Next-generation sequencing RNA sequencing primer

Platform

There are several platforms that are categorized as next-generation sequencing (NGS; aka massively parallel sequencing methods), but the Illumina platform is the most widely used and flexible for RNA sequencing (RNA-Seq) studies (Van Verk *et al*., 2013). Other methods such as 454, PacBio, IonTorrent and SOLID platforms are also available. In RNA-Seq studies, the RNA from a single cell, tissue or whole animal (as with small insects) is isolated and purified by taking advantage of the polyA tail on the mRNA, which is present in most translated transcripts. This step separates the RNAs that will eventually be translated into proteins from nontranslated cellular RNAs (e.g. ribosomal RNAs). The purified RNA is then made into cDNA, and this cDNA is amplified using standard PCR approaches. The PCR product is fragmented to a specific length that is dependent on the platform used (Illumina, 454, etc.) and the size selected cDNA is converted into a 'library' by the addition of platform-specific adaptors to the ends of the cDNA.

The Illumina method has several platforms with different capacities. For instance, the miSEQ platform generates paired end fragment size reads (currently 300 base pairs each read), and 25 000 000 reads per run, while the Illumina HiSEQ2500 can generate paired end reads of 125 bases each but generates 4 billion reads. Both the Illumina and 454 methods allow samples to be multiplexed, meaning that several tissues or tissues from several organisms can be run at the same time. In this approach, the target cDNA is amplified with primers that harbour specific DNA sequences called 'barcodes'. The sequences from the various mRNA samples from different developmental stages, tissues or even individuals can then be sorted using informatics techniques. In this review we focus mostly on studies where NGS approaches have been used to address questions about insect biology, though some previous studies using expressed sequence tag analysis and microarrays are also discussed.

Assembly and Annotation

Assembly refers to the process of piecing the small NGS sequence reads back together to get full-length reads for the different RNAs present in the tissue. Assembling short reads (<200 base pairs) is in general more difficult than assembling longer reads unless a decent scaffold sequence is available. Thus, model organisms whose genomes and transcriptomes are already well characterized are much easier to work with, especially when short read lengths are used. Commonly used assembly programs include Trinity (Grabherr *et al*. 2011), Trans-ABySS (Birol *et al*., 2009), Velvet-Oases (Schulz *et al*., 2012), and SOAPdenovo (Xie *et al*., 2014). Another advance in the assembly area are the occurrence of 'Assemblathons' (https://genome10k.soe.ucsc.edu/assemblathon; Bradnam *et al*., 2013). These are semi-competitions where different informatics teams in the field use their software to assemble a target genome provided by the organizers of the assemblathon, with each participant having the same amount of time to do the assembly. These assemblathons should be consulted regularly for advances that occur as a result of these competitions (Bradnam *et al*., 2013).

Quantitation of transcript abundance

Knowing what transcripts are present in a tissue or cell type supplies only one dimension of data for an RNA-Seq study. Additional information can be obtained by looking at the relative abundance of different transcripts to generate an 'expression profile' for different genes or gene categories. Phylogenetic studies usually ignore transcript abundance and focus on the identification and comparison of orthologous sequences present in different taxa. Other types of studies explicitly compare expression profiles as well as enumerating which transcripts are present. Most traits related to insect biology are potentially affected by variation in expression levels, and studies of such traits require fairly accurate estimation of abundance of transcripts. An updated list of read mapping software is provided by http://wwwdev.ebi.ac.uk/fg/hts_mappers/.

highest expression level in eggs, suggesting that they may have roles in embryogenesis and development as well as in immune function.

Phenotypic plasticity

All insects must complete development to adulthood, but in some species the precise outcome of this developmental trajectory is flexible. Such phenotypic plasticity – the ability of a single genotype to give rise to a range of phenotypes – is an obvious candidate for transcriptomic approaches, because genomics alone cannot explain how various phenotypes are generated. Eusocial insects in the orders Hymenoptera, Isoptera and Hemiptera offer an extreme version of plasticity, known as polyphenism, in

Box 2. Pitfalls of RNA-Seq

Non-Model Problem: Applying RNA-Seq methods to non-model organisms poses specific problems with assembly that are not faced in model organisms. Because assembly must be conducted *de novo* rather than based on an existing reference, specialized methods are required (Grabherr *et al*., 2011; Martin & Wang, 2011). While these issues are less extreme with longer reads (because a sufficiently long read will cover the entire transcript, eliminating the need for assembly), it is currently impossible in most non-model studies to know whether the 'right' assembly has been found. The overall quality of an assembly is assessed by looking at the number and length of the transcripts identified, and any sequencing errors that appear in the final assembly can be corrected. Following assembly, an attempt is made to determine the identity and function of the final transcripts. This annotation relies on finding orthology between the assembled transcript sequence and sequences previously characterized in other organisms. Orthology determination may rely on sequence similarity, predicted structural similarity between proteins, or other traits, based on Bayesian statistics and phylogenetic tree approaches. Transcripts that cannot be annotated (i.e. those with no significant similarity to any previously identified sequence) are characterized as novel, and may be unique to the taxa being investigated.

Because well-annotated genomes and transcriptomes are available for so few non-model organisms, researchers will often find that 50 to 80% of the transcripts they identify are unannotated (Riesgo *et al*. 2012). Until additional resources are available, it is difficult to know what role, if any, these transcripts might play in the trait under study. While this problem can be avoided by using a candidate gene approach (which ensures, *ipso facto*, that no unknown transcripts will be found), candidate gene studies run the risk of finding a 'shared' genetic basis for traits in different taxa because they only look at a small fraction of the transcripts present. While unbiased whole transcriptome studies are challenging, their potential for uncovering new molecular pathways and processes make them the preferred approach.

Transcript Length and Source Problems: There are several problems that researchers need to consider in this context. Because many of the current methods are dependent on a manipulation of target mRNA (such as cDNA synthesis, amplification steps and ligation), accurate downstream quantitation of transcript abundance can be problematic. Another problem is that transcript length bias can occur because of the high degree of fragmentation and size selection of transcripts or cDNA fragments used in RNA-Seq methods. Yet another concern involves depth of coverage in an RNA-Seq study: rare or infrequently expressed transcripts are difficult to detect, and at low to moderate coverage levels it is impossible to know whether the absence of a transcript results from its absence in the genome, or from its rarity/low level of expression (Tarazona *et al*., 2011; Maza *et al*., 2013). The tissue chosen for comparison is also a critical aspect of RNA-Seq studies; for instance, if one is comparing expression of digestive transcripts between organisms, it is critical that only digestive tract tissue be sampled. Even with appropriate tissue sampling, highly composite parts of the anatomy such as the abdomen (which may contain multiple specialized structures) can result in many false-positives (Johnson *et al*., 2013), so great care must be taken in extracting and processing the tissue of interests. For this reason, RNA-Seq studies that seek to characterize transcriptome abundance should not use composite structures or whole insects as a tissue source (though for studies examining transcript *presence* rather than abundance, these tissue sources are acceptable).

Quantitating Abundance Problems: Another problem that arises when estimating abundance of transcripts involves isoforms (alternative versions of a specific transcript) and spliced transcripts. Identification of isoforms is sometimes difficult, and the validation of the existence of an alternative transcript produced by splicing is also a shortcoming in RNA-Seq studies. A promising alternative to cDNA RNA-Seq methods is to directly sequence RNA (DRS) for a sample (Ozsolak & Milos, 2011; Mannello *et al*., 2012). This approach requires single molecule sequencing approaches, which are currently on the horizon, or the development of direct RNA amplification without cDNA synthesis steps.

which a single genotype is responsible for morphologically and/or behaviourally distinct castes of workers and reproductives. Only slightly less striking is the ability of some Orthopterans to switch between two phases, solitary and gregarious, which differ in morphology, behaviour and physiology. In many cases, the proximal causes of phenotype determination are well established [e.g. diet-induced caste determination in honey bees (Shuel & Dixon, 1960) and pheromone-induced caste determination in termites (Wilson, 1971)], but only recently have the necessary tools emerged to allow for exploration of the gene expression patterns that underlie these phenotypes.

Polyphenism

Studies of eusociality in Hymenoptera abound (Amdam *et al*., 2004; Patel *et al*., 2007; Toth & Robinson, 2007; Boomsma, 2009; Linksvayer & Wade, 2009; Johnson &

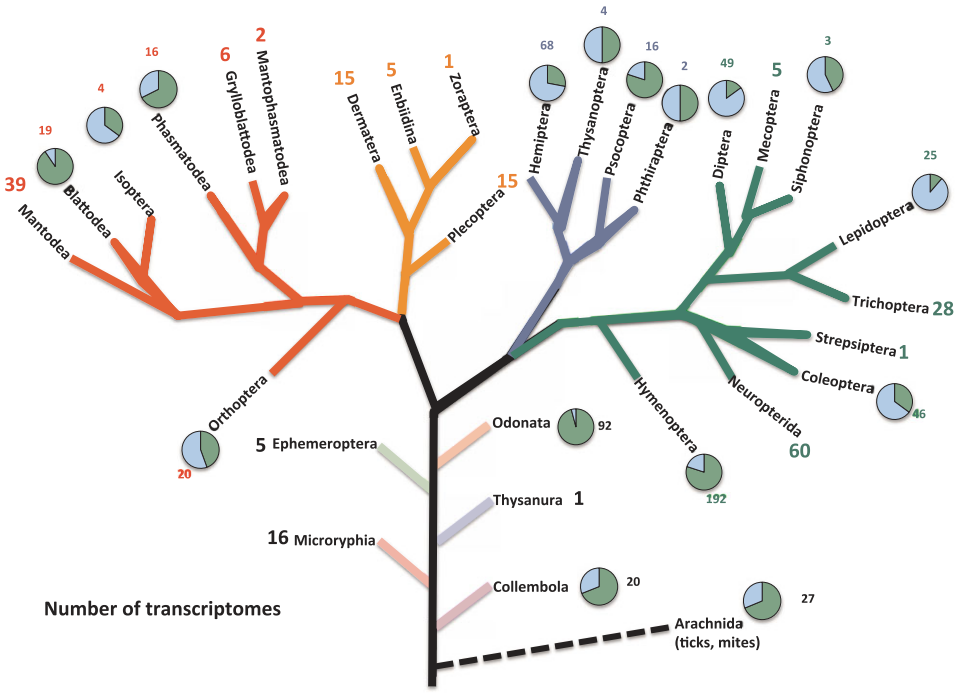

Figure 1. Phylogenetic tree showing the broad phylogenetic distribution of insect transcriptomes. Tree modified after Bugwood (http://wiki.bugwood.org/Main_Page). The number of transcriptomes for each insect order was compiled from the literature. Each order with transcriptomes generated independent of 1KITE (http://www.1kite.org) and surveyed as part of our literature search (key words: 'transcriptome', 'RNA-Seq', 'microarray', or '454'; November 2013) has a pie graph associated with it. The green (dark) section of the graph represents the number of species with transcriptomes generated for that order by 1KITE (http://www.1kite.org/downloads/1KITE_species.txt; November 2013). The blue (light) section represents the number of taxa with transcriptional studies done independently of the 1KITE initiative. The number in small font indicates the number of taxa slated for study by the 1KITE initiative. The number in large font represents the number of 1KITE species examined or slated for examination in a given order, where no studies independent of 1KITE have been done.

Linksvayer, 2010; Toth *et al*., 2010; Daugherty *et al*., 2011; Johnson & Tsutsui, 2011; Le Conte *et al*., 2011; Ometto *et al*., 2011; Woodard *et al*., 2011; Greenberg *et al*., 2012; Huang *et al*., 2012; Liang *et al*., 2012; Baek *et al*., 2013; Ferreira *et al*., 2013). In the honey bee, *Apis mellifera*, research on the basis of polyphenism is well advanced, and has been reviewed elsewhere (Evans & Wheeler, 2001; Smith *et al*., 2008; Simpson *et al*., 2011). We touch here only on more recent studies using transcriptomic approaches, which have confirmed existing knowledge of the genes involved in eusociality and also identified novel genes and mechanisms. For instance, Simola *et al*. (2013) conducted genome comparisons of eusocial and solitary species to examine the importance of regulatory vs. compositional changes in the evolution of eusociality. They found nearly 2000 genes (largely involved in neuronal and hormonal functions) with similar regulatory changes in all eusocial vs. solitary insects, and identified specific transcription factors (empty spiracles and cAMP response element-binding protein) that appear to be involved in the evolution of eusociality. In another study, Ferreira *et al*. (2013) investigated whether expression changes in a conserved toolkit of genes (e.g. vitellogenin, insulin, major royal jelly protein, juvenile hormone (JH), methyltransferases, hexamerin, Malvolio, Amfor and P450) were responsible for caste evolution across social taxa, as suggested by some investigators (Toth & Robinson, 2007; Johnson & Linksvayer, 2010). Using a candidate gene approach, previous studies have concluded that these toolkit genes and processes share conserved roles in caste regulation in many eusocial taxa (Amdam *et al*., 2004; Toth *et al*., 2010; Daugherty *et al*., 2011; Woodard *et al*., 2011; Huang *et al*., 2012).

Unbiased surveys of entire transcriptomes of social insects are still rare, however, and experiments using a candidate gene approach are inherently unlikely to find novel explanations for the investigated phenotypes. Thus, some workers have suggested that many of the molecular processes responsible for polyphenism in different social lineages will differ substantially, because extensive molecular and developmental re-wiring may be required in the evolution of caste commitment and eusociality (Boomsma, 2009; Linksvayer & Wade, 2009), and that each lineage may have a unique set of genes responsible for polyphenism. In the case of the primitively eusocial wasp *Polistes canadensis*, Ferreira *et al*. (2013) identified 2442 genes with caste-biased expression levels. Seventy-five percent of these were novel genes, and there was

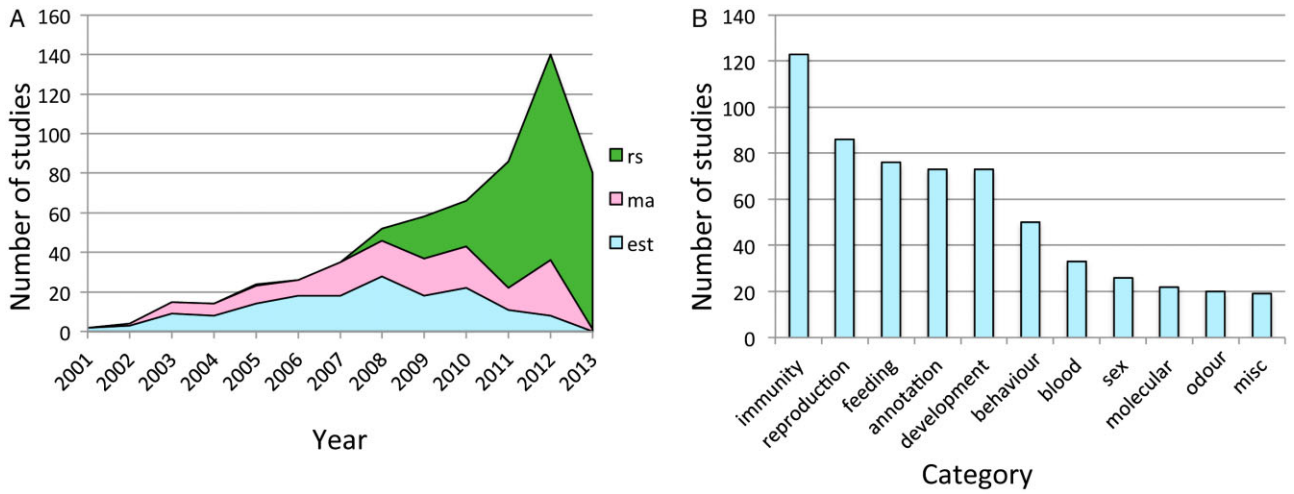

Figure 2. (A) Compilation of studies using transcriptomics in non-model (i.e. non *Droosophila melanogaster*) insects. The graph is a cumulative count of number of insect transcriptomes analyzed by year since 2001. Blue indicates clone and sequence expressed sequence tag studies, red indicates microarray studies and green inidicates the number of RNA-Seq studies done with next generation sequencing. (B) Bar graph showing the number of insect transcriptomes analyzed and the category of research the transcriptome was generated for.

very little overlap in the identity and expression pattern of caste-biased genes from *P. canadensis* and those in other social hymenopterans. Only 6.5% of honey bee caste-biased genes were also caste-biased in *P. canadensis*, and there was no correlation with caste-biased genes of the eusocial fire ant, *Solenopsis invicta*.

Isoptera (termites) have also been the focus of research on phenotypic plasticity with respect to social behaviour (Korb *et al*., 2009; Husseneder *et al*., 2012; Tarver *et al*., 2012; Sen *et al*., 2013). Termites exhibit JH-dependent caste differentiation, and Tarver *et al*. (2012) examined the role of a termite-specific cytochrome P450 gene (Cyp15F1) in this process. Genes in the cytochrome P450 family 15 are known to play a role in JH biosynthesis and degradation, and elevated JH titres in *Reticulitermes flavipes* workers induce differentiation into the soldier caste via the intermediate presoldier stage. Tarver *et al*. (2012) tested whether expression levels of Cyp15F1 change in response to primer pheromones derived from the heads of soldiers. They found that expression levels of Cyp15F1 were elevated by one soldier pheromone (CAD), and that RNA interference (RNAi) against Cyp15F1 reduced presoldier formation by 78% compared with the no-RNAi control. The authors conclude that primer pheromones, acting as socio-environmental cues, directly influence Cyp15F1 expression and help regulate caste differentiation by mediating JH signalling.

Phase switching

Phase change in locusts can be induced rapidly and is fully reversible. The solitary phase, in which locusts actively avoid encountering each other and are cryptically coloured, may persist for many generations. When population densities increase to the point where avoidance is impossible, a shift to the gregarious phase, in which locusts are brightly coloured and can accumulate into devastating swarms, is triggered. While the complete shift from one phase to another – which involves changes in behaviour, colour, morphology, life history and physiology (Uvarov, 1966) – takes several generations, behavioural changes are evident within hours of the phase change-inducing stimulus (Roessingh *et al*., 1998). Although the number and function of the genes involved in this transition are not yet thoroughly known, two recent studies detected similar numbers of transcripts with phase-related expression differences: Badisco *et al*. (2011) conducted a microarray-based comparison of genes expressed in the central nervous system and found 214 genes (out of 20 755) differentially expressed in solitary and gregarious phases, while Chen *et al*. (2010), using a *de novo* approach, found 242 (out of 11 490) genes with expression differences. Wang *et al*. (2014) examined methylomes (i.e. DNA methylation patterns), brain transcriptomes and alternative pre-messenger RNA splice transcripts of solitary and gregarious locusts. Many of the genes that displayed differences are involved in the regulation of cytoskeletal microtubules (essential neuronal structures involved in neuronal polarization, development and plasticity). The authors conclude that the abundant differences between solitary and gregarious locusts in expression levels, methylation states and splice forms contribute to the neuronal plasticity that underlies the behavioural changes required for phase switching.

Consistent with the need to respond to the diverse sensory inputs inherent in swarming and migration, genes

involved in environmental interaction are upregulated in gregarious locusts (Chen *et al*., 2010; Badisco *et al*., 2011; Guo *et al*., 2011). Such genes include those associated with the regulation and development of neuronal structures; regulation of neurotransmitter activities; and detection of chemosensory cues. High expression levels in these categories might help gregarious locusts deal with sensory complexity, because enhanced nervous system plasticity should facilitate rapid adaptation to new environments and stimuli. Acute stress genes [HSPs and pathogen resistance proteins (Badisco *et al*., 2011)] are also upregulated in gregarious locusts, possibly mitigating the increase in disease risk associated with swarm life. In contrast, 'anti-aging' genes, including those associated with oxidative stress resistance, detoxification and anabolic renewal (Badisco *et al*., 2011) and biosynthetic pathways (Chen *et al*., 2010) are downregulated in gregarious locusts, which may represent a trade-off to free metabolic resources for other functions at the cost of a reduction in lifespan (Uvarov, 1966).

Physiological response to the environment

Response to environmental stress and infection

Insects can respond to environmental stress in many ways. Stressors may be predictable (e.g. seasonal changes in climate, bacteria associated with specific environments) or unpredictable (e.g. unseasonable temperatures, viral infection), and some may be constant (e.g. freezing temperatures in polar environments); thus, the particular environment of each species will determine the genes that are differentially expressed in response to stress. Several studies have explored changes in gene regulation in response to cold stress (Timmermans *et al*., 2009; Teets *et al*., 2012; Dunning *et al*., 2013; Teets & Denlinger, 2013). Teets *et al*. (2012) examined gene expression changes in response to extreme cold in the Antarctic midge, *Belgica antarctica*, which is the world's southernmost insect and the only insect known to be endemic to Antarctica. They found upregulation of cellular recycling pathways and coordinated upregulation of numerous genes involved in autophagy (Teets & Denlinger, 2013), suggesting that autophagy and related processes are important aspects of stress tolerance in the Antarctic environment. The authors compared these results with data from the Arctic collembolan, *Megaphorura arctica*, and found very little overlap in expression profiles, indicating that these polar species have developed different gene expression mechanisms to withstand similar desiccating conditions. High temperatures can be as dangerous as low ones, and recent transcriptome studies have shown that, in spite of their name, antifreeze proteins play a role in tolerance of both high and low temperatures (Ma *et al*., 2012; Qiu *et al*., 2013).

In predictable environments, dormancy for part of the year might be a reasonable response to environmental stress. Diapause and quiescence are both forms of dormancy that allow insects to survive stressful conditions, but they are distinct in that diapause involves a preprogrammed developmental arrest that is anticipatory in nature, while quiescence is characterized by an immediate developmental arrest in response to unfavourable conditions. These processes have been examined by comparing the transcriptomes of diapausing and quiescent organisms (Poelchau *et al*., 2013b), and of developmental stages in different phases of diapause (Poelchau *et al*., 2013a). For instance, Gong *et al*. (2013) compared the transcriptomes of the diapausing and non-diapausing orange wheat blossom midge, *Sitodiplosis mosellana*, an economically important wheat pest. In diapausing individuals, Gong *et al*. found upregulation of heat shock proteins, carbohydrate metabolism genes, and genes involved in the suppression of JH signalling and response.

The expression of immune response genes (i.e. those transcribed in response to infection by viruses, bacteria or parasites) has been examined in several species (Table S1). In the harlequin ladybird beetle, *Harmomia axyridis*, a widespread invasive species that has acquired strong resistance to bacterial and fungal infection, Vilcinskas *et al*. (2013) compared transcriptome profiles before and after infection. Their findings indicate that immunity evolved through the rapid expansion of genes coding for antimicrobial peptides and proteins. Similarly rapid evolution of immune response genes was also found in studies of *Nasonia vitripennis*. Sackton *et al*. (2013) found that many novel (i.e. found only *N. vitripennis*) immune genes were induced by infection, producing a species-specific immune system for this wasp.

Blood meal dynamics

Because blood-feeding mosquitos serve as vectors for malaria, West Nile Virus, and other diseases, the transcriptional regulation of blood feeding is of particular interest to human health. Two recent studies examined the role of olfaction in blood meal timing in *Anopheles gambiae*: Rinker *et al*. (2013) found that moderate changes in odorant receptor (*Or*) transcript abundance can modulate oviposition behaviours after a blood meal, and Rund *et al*. (2013) identified well-defined cyclic expression patterns of *Or*s and odorant binding proteins, a family of soluble proteins found in insect olfactory organs. Given that the peaks in expression correspond with times of increased blood-feeding behaviour, such cyclic expression may be responsible for coordinating the olfactory system with *An. gambiae*'s pre-dusk/dusk circadian niche. Researchers seeking targets for control measures against *Aedes aegypti*, the principal mosquito vector of dengue viruses,

have examined fat body transcriptomes in blood-fed and non-blood-fed *Ae. aegypti* (Price *et al*., 2011), while other studies characterizing blood feeding have examined salivary gland and midgut tissues (Table S1).

Response to plant defence and pesticides

Although the ability of herbivores to cope with naturally occurring plant defences and the evolution of resistance to synthetic insecticides might seem like separate topics, recent research using RNA-Seq has revealed a surprising amount of overlap between them (Bass *et al*., 2013; Dermauw *et al*., 2013). For example, the spider mite *Tetranychus urticae* is polyphagous, and many strains have evolved resistance to all insecticides. To investigate whether the genes that allow *T. urticae* to feed on many chemically defended hosts are also involved in its ability to rapidly evolve pesticide resistance, Dermauw *et al*. (2013) compared transcriptomes from two different insecticide-resistant strains to those of a strain selected for five generations on tomato (a challenging host for *T. urticae*). They found >40% of the differentially regulated genes were shared by all three groups, and prominent amongst these were some 'usual suspects' commonly implicated in detoxification or transport of xenobiotics (e.g. cytochrome P450s, carboxyl/cholinesterases, and glutathione-S-transferases), although genes of unknown function had some of the most marked expression changes. A more striking result, however, was the detection of expression differences in some previously unexplored gene categories, namely binding proteins (e.g. lipocalins) and transport proteins (e.g. the major facilitator superfamily). While the role of these proteins in polyphagy and insecticide resistance is not yet known, their detection – which would have been impossible using a candidate gene based approach – illustrates the value of unbiased whole transcriptome comparisons.

Based on such results, it seems likely that many of the studies that have looked only at the transcriptional response to either naturally occurring plant compounds or only at the response to insecticides will in fact shed light on the reciprocal question. By characterizing differential gene regulation in pest insects in response to plant defences, new targets for pest management could emerge. Many recent studies have examined gene expression changes associated with feeding on well-defended host plants. Chi *et al*. (2009) found that feeding on defended soybean is energetically costly for the cowpea bruchid, *Callosobruchus maculatus*: the expression of structural, defence and stress-related genes was much lower on defended plants, suggesting that insects reallocate resources to deal with plant compounds at a cost to their ability to respond to other potential challenges. In the case of the whitefly, *Bemisia tabaci*, however, Alon *et al*. (2012) found that expression changes in response to phenylpropanoids (a large group of plant secondary defence metabolites) do not appear to be costly. They examined reproductive performance and transcriptome profiles after feeding on phenylpropanoid-enriched tobacco and concluded that the resulting elevated expression of genes involved in metabolism, protein synthesis and defence was not accompanied by reduced reproductive performance.

The response of the fall armyworm, *Spodoptera frugiperda*, to toxins present in maize was examined by Fescemyer *et al*. (2013). They found upregulation of genes involved in the production of midgut constituents (to counteract the peritrophic membrane depletion the maize toxins cause) and of genes involved in the production of digestive enzymes (to help clear the toxins more rapidly). Similarly, Asian rice gall midges, *Orseolia oryzae*, feeding on resistant vs. susceptible rice varieties upregulate genes for proteolysis and protein phosphorylation, possibly in an attempt to break down the toxic proteins more quickly (Sinha *et al*., 2012).

Some researchers have used the transcriptional response of herbivores to a novel or recently adopted host as a proxy for the evolution of host plant range, hypothesizing that the genes expressed might encompass those involved in historical transitions to new hosts (Liu *et al*., 2010; Matzkin, 2012; Whiteman *et al*., 2012; Celorio-Mancera *et al*., 2013; Luan *et al*., 2013). These studies involve the examination of experimentally manipulated insects using specific plants as hosts. For instance, (Matzkin, 2012) looked at host range evolution in the cactophilic fruit fly, *Drosophila mojavensis*, which exhibits population-level specialization on different cactus species. Matzkin found that >20% of the genome was differentially regulated in response to a challenging cactus species. Genes associated with mRNA binding and toxin response were upregulated, while genes associated with energy and carbohydrate metabolism and structural components were downregulated. In another study, Celorio-Mancera *et al*. (2013) examined the evolution of host range in the butterfly *Polygonia c-album*, a polyphagous, widely distributed butterfly. They compared the transcriptomes of larvae feeding on an established host plant (*Urtica dioica*) vs. one that represents a recent evolutionary shift onto a very divergent host (*Ribes uva-crispa*). Like Dermauw *et al*. (2013) they found unexpectedly high levels of differential regulation of previously unsuspected proteins, namely genes coding for transporter and membrane-binding proteins (including major facilitator superfamily and monocarboxylate types), in response to the newer host. In addition, they found upregulation of cuticular proteins, which they suggest may strengthen the insect gut and prevent toxic plant compounds from entering the insect body.

Sexual dimorphism

Sexual selection is the driving force responsible for a vast array of novel morphological and behavioural phenotypes found throughout insects (Wilkinson & Dodson, 1997; Emlen, 2008; Allen *et al*., 2011). The extent of sexual dimorphism ranges from minor differences in somatic structures to dramatic showy or exaggerated characters – such as beetle horns, butterfly wing patterning and cricket calls – involved in competition, sexual advertisement and mate selection. Despite the widespread study of sexual dimorphism, we still know little about how variation at the molecular level controls this morphological diversity. A fundamental aspect of the genetic basis of sexual differences is that males and females have virtually identical genomes. Other than the few genes located on the sex-limited copy of the heteromorphic sex chromosomes (e.g. the Y chromosome in XY sex determination systems), there are no genes that exist exclusively in only one sex. As a result, sexual dimorphism is shaped primarily by differential patterns of gene expression. Research examining the specifics of sexual differences in gene expression has flourished recently (for reviews, see Ellegren & Parsch, 2007; Parsch & Ellegren, 2013), motivated by the advent of next-generation genomic techniques and an interest in a number of different evolutionary phenomena, including: the resolution of sexually antagonistic selection pressures; the molecular evolution of reproductive proteins; the chromosomal distribution of sex-biased genes; and the origin and evolution of dosage compensation. The vast majority of this research has focused on *Drosophila*, but RNA-Seq offers the opportunity to explore these topics in other insect systems distinguished by dramatic sexually dimorphic phenotypes. Below we summarize the details associated with much of this research and highlight some recent studies on non-*Drosophila* systems.

Genetic basis of sexual dimorphism

Sexually dimorphic traits arise because different selective pressures affect male and female phenotypes (Lande, 1980). Differences in gene expression between the sexes (termed sex-biased or sex-specific gene expression) evolve to resolve the conflict between male- and female-specific adaptive peaks so that each sex can approach its own adaptive optima (Parsch & Ellegren, 2013); therefore, cataloguing genes that are over-expressed in one sex relative to the other in sexually dimorphic tissue can help identify candidate genes and pathways responsible for sexually dimorphic development. Studies in *Drosophila* have shown that sex-biased gene expression is pervasive and evolutionarily dynamic (Meiklejohn *et al*., 2003; Ranz *et al*., 2003; Zhang *et al*., 2007; Assis *et al*., 2012). Depending on the study, up to 50% of the genes examined exhibit differential expression between males and females. Substantial differences in the patterns of sex-biased expression are also found between closely related species, particularly for genes serving male reproductive functions (Meiklejohn *et al*., 2003; Ranz *et al*., 2003; Zhang *et al*., 2007; Assis *et al*., 2012). These studies, however, have focused primarily on whole-body gene expression where much of the sex-biased expression results from reproductive tissues. Few studies, even in *Drosophila*, have examined sex-biased expression in specific somatic tissues (but see Chang *et al*., 2011), and fewer still in a somatic tissue with respect to a specific sexually dimorphic phenotype in an attempt to associate expression differences with phenotypic variation.

In non-*Drosophila* insects, we are aware of only two studies comparing male and female expression levels at a genomic scale in specific (not whole-body) somatic tissue. Baker *et al*. (2011) examined sex-biased gene expression in several tissues of the mosquito, *An. gambiae*, and found that, depending on the tissue, 5–15% of the genome was differentially expressed between males and females. A more directed approach was taken by Wilkinson *et al*. (2013) in an attempt to understand the expression differences underlying extreme sexual dimorphism in the eyestalks of stalk-eyed flies (Diopsidae). Using microarrays designed from an EST library (Baker *et al*., 2009), they identified nearly 900 genes that were sex-biased in the developing eyestalk tissue but not in the adult head. As might be expected because males possess a larger eyestalk, male-biased genes were over-represented for genes involved in growth and mitochondrial function. Results from both of these studies suggest that transcriptome analysis provides a valuable tool for understanding the overall architecture of sexual dimorphism, but, because sex-biased expression is so widespread, such analyses are unlikely to directly identify candidate genes. It is also important to note that the causal loci directly responsible for sexually dimorphism may not exhibit any expression differences between the sexes, but rather exert their influence through cis-regulatory effects or alternative splicing of pre-mRNA. It will be critical for future studies to identify essential upstream regulators of sexual dimorphism and exaggerated traits, such as *doublesex* and *Insulin-like receptor* (Emlen *et al*., 2012; Gotoh *et al*., 2014), and to examine the transcriptional impact of these genes through functional manipulation (e.g. RNAi).

Evolution of reproductive proteins

Genes involved in mating and reproduction, particularly for males, are among the most rapidly evolving loci in the genome, and often show evidence of positive selection (Haerty *et al*., 2007; Singh & Artieri, 2010). Because of

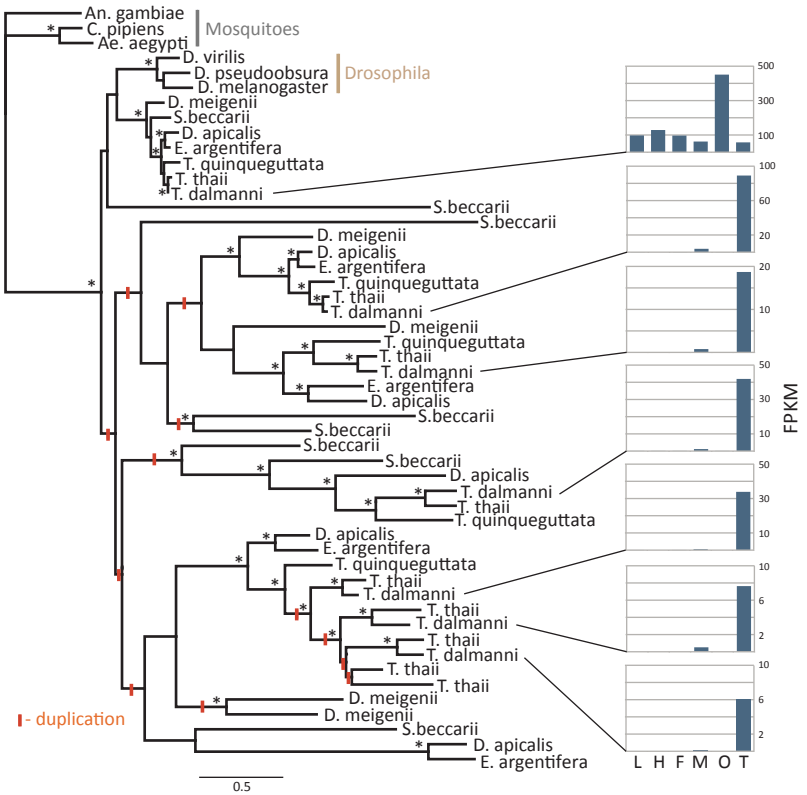

Figure 3. Phylogeny of the *Nap1* gene family depicting abundant gene expansion within stalk-eyed flies. Contig sequences homologous to *Drosophila Nap1* were extracted from an assembly of a testes transcriptome for each stalk-eyed fly species except *Teleopsis dalmanni* whose sequence data was derived from a multi-tissue assembly (GenBank accession numbers: KM821168-KM821212). The maximum likelihood tree was generated in PhyML (Guindon and Gascuel 2003) using an LG + G model with 100 bootstrap replicates. Branches with bootstrap values >90 are indicated by an asterisk. Red marks on the branches denote putative duplication events. Gene expression values (FPKM) for each *T. dalmanni* paralogue are presented to the right of the tree. Tissue abbreviations: L,larva; H, adult head; F, female carcass (whole body with reproductive tissues removed); M, male carcass, O, ovaries; T, testes. Stalk-eyed fly generic abbreviations: D. meigenii, *Diasemopsis meigenii*; D. apicalis, *Diopsis apicalis*; E., *Eurydiopsis*; S., *Sphyracephala*; T., *Teleopsis*.

their adaptive significance, reproductive proteins have been the subject of many evolutionary studies, and transcriptome analyses offer an efficient way of examining their molecular evolution. While few non-*Drosophila* insect studies have directly measured sex-biased gene expression on a genomic scale, examination of transcript diversity in a sexually dimorphic tissue for one sex is much more common. These studies include gene expression analysis of the sex pheromone glands in the moth *Heliothis virescens* (Vogel *et al*., 2010), the developing horn tissue in the beetle *Onthophagus taurus* (Choi *et al*., 2010), and the male reproductive organs of several insects – two *Heliconius* species (Walters & Harrison, 2010, 2011), *Ceratitis capitata* (Scolari *et al*., 2012), *Lutzomyia longipalis* (Azevedo *et al*., 2012), *Teleopsis dalmanni* (Baker *et al*., 2012; Reinhardt *et al*., 2014), *Teleogryllus oceanicus* (Bailey *et al*., 2013) and two *Gryllus* species (Andres *et al*., 2013). Protein sequences identified through transcriptome surveys can be compared with protein sequences from other species to test various hypotheses of molecular evolution. For instance, Walters and Harrison (2011) recently used a transcriptome analysis of seminal fluid proteins in *Heliconius* butterflies to examine the correlation between mating system and the rate of protein evolution. Contrary to expectation, they found that rates of protein evolution were not accelerated in species with promiscuous vs. monogamous female mating behaviour.

Another critical feature of transcriptome studies is that they facilitate the identification of lineage-specific genes that have no direct homologue in other insect groups. Novel genes can arise in many ways (e.g. gene duplication, lateral gene transfer, *de novo* gene creation from noncoding regions of DNA) and may rapidly acquire adaptive functional roles (Ranz & Parsch, 2012; Chen *et al*., 2013). Studies in *Drosophila* have shown that novel genes represent a substantial component of male-biased gene expression and play a critical role in spermatogenesis. A recent study in stalk-eyed flies (Baker *et al*., 2012) highlights the ability of RNA-Seq to uncover novel genetic diversity associated with reproductive proteins. More recently, testes transcriptomes from additional stalk-eyed fly species have been analysed, and numerous lineage-specific novel genes have been found (see Fig. 3 for an example). The gene family represented in Fig. 3 comprises homologues of the *D. melanogaster* gene, *Nucleosome assembly protein 1* (*Nap1*), an important histone chaperone gene that is expressed in the testes (Kimura, 2013). In *Drosophila* and the mosquito species, *Nap1* is present as a single homologue, while stalk-eyed fly species exhibit as many as nine paralogues of *Nap1*. Overall, the tree suggests there have been at least 14 duplications of the *Nap1* gene within the family, and it appears these new genes function primarily in spermatogenesis. Of the seven paralogous copies in *T. dalmanni*, the one with the closest homology to the *D. melanogaster*

Nap1 protein is expressed ubiquitously throughout the body, while the other six copies exhibit testes-specific gene expression (Fig. 3). DNA undergoes dramatic condensation during spermatogenesis involving the replacement of histones with protamines (White-Cooper, 2010), so it is possible these new Nap1 proteins play some role in this process. Overall, the use of transcriptomes for gene discovery has some inherent limitations because variation in expression levels across taxa will impact their representation in any gene family analysis. Even with this caveat, transcriptome surveys are an incredibly powerful tool for exploring the genetic diversity that is unique, and therefore of particular relevance, to non-model organism groups.

Evolution of sex chromosomes

Perhaps the most widespread use of sex transcriptome analysis involves addressing evolutionary questions regarding the origin and composition of sex chromosomes. Most insect species have a genetic sex determination system with one sex possessing heteromorphic sex chromosomes comprised of one gene-rich chromosome and one degenerate, heterochromatic, gene-poor chromosome (for reviews of the evolutionary processes driving this state see: Charlesworth, 1991; Charlesworth *et al*., 2005; Bachtrog, 2006). Which sex carries the heteromorphic pair varies among taxa (e.g. XY males/XX females in Diptera, ZZ males/ZW females in Lepidoptera), but in either case, the result is a dose imbalance for sex-limited genes in the heterogametic sex. To offset the potential fitness costs of having a single gene copy, dosage compensation mechanisms that balance gene expression levels across the genome are expected to evolve.

The abundant variation found among insects in both sex-determination systems and sex chromosome composition (Baker & Wilkinson, 2010; Gempe & Beye, 2011; Sahara *et al*., 2012; Vicoso & Bachtrog, 2013) has motivated a number of studies using RNA-Seq to examine the evolution of dosage compensation by comparing the expression levels of autosomal and sex-linked genes (Vicoso & Charlesworth, 2009; Prince *et al*., 2010; Harrison *et al*., 2012; Wilkinson *et al*., 2013). These studies have confirmed the observation that hypertranscription of the heterogametic sex chromosome occurs in XY systems (Prince *et al*., 2010; Wilkinson *et al*., 2013), but not ZW systems (Zha *et al*., 2009; Harrison *et al*., 2012). However, the mechanism of dosage compensation appears to differ between insect groups: beetles exhibit hypertranscription of the X chromosome in both sexes (Prince *et al*., 2010), while, in stalk-eyed flies, hypertranscription of the X chromosome appears to occur only in males (Wilkinson *et al*., 2013). In addition, ambiguity remains regarding the absence of dosage compensation in ZW species (Walters & Hardcastle, 2011). Transcriptome analyses in several more insect groups will be needed to map the evolutionary origin of dosage compensation systems and to understand the specific genetic mechanisms controlling dosage effects.

The existence of heteromorphic chromosomes affects the evolution of sex-biased gene expression well beyond the effects of dosage compensation. Sex-linked genes in species with heteromorphic chromosomes are found more often in the homogametic sex (e.g. the X chromosome spends 2/3 of its time in females in XY systems), a condition that has strong implications for the resolution of sexually antagonistic selective pressures (Rice, 1992; Bonduriansky & Chenoweth, 2009). Sex chromosomes are predicted to be a preferred location for sex-biased genes, and early transcriptome surveys of sex-biased gene expression in *Drosophila* revealed dramatic differences in the distribution of sex-biased genes, with the X chromosome significantly over-represented for female-biased genes (Parisi *et al*., 2003; Ranz *et al*., 2003; Sturgill *et al*., 2007). This research sparked an interest in the composition of sex chromosomes, and every study to date has found some form of differential representation of sex-biased genes on the sex chromosomes. For XY species, there is evidence of over-representation of female-biased genes on the X chromosome (Prince *et al*., 2010; Wilkinson *et al*., 2013) and under-representation of male-biased genes on the X chromosome (Magnusson *et al*., 2012). In contrast, the Z chromosome in the silk moth, *Bombyx mori*, is enriched for genes that function in the testes and is the source of excessive gene movement of female-biased genes onto autosomes (Arunkumar *et al*., 2009; Wang *et al*., 2012), while pea aphids (*Acyrthosiphon pisum*), which have an XO sex-determination system, have a highly masculinized X chromosome that is consistent with predictions based on the unusual mode of X chromosome inheritance in this species (Jaquiery *et al*., 2013). Given the rapid turnover in sex chromosome identity within insects (Kaiser & Bachtrog, 2010; Pease & Hahn, 2012; Vicoso & Bachtrog, 2013), the effect of chromosomal location on the pattern of sex-biased gene expression is likely to provide a major evolutionary force shaping the phenotypic differences between the sexes.

Conclusion

The democratization of RNA-Seq technology in terms of affordability, accessibility and interpretability, has paved the way to exciting progress in understanding the transcriptional basis of phenotypic diversity in non-model insects. Over the past few years, the number of published studies using this approach has more than doubled, allow-

ing researchers to address previously intractable questions. RNA-Seq has been especially useful in cases where extreme phenotypic divergence occurs between individuals with similar or identical genomes, such as sexual dimorphism and social polyphenism.

RNA-Seq has been used in a variety of ways in non-model insects. At its most fundamental level, RNA-Seq allows for the identification, annotation and comparison of the sets of genes expressed in different species, individuals or phenotypes. Such annotation-based studies have led to the identification of gene families that are evolving especially rapidly, and thus are plausible candidates for explaining the evolution of divergent phenotypes. Beyond cataloging the genes present in a transcriptome, RNA-Seq has further been used to compare the expression of genes across species, under various environmental conditions, and at different developmental stages. A third, and very recent, application of RNA-Seq is for examination of alternative pre-messenger RNA splicing. Alternate splice sites give rise to different protein isoforms, and these can radically alter protein function. In humans, recent studies suggest that >90% of all genes encode two or more proteins via alternative splicing of pre-messenger RNA (Wang *et al*., 2008), and in cases where both genetic sequence and gene expression levels are identical, it may be the case that alternative splicing is responsible for phenotypic divergence. While detailed surveys of the prevalence and impact of alternative splicing on insect phenotypes are still lacking, work on the *Dscam* gene, which plays a role in axonal guidance and can generate 38 016 distinct mRNA isoforms in *D. melanogaster* (Schmucker *et al*., 2000), suggests that hypervariable splicing of this gene is common throughout the Arthopoda (Armitage *et al*., 2012). In non-model insects, preliminary results from the tarnished plant bug (*Lygus lineolaris*) suggest that alternative splicing may contribute to the divergence of odorant binding proteins (Hull *et al*., 2014). Because odorant binding proteins mediate odorant detection and discrimination, alternative splicing of the genes coding for these proteins could affect oviposition, host plant use, mate preference and many other evolutionarily important behaviours.

One of the most striking results of the increasingly large number of *de novo* (rather than candidate-based) RNA-Seq studies published to date is the realization that many genes of unknown function are being expressed in virtually every species that has been examined. Because transcriptome annotation necessarily relies on sequence similarity to previously identified genes, the finding that 50% (or more) of the differentially expressed (or alternatively spliced) genes in one's study are unannotated means that it is difficult to know what role, if any, these genes have in explaining the phenotype of interest. Now that the most vexing aspects of RNA isolation, library preparation, sequencing, assembly and analysis are fairly well worked out, the next great challenge for those who study non-model organisms is to simplify the process of figuring out just what all those genes are for.

Acknowledgements

We thank an anonymous reviewer for constructive comments. We thank the Korein Foundation and the Lewis and Dorothy Cullman Foundation for their continued generous support. The Sackler Institute for Comparative Genomics at the American Museum of Natural History provided a vibrant community for discussion of the issues addressed here. Apurva Narechania and Josie Reinhardt provided valuable assistance with transcriptome assemblies and bioinformatic analysis. S.S. would like to thank the German Academic Exchange Service (DAAD) and the Biosystematics Group at Wageningen University for support. S.O. is supported by NSF DBI-1307844. Funding for R.B.'s research on stalk-eyed flies was provided by NSF DEB-0951816.

References

Allen, C.E., Zwaan, B.J. and Brakefield, P.M. (2011) Evolution of sexual dimorphism in the Lepidoptera. *Annu Rev Entomol* **56**: 445–464.

Alon, M., Elbaz, M., Ben-Zvi, M.M., Feldmesser, E., Vainstein, A. and Morin, S. (2012) Insights into the transcriptomics of polyphagy: bemisia tabaci adaptability to phenylpropanoids involves coordinated expression of defense and metabolic genes. *Insect Biochem Mol Biol* **42**: 251–263.

Amdam, G.V., Norberg, K., Fondrk, M.K. and Page, R.E. (2004) Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. *Proc Natl Acad Sci U S A* **101**: 11350–11355.

Andres, J.A., Larson, E.L., Bogdanowicz, S.M. and Harrison, R.G. (2013) Patterns of transcriptome divergence in the male accessory gland of two closely related species of field crickets. *Genetics* **193**: 501–513.

Armitage, S.A.O., Freiburg, R.Y., Kurtz, J. and Bravo, I.G. (2012) The evolution of *Dscam* genes across the arthropods. *BMC Evol Biol* **12**: 53–67.

Arunkumar, K.P., Mita, K. and Nagaraju, J. (2009) The silkworm Z chromosome is enriched in testis-specific genes. *Genetics* **182**: 493–501.

Assis, R., Zhou, Q. and Bachtrog, D. (2012) Sex-biased transcriptome evolution in *Drosophila*. *Genome Biol Evol* **4**: 1189–1200.

Azevedo, R.V., Dias, D.B., Bretas, J.A., Mazzoni, C.J., Souza, N.A., Albano, R.M. *et al*. (2012) The transcriptome of *Lutzomyia longipalpis* (Diptera: Psychodidae) male reproductive organs. *PLoS One* **7**: e34495.

Bachtrog, D. (2006) A dynamic view of sex chromosome evolution. *Curr Opin Genet Dev* **16**: 578–585.

Badisco, L., Ott, S.R., Rogers, S.M., Matheson, T., Knapen, D., Vergauwen, L. *et al*. (2011) Microarray-based transcriptomic analysis of differences between long-term gregarious and solitarious desert locusts. *PLoS One* **6**: e28110.

Baek, J.H., Oh, J.H., Kim, Y.H. and Lee, S.H. (2013) Comparative transcriptome analysis of the venom sac and gland of social wasp *Vespa tropica* and solitary wasp *Rhynchium brunneum*. *J Asia-Pac Entomology* **16**: 497–502.

Bailey, N.W., Veltsos, P., Tan, Y.F., Millar, A.H., Ritchie, M.G. and Simmons, L.W. (2013) Tissue-specific transcriptomics in the field cricket *Teleogryllus oceanicus*. *Genes Genomes Genet* **3**: 225–230.

Baker, D., Nolan, T., Fischer, B., Pinder, A., Crisanti, A. and Russell, S. (2011) A comprehensive gene expression atlas of sex- and tissue-specificity in the malaria vector, *Anopheles gambiae*. *BMC Genomics* **12**: 296–307.

Baker, R.H. and Wilkinson, G.S. (2010) Comparative genomic hybridization (CGH) reveals a neo-X chromosome and biased gene movement in stalk-eyed flies (Genus *Teleopsis*). *PLoS Genet* **6**: e1001121.

Baker, R.H., Morgan, J., Wang, X., Boore, J.L. and Wilkinson, G.S. (2009) Genomic analysis of a sexually-selected character: EST sequencing and microarray analysis of eye-antennal imaginal discs in the stalk-eyed fly *Teleopsis dalmanni* (Diopsidae). *BMC Genomics* **10**: 1–20.

Baker, R.H., Narechania, A., Johns, P.M. and Wilkinson, G.S. (2012) Gene duplication, tissue-specific gene expression and sexual conflict in stalk-eyed flies (Diopsidae). *Philos Trans R Soc B Biol Sci* **367**: 2357–2375.

Bao, Y.Y., Qu, L.Y., Zhao, D., Chen, L.B., Jin, H.Y., Xu, L.M. *et al*. (2013) The genome- and transcriptome-wide analysis of innate immunity in the brown planthopper, *Nilaparvata lugens*. *BMC Genomics* **14**: 160–182.

Bass, C., Zimmer, C.T., Riveron, J.M., Wilding, C.S., Wondji, C.S., Kaussmann, M. *et al*. (2013) Gene amplification and microsatellite polymorphism underlie a recent insect host shift. *Proc Natl Acad Sci U S A* **110**: 19460–19465.

Birol, I., Jackman, S., Nielsen, C., Qian, J., Varhol, R., Stazyk, G. *et al*. (2009) De novo transcriptome assembly with ABySS. *Bioinformatics* **25**: 2872–2877.

Bonduriansky, R. and Chenoweth, S.F. (2009) Intralocus sexual conflict. *Trends Ecol Evol* **24**: 280–288.

Boomsma, J.J. (2009) Lifetime monogamy and the evolution of eusociality. *Philosophical Trans R Soc B-Biol Sci* **364**: 3191–3207.

Bradnam, K.R., Fass, J.N., Alexandrov, A., Baranay, P., Bechner, M., Birol, I. *et al*. (2013) Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. *GigaScience* **2**: 1–31.

Catalan, A., Hutter, S. and Parsch, J. (2012) Population and sex differences in Drosophila melanogaster brain gene expression. *BMC Genomics* **13**: 654–666.

Celorio-Mancera, M.D., Wheat, C.W., Vogel, H., Soderlind, L., Janz, N. and Nylin, S. (2013) Mechanisms of macroevolution: polyphagous plasticity in butterfly larvae revealed by RNA-Seq. *Mol Ecol* **22**: 4884–4895.

Chang, P.L., Dunham, J.P., Nuzhdin, S.V. and Arbeitman, M.N. (2011) Somatic sex-specific transcriptome differences in *Drosophila* revealed by whole transcriptome sequencing. *BMC Genomics* **12**: 364–384.

Charlesworth, B. (1991) The evolution of sex chromosomes. *Science* **251**: 1030–1033.

Charlesworth, D., Charlesworth, B. and Marais, G. (2005) Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**: 118–128.

Chen, S., Yang, P., Jiang, F., Wei, Y., Ma, Z. and Kang, L. (2010) De novo analysis of transcriptome dynamics in the migratory locust during the development of phase traits. *PLoS One* **5**: e15633.

Chen, S., Krinsky, B.H. and Long, M. (2013) New genes as drivers of phenotypic evolution. *Nat Rev Genet* **14**: 645–660.

Chi, Y.H., Salzman, R.A., Balfe, S., Ahn, J.E., Sun, W., Moon, J. *et al*. (2009) Cowpea bruchid midgut transcriptome response to a soybean cystatin: costs and benefits of counter-defense. *Insect Mol Biol* **18**: 97–110.

Choi, J.H., Kijimoto, T., Snell-Rood, E., Tae, H., Yang, Y., Moczek, A.P. *et al*. (2010) Gene discovery in the horned beetle *Onthophagus taurus*. *BMC Genomics* **11**: 703–722.

Daugherty, T.H.F., Toth, A.L. and Robinson, G.E. (2011) Nutrition and division of labor: effects on foraging and brain gene expression in the paper wasp *Polistes metricus*. *Mol Ecol* **20**: 5337–5347.

Dermauw, W., Wybouw, N., Rombauts, S., Menten, B., Vontas, J., Grbic, M. *et al*. (2013) A link between host plant adaptation and pesticide resistance in the polyphagous spider mite *Tetranychus urticae*. *Proc Natl Acad Sci U S A* **110**: E113–E122.

Dunning, L.T., Dennis, A.B., Park, D., Sinclair, B.J., Newcomb, R.D. and Buckley, T.R. (2013) Identification of cold-responsive genes in a New Zealand alpine stick insect using RNA-Seq. *Comp Biochem Phys D-Genomics Proteomics* **8**: 24–31.

Ellegren, H. and Parsch, J. (2007) The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet* **8**: 689–698.

Emlen, D.J. (2008) The Evolution of Animal Weapons. *Annu Rev Ecol Evol Systs* **39**: 387–413.

Emlen, D.J., Warren, I.A., Johns, A., Dworkin, I. and Lavine, L.C. (2012) A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. *Science* **337**: 860–864.

Evans, J.D. and Wheeler, D.E. (2001) Gene expression and the evolution of insect polyphenisms. *Bioessays* **23**: 62–68.

Ewen-Campen, B., Shaner, N., Panfilio, K.A., Suzuki, Y., Roth, S. and Extavour, C.G. (2011) The maternal and early embryonic transcriptome of the milkweed bug *Oncopeltus fasciatus*. *BMC Genomics* **12**: 61–83.

Ferreira, P.G., Patalano, S., Chauhan, R., Ffrench-Constant, R., Gabaldon, T., Guigo, R. *et al*. (2013) Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biol* **14**: R20.

Fescemyer, H.W., Sandoya, G.V., Gill, T.A., Ozkan, S., Marden, J.H. and Luthe, D.S. (2013) Maize toxin degrades peritrophic matrix proteins and stimulates compensatory transcriptome responses in fall armyworm midgut. *Insect Biochem Mol Biol* **43**: 280–291.

Gempe, T. and Beye, M. (2011) Function and evolution of sex determination mechanisms, genes and pathways in insects. *Bioessays* **33**: 52–60.

Glaser-Schmitt, A., Catalan, A. and Parsch, J. (2013) Adaptive divergence of a transcriptional enhancer between populations of *Drosophila melanogaster*. *Philosophical Trans R Soc B Biol Sci* **368**: 20130024.

Gong, Z.J., Wu, Y.Q., Miao, J., Duan, Y., Jiang, Y.L. and Li, T. (2013) Global transcriptome analysis of orange wheat blossom midge, *Sitodiplosis mosellana* (Gehin) (Diptera:

Cecidomyiidae) to identify candidate transcripts regulating diapause. *PLoS One* **8**: e71564.

Gotoh, H., Miyakawa, H., Ishikawa, A., Ishikawa, Y., Sugime, Y., Emlen, D.J. *et al*. (2014) Developmental link between sex and nutrition; doublesex regulates sex-specific mandible growth via juvenile hormone signaling in stag beetles. *PLoS Genet* **10**: e1004098.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I. *et al*. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**: 644–652.

Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L. *et al*. (2011) The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471**: 473–479.

Greenberg, J.K., Xia, J., Zhou, X., Thatcher, S.R., Gu, X., Ament, S.A. *et al*. (2012) Behavioral plasticity in honey bees is associated with differences in brain microRNA transcriptome. *Genes Brain Behav* **11**: 660–670.

Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**: 696–704.

Guo, W., Wang, X.H., Ma, Z.Y., Xue, L.A., Han, J.Y., Yu, D. *et al*. (2011) CSP and takeout genes modulate the switch between attraction and repulsion during behavioral phase change in the migratory locust. *PLoS Genet* **7**: e1001291.

Haerty, W., Jagadeeshan, S., Kulathinal, R.J., Wong, A., Ram, K.R., Sirot, L.K. *et al*. (2007) Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*. *Genetics* **177**: 1321–1335.

Hardison, R.C. and Taylor, J. (2012) Genomic approaches towards finding cis-regulatory modules in animals. *Nat Rev Genet* **13**: 469–483.

Harrison, P.W., Mank, J.E. and Wedell, N. (2012) Incomplete sex chromosome dosage compensation in the Indian meal moth, *Plodia interpunctella*, based on de novo transcriptome assembly. *Genome Biol Evol* **4**: 1118–1126.

Huang, Q., Sun, P., Zhou, X. and Lei, C. (2012) Characterization of head transcriptome and analysis of gene expression involved in caste differentiation and aggression in *Odontotermes formosanus* (Shiraki). *PLoS One* **7**: e50383.

Hull, J.J., Perera, O.P. and Snodgrass, G.L. (2014) Cloning and expression profiling of odorant-binding proteins in the tarnished plant bug, *Lygus lineolaris*. *Insect Mol Biol* **23**: 78–97.

Husseneder, C., McGregor, C., Lang, R.P., Collier, R. and Delatte, J. (2012) Transcriptome profiling of female alates and egg-laying queens of the Formosan subterranean termite. *Comp Biochem Physiol D-Genomics Proteomics* **7**: 14–27.

Jaquiery, J., Rispe, C., Roze, D., Legeai, F., Le Trionnaire, G., Stoeckel, S. *et al*. (2013) Masculinization of the x chromosome in the pea aphid. *PLoS Genet* **9**: e1003690.

Johnson, B.R. and Linksvayer, T.A. (2010) Deconstructing the superorganism: social physiology, groundplans, and sociogenomics. *Quart Rev Biol* **85**: 57–79.

Johnson, B.R. and Tsutsui, N.D. (2011) Taxonomically restricted genes are associated with the evolution of sociality in the honey bee. *BMC Genomics* **12**: 164–174.

Johnson, B.R., Atallah, J. and Plachetzki, D.C. (2013) The importance of tissue specificity for RNA-seq: highlighting the errors of composite structure extractions. *BMC Genomics* **14**: 586–594.

Kaiser, V.B. and Bachtrog, D. (2010) Evolution of sex chromosomes in insects. *Annu Rev Genet* **44**: 91–112.

Kimura, S. (2013) The Nap family proteins, CG5017/Hanabi and Nap1, are essential for *Drosophila* spermiogenesis. *FEBS Lett* **587**: 922–929.

King, M.C. and Wilson, A.C. (1975) Evolution at two levels in humans and chimpanzees. *Science* **188**: 107–116.

Korb, J., Weil, T., Hoffmann, K., Foster, K.R. and Rehli, M. (2009) A gene necessary for reproductive suppression in termites. *Science* **324**: 758–758.

Lande, R. (1980) Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* **34**: 292–305.

Le Conte, Y., Alaux, C., Martin, J.F., Harbo, J.R., Harris, J.W., Dantec, C. *et al*. (2011) Social immunity in honeybees (*Apis mellifera*): transcriptome analysis of varroa-hygienic behaviour. *Insect Mol Biol* **20**: 399–408.

Li, L.T., Zhu, Y.B., Ma, J.F., Li, Z.Y. and Dong, Z.P. (2013) An analysis of the *Athetis lepigone* transcriptome from four developmental stages. *PLoS ONE* **8**: e73911.

Liang, Z.S., Trang, N., Mattila, H.R., Rodriguez-Zas, S.L., Seeley, T.D. and Robinson, G.E. (2012) Molecular determinants of scouting behavior in honey bees. *Science* **335**: 1225–1228.

Linksvayer, T.A. and Wade, M.J. (2009) Genes with social effects are expected to harbor more sequence variation within and between species. *Evolution* **63**: 1685–1696.

Liu, J., Li, M., Li, J.-M., Huang, C.-J., Zhou, X.-P., Xu, F.-C. *et al*. (2010) Viral infection of tobacco plants improves performance of Bemisia tabaci but more so for an invasive than for an indigenous biotype of the whitefly. *J Zhejiang Univ Sci B* **11**: 30–40.

Luan, J.B., Wang, Y.L., Wang, J., Wang, X.W. and Liu, S.S. (2013) Detoxification activity and energy cost is attenuated in whiteflies feeding on Tomato yellow leaf curl China virus-infected tobacco plants. *Insect Mol Biol* **22**: 597–607.

Ma, J., Wang, J., Mao, X.F. and Wang, Y. (2012) Differential expression of two antifreeze proteins in the desert beetle *Anatolica polita* (Coleoptera: Tenebriondae): seasonal variation and environmental effects. *Cryoletters* **33**: 337–348.

Magnusson, K., Lycett, G., Mendes, A., Lynd, A., Papathanos, P.-A., Crisanti, A. *et al*. (2012) Demasculinization of the *Anopheles gambiae* X chromosome. *BMC Evol Biol* **12**: 69–82.

Mannello, F., Ligi, D. and Magnani, M. (2012) Deciphering the single-cell omic: innovative application for translational medicine. *Expert Rev Proteomics* **9**: 635–648.

Martin, J.A. and Wang, Z. (2011) Next-generation transcriptome assembly. *Nat Rev Genet* **12**: 671–682.

Matzkin, L.M. (2012) Population transcriptomics of cactus host shifts in *Drosophila mojavensis*. *Mol Ecol* **21**: 2428–2439.

Maza, E., Frasse, P., Senin, P., Bouzayen, M. and Zouine, M. (2013) Comparison of normalization methods for differential gene expression analysis in RNA-Seq experiments: a matter of relative size of studied transcriptomes. *Communicative Integr Biol* **6**: e25849.

Meiklejohn, C.D., Parsch, J., Ranz, J.M. and Hartl, D.L. (2003) Rapid evolution of male-biased gene expression in *Drosophila*. *Proc Natl Acad Sci U S A* **100**: 9894–9899.

Ometto, L., Shoemaker, D., Ross, K.G. and Keller, L. (2011) Evolution of gene expression in fire ants: the effects of

developmental stage, caste, and species. *Mol Biol Evol* **28**: 1381–1392.

Ozsolak, F. and Milos, P.M. (2011) Single-molecule direct RNA sequencing without cDNA synthesis. *Wiley Interdisciplinary Rev-RNA* **2**: 565–570.

Parisi, M., Nuttall, R., Naiman, D., Bouffard, G., Malley, J., Andrews, J. *et al*. (2003) Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* **299**: 697–700.

Parsch, J. and Ellegren, H. (2013) The evolutionary causes and consequences of sex-biased gene expression. *Nat Rev Genet* **14**: 83–87.

Patel, A., Fondrk, M.K., Kaftanoglu, O., Emore, C., Hunt, G., Frederick, K. *et al*. (2007) The making of a queen: TOR pathway Is a key player in diphenic caste development. *PLoS One* **2**: e509.

Pease, J.B. and Hahn, M.W. (2012) Sex chromosomes evolved from independent ancestral linkage groups in winged insects. *Mol Biol Evol* **29**: 1645–1653.

Picelli, S., Björklund, Å., Faridani, O., Winberg, G., Sagasser, S. and Sandberg, R. (2013) Improved Smart-Seq for sensitive full-length transcriptome profiling in single cells. *Nat Methods* **10**: 1096–1098.

Poelchau, M.F., Reynolds, J.A., Denlinger, D.L., Elsik, C.G. and Armbruster, P.A. (2013a) Transcriptome sequencing as a platform to elucidate molecular components of the diapause response in the Asian tiger mosquito *Aedes albopictus*. *Physiol Entomol* **38**: 173–181.

Poelchau, M.F., Reynolds, J.A., Elsik, C.G., Denlinger, D.L. and Armbruster, P.A. (2013b) RNA-Seq reveals early distinctions and late convergence of gene expression between diapause and quiescence in the Asian tiger mosquito, *Aedes albopictus*. *J Exp Biol* **216**: 4082–4090.

Price, D.P., Nagarajan, V., Churbanov, A., Houde, P., Milligan, B., Drake, L.L. *et al*. (2011) The fat body transcriptomes of the yellow fever mosquito *Aedes aegypti*, pre- and post- blood meal. *PLoS One* **6**: e22573.

Prince, E.G., Kirkland, D. and Demuth, J.P. (2010) Hyperexpression of the X chromosome in both sexes results in extensive female bias of X-linked genes in the flour beetle. *Genome Biol Evol* **2**: 336–346.

Qiu, L.M., Mao, X.F., Hou, F. and Ma, J. (2013) A novel function; Thermal protective properties of an antifreeze protein from the summer desert beetle *Microdera punctipennis*. *Cryobiology* **66**: 60–68.

Ranz, J.M. and Parsch, J. (2012) Newly evolved genes: moving from comparative genomics to functional studies in model systems. How important is genetic novelty for species adaptation and diversification? *Bioessays* **34**: 477–483.

Ranz, J.M., Castillo-Davis, C.I., Meiklejohn, C.D. and Hartl, D.L. (2003) Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* **300**: 1742–1745.

Reinhardt, J.A., Brand, C.L., Paczolt, K.A., Johns, P.M., Baker, R.H. and Wilkinson, G.S. (2014) Meiotic drive impacts expression and evolution of x-linked genes in stalk-eyed flies. *PLoS Genet* **10**: e1004362.

Rice, W.R. (1992) Sexually antagonistic genes: experimental evidence. *Science* **256**: 1436–1439.

Riesgo, A., Andrade, S.C., Sharma, P., Novo, M., Pérez-Porro, A., Vahtera, V. *et al*. (2012) Comparative description of ten transcriptomes of newly sequenced invertebrates and efficiency estimation of genomic sampling in non-model taxa. *Front Zool* **9**: 33–57.

Rinker, D.C., Pitts, R.J., Zhou, X., Suh, E., Rokas, A. and Zwiebel, L.J. (2013) Blood meal-induced changes to antennal transcriptome profiles reveal shifts in odor sensitivities in *Anopheles gambiae*. *Proc Natl Acad Sci U S A* **110**: 8260–8265.

Roessingh, P., Bouaichi, A. and Simpson, S.J. (1998) Effects of sensory stimuli on the behavioural phase state of the desert locust, *Schistocerca gregaria*. *J Insect Physiol* **44**: 883–893.

Rund, S.S., Bonar, N.A., Champion, M.M., Ghazi, J.P., Houk, C.M., Leming, M.T. *et al*. (2013) Daily rhythms in antennal protein and olfactory sensitivity in the malaria mosquito *Anopheles gambiae*. *Scientific Rep* **3**: 2494.

Sackton, T.B., Werren, J.H. and Clark, A.G. (2013) Characterizing the infection-induced transcriptome of *Nasonia vitripennis* reveals a preponderance of taxonomically-restricted immune genes. *PLoS One* **8**: e83984.

Sahara, K., Yoshido, A. and Traut, W. (2012) Sex chromosome evolution in moths and butterflies. *Chromosome Res* **20**: 83–94.

Saminadin-Peter, S.S., Kemkemer, C., Pavlidis, P. and Parsch, J. (2012) Selective sweep of a cis-regulatory sequence in a non-African population of *Drosophila melanogaster*. *Mol Biol Evol* **29**: 1167–1174.

Schmucker, D., Clemens, J.C., Shu, H., Worby, C.A., Xiao, J., Muda, M. *et al*. (2000) Drosophila *Dscam* is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* **101**: 671–684.

Schulz, M.H., Zerbino, D.R., Vingron, M. and Birney, E. (2012) Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* **28**: 1086–1092.

Scolari, F., Gomulski, L.M., Ribeiro, J.M., Siciliano, P., Meraldi, A., Falchetto, M. *et al*. (2012) Transcriptional profiles of mating-responsive genes from testes and male accessory glands of the Mediterranean fruit fly, *Ceratitis capitata*. *PLoS One* **7**: e46812.

Sen, R., Raychoudhury, R., Cai, Y., Sun, Y., Lietze, V.-U., Boucias, D.G. *et al*. (2013) Differential impacts of juvenile hormone, soldier head extract and alternate caste phenotypes on host and symbiont transcriptome composition in the gut of the termite *Reticulitermes flavipes*. *BMC Genomics* **14**: 1–18.

Shuel, R.W. and Dixon, S.E. (1960) The early establishment of dimorphism in the female honeybee, *Apis mellifera* L. *Insectes Sociaux* **7**: 265–282.

Simola, D.F., Wissler, L., Donahue, G., Waterhouse, R.M., Helmkampf, M., Roux, J. *et al*. (2013) Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. *Genome Res* **23**: 1235–1247.

Simpson, S.J., Sword, G.A. and Lo, N. (2011) Polyphenism in Insects. *Curr Biol* **21**: R738–R749.

Singh, R.S. and Artieri, C.G. (2010) Male sex drive and the maintenance of sex: evidence from Drosophila. *J Hered* **101** (Suppl 1): S100–S106.

Sinha, D.K., Nagaraju, J., Tomar, A., Bentur, J.S. and Nair, S. (2012) Pyrosequencing-based transcriptome analysis of the Asian rice gall midge reveals differential response during compatible and incompatible interaction. *Int J Mol Sci* **13**: 13079–13103.

Smith, C.R., Toth, A.L., Suarez, A.V. and Robinson, G.E. (2008) Genetic and genomic analyses of the division of labour in insect societies. *Nat Rev Genet* **9**: 735–748.

Sturgill, D., Zhang, Y., Parisi, M. and Oliver, B. (2007) Demasculinization of X chromosomes in the *Drosophila* genus. *Nature* **450**: 238–241.

Tarazona, S., Garcia-Alcalde, F., Dopazo, J., Ferrer, A. and Conesa, A. (2011) Differential expression in RNA-seq: a matter of depth. *Genome Res* **21**: 2213–2223.

Tarver, M.R., Coy, M.R. and Scharf, M.E. (2012) Cyp15F1: a novel cytochrome P450 gene linked to juvenile hormone-dependent caste differentiation in the termite *Reticulitermes flavipes*. *Arch Insect Biochem Physiol* **80**: 92–108.

Teets, N.M. and Denlinger, D.L. (2013) Autophagy in Antarctica Combating dehydration stress in the world's southernmost insect. *Autophagy* **9**: 629–631.

Teets, N.M., Peyton, J.T., Colinet, H., Renault, D., Kelley, J.L., Kawarasaki, Y. *et al*. (2012) Gene expression changes governing extreme dehydration tolerance in an Antarctic insect. *Proc Natl Acad Sci U S A* **109**: 20744–20749.

Timmermans, M.J.T.N., Roelofs, D., Nota, B., Ylstra, B. and Holmstrup, M. (2009) Sugar sweet springtails: on the transcriptional response of *Folsomia candida* (Collembola) to desiccation stress. *Insect Mol Biol* **18**: 737–746.

Toth, A.L. and Robinson, G.E. (2007) Evo-devo and the evolution of social behavior. *Trends Genet* **23**: 334–341.

Toth, A.L., Varala, K., Henshaw, M.T., Rodriguez-Zas, S.L., Hudson, M.E. and Robinson, G.E. (2010) Brain transcriptomic analysis in paper wasps identifies genes associated with behaviour across social insect lineages. *Proc R Soc B Biol Sci* **277**: 2139–2148.

Uvarov, B.P. (1966) *Grasshoppers and Locusts: A Handbook of General Acridology*, Vol. 1. Anti-Locust Research Centre (Cambridge University Press), Cambridge, UK.

Van Verk, M.C., Hickman, R., Pieterse, C.M. and Van Wees, S.C. (2013) RNA-Seq: revelation of the messengers. *Trends Plant Sci* **18**: 175–179.

Vicoso, B. and Bachtrog, D. (2013) Reversal of an ancient sex chromosome to an autosome in *Drosophila*. *Nature* **499**: 332–335.

Vicoso, B. and Charlesworth, B. (2009) The deficit of male-biased genes on the *D. melanogaster* X chromosome Is expression-dependent: a consequence of dosage compensation? *J Mol Evol* **68**: 576–583.

Vilcinskas, A., Stoecker, K., Schmidtberg, H., Rohrich, C.R. and Vogel, H. (2013) Invasive harlequin ladybird carries biological weapons against native competitors. *Science* **340**: 862–863.

Vogel, H., Heidel, A.J., Heckel, D.G. and Groot, A.T. (2010) Transcriptome analysis of the sex pheromone gland of the noctuid moth *Heliothis virescens*. *BMC Genomics* **11**: 29–50.

Walters, J.R. and Hardcastle, T.J. (2011) Getting a full dose? Reconsidering sex chromosome dosage compensation in the silkworm, *Bombyx mori*. *Genome Biol Evol* **3**: 491–504.

Walters, J.R. and Harrison, R.G. (2010) Combined EST and proteomic analysis identifies rapidly evolving seminal fluid proteins in *Heliconius* butterflies. *Mol Biol Evol* **27**: 2000–2013.

Walters, J.R. and Harrison, R.G. (2011) Decoupling of rapid and adaptive evolution among seminal fluid proteins in *Heliconius* butterflies with divergent mating systems. *Evolution* **65**: 2855–2871.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C. *et al*. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**: 470–476.

Wang, J., Long, M. and Vibranovski, M.D. (2012) Retrogenes moved out of the z chromosome in the silkworm. *J Mol Evol* **74**: 113–126.

Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D. *et al*. (2014) The locust genome provides insight into swarm formation and long-distance flight. *Nat Commun* **5**: 1–9.

Wang, Z., Gerstein, M. and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**: 57–63.

White-Cooper, H. (2010) Molecular mechanisms of gene regulation during *Drosophila* spermatogenesis. *Reproduction* **139**: 11–21.

Whiteman, N.K., Gloss, A.D., Sackton, T.B., Groen, S.C., Humphrey, P.T., Lapoint, R.T. *et al*. (2012) Genes involved in the evolution of herbivory by a leaf-mining drosophilid fly. *Genome Biol Evol* **4**: 900–916.

Wilkinson, G.S. and Dodson, G. (1997) Function and evolution of antlers and eye stalks in flies. In *The Evolution of Mating Systems in Insects and Arachnids* (Choe, J. and Crespi, B., eds), pp. 310–328. Cambridge University Press, Cambridge, UK.

Wilkinson, G.S., Johns, P.M., Metheny, J.D. and Baker, R.H. (2013) Sex-biased gene expression during head development in a sexually dimorphic stalk-eyed fly. *PLoS One* **8**: e59826.

Wilson, E.O. (1971) *The Insect Societies*. Belknap Press of Harvard University Press, Cambridge, MA, USA.

Wolf, J.B. (2013) Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Mol Ecol Resour* **13**: 559–572.

Woodard, S.H., Fischman, B.J., Venkat, A., Hudson, M.E., Varala, K., Cameron, S.A. *et al*. (2011) Genes involved in convergent evolution of eusociality in bees. *Proc Natl Acad Sci U S A* **108**: 7472–7477.

Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S. *et al*. (2014) SOAPdenovo-Trans: De novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics* **30**: 1660–1666.

Zeng, V., Ewen-Campen, B., Horch, H.W., Roth, S., Mito, T. and Extavour, C.G. (2013) Developmental gene discovery in a hemimetabolous insect: de novo assembly and annotation of a transcriptome for the cricket *Gryllus bimaculatus*. *PLoS ONE* **8**: e61479.

Zha, X., Xia, Q., Duan, J., Wang, C., He, N. and Xiang, Z. (2009) Dosage analysis of Z chromosome genes using microarray in silkworm, *Bombyx mori*. *Insect Biochem Mol Biol* **39**: 315–321.

Zhang, Y., Sturgill, D., Parisi, M., Kumar, S. and Oliver, B. (2007) Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. *Nature* **450**: 233–238.

Zheng, W., Peng, T., He, W. and Zhang, H. (2012) High-throughput sequencing to reveal genes involved in reproduction and development in *Bactrocera dorsalis* (Diptera: Tephritidae). *PLoS ONE* **7**: e36463.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Recent transcriptome studies in non-model insects.