Epigenetics and locust life phase transitions

Ulrich R. Ernst, Matthias B. Van Hiel, Geert Depuydt, Bart Boerjan, Arnold De Loof and Liliane Schoofs*

ABSTRACT
Insects are one of the most successful classes on Earth, reflected in an enormous species richness and diversity. Arguably, this success is partly due to the high degree to which polyphenism, where one genotype gives rise to more than one phenotype, is exploited by many of its species. In social insects, for instance, larval diet influences the development into distinct castes; and locust polyphenism has tricked researchers for years into believing that the drastically different solitarious and gregarious phases might be different species. Solitarious locusts behave much as common grasshoppers. However, they are notorious for forming vast, devastating swarms upon crowding. These gregarious animals are shorter lived, less fecund and transmit their phase characteristics to their offspring. The behavioural gregarisation occurs within hours, yet the full display of gregarious characters takes several generations, as does the reversal to the solitarious phase. Hormones, neuropeptides and neurotransmitters influence some of the phase traits; however, none of the suggested mechanisms can account for all the observed differences, notably imprinting effects on longevity and fecundity. This is why, more recently, epigenetics has caught the interest of the polyphenism field. Accumulating evidence points towards a role for epigenetic regulation in locust phase polyphenism. This is corroborated in the economically important locust species Locusta migratoria and Schistocerca gregaria. Here, we review the key elements involved in phase transition in locusts and possible epigenetic regulation. We discuss the relative role of DNA methylation, histone modification and small RNA molecules, and suggest future research directions.

KEY WORDS: Locust phase, Polyphenism, Locust swarming, Locusta migratoria, Schistocerca gregaria, Apis mellifera, Invertebrate, DNA methylation, Histone modification, Methylome

Introduction
The term epigenetics tends to take a variety of meanings (Haig, 2004; Jablonka and Lamb, 2002). In its narrow sense, it can be defined as ‘meiotically and mitotically heritable changes in gene expression, not based on DNA sequence alterations’ (Riggs et al., 1996). In a broader sense, it can also be interpreted as ‘modifications of chromosome structure’ (Bird, 2007). Independent of the interpretation, however, epigenetics did not receive much attention in insect research until recent years.

The most prominent epigenetic mechanisms, namely (1) methylation of cytosine in DNA, (2) modifications of histone proteins and (3) ncDNA positioning and regulation by non-coding RNA, were only rarely the focus of the entomology field. This may have been due to the general low, often almost undetectable, levels of methylation in insects and other invertebrates (Glastad et al., 2011), including the prime model organisms Drosophila melanogaster and Caenorhabditis elegans, suggesting that DNA methylation in insects is of minor importance (Glastad et al., 2011). However, it should be noted that the detection of methylated bases in DNA and/or RNA is prone to errors, especially when the methylation rate is relatively low. As a result, methylation levels, or even its occurrence per se, have been heavily debated in C. elegans (Klass et al., 1983; Simpson et al., 1986) and D. melanogaster (Achwal et al., 1983; Lyko et al., 2000; Raddatz et al., 2013).

The discovery of a functional DNA methylation system in the European honeybee Apis mellifera (Wang et al., 2006) triggered a renewed interest in the role of epigenetics in insect biology. In addition, recent advances in sequencing technologies have tremendously facilitated the systematic study of epigenomes and epigenetics in a wide range of insects (Beeler et al., 2014; Bonasio et al., 2012; Krauss et al., 2009; Lo et al., 2012; Lyko et al., 2010; Smith et al., 2012; Walsh et al., 2010; Weiner et al., 2013; Xiang et al., 2013; Ye et al., 2013; Zemach et al., 2010; Zwier et al., 2012).

The diversity of epigenetic systems in insects has made them interesting models for understanding DNA methylation (Lyko and Maleszka, 2011). The observation that locust DNA may be the subject of slightly higher methylation levels (1.3–1.9% of total cytosines, depending on the tissue) than reported for other insect species (Boerjan et al., 2011) revived interest in epigenetic control in locust phase polyphenism. Polyphenism, where one genotype produces several phenotypes, is relatively common in the Animal Kingdom (Simpson et al., 2011). Popular examples in vertebrates include temperature-dependent sex determination in some fish and reptiles, where the ambient temperature experienced during a specific time in development triggers the development into male or female (Navarro-Marti et al., 2011). The insect clade provides some of the very best models for polyphenism (reviewed in Moczek, 2010; Simpson et al., 2011; Whitman and Ananthakrishnan, 2009). Metamorphosis in holometabolic insects, where larva, pupa and imago often differ dramatically in various traits, is just one example. Other well-studied examples include the formation of eye-spots in the butterfly Bicyclus anynana (reviewed in Brakefield and Frankino, 2009), or the striking difference in horn size in dung beetles (Onthophagus) (Kijimoto et al., 2013). In aphids it has been shown that both wing polyphenism – i.e. the development of winged or wingless morphs – and morph polyphenism – i.e. oviparity or viviparity – are maternally regulated (Hartfelder and Emlen, 2012; Ogawa and Miura, 2014; Zera and Denno, 1997). In addition, research into polyphenism and developmental plasticity has often relied on the caste phenomenon in social insects.

Locusts undergo similar drastic changes when they start swarming, a behaviour that has been extensively studied for many years (Burrows et al., 2011; Pener and Simpson, 2009; Wang and Kang, 2014). When solitary locusts become gregarious, they form enormous groups of countless individuals, spanning occasionally hundreds of square kilometres (Ferenz, 1990). The socio-economic impact of these swarms is estimated to be up to several billion US...
Schistocerca gregaria material Table Uvarov, 1966; Verlinden et al., 2009) (Fig. 1; supplementary physiology (Pener and Simpson, 2009; Pener and Yerushalmi, 1998; including locust morphology, behaviour, neurochemistry and encompasses extreme phenotypic plasticity at multiple levels Phase transition between the solitarious and the gregarious form

migratory locust transitions in the most important and best-studied species, the review, we will focus on the epigenetic aspects of polyphenetic dollars (www.fao.org/docrep/018/i2940e/i2940e17.pdf). In this study, we will focus on the epigenetic aspects of polyphenetic differences can be seen in the shape and size of the eyes, wings, antennae and jumping hindlegs, as well as in the distribution of sensory receptors (Pener, 1991; Pener and Yerushalmi, 1998). Solitarious individuals are generally bigger and cryptically coloured compared with their long-term gregarious counterparts, which have a conspicuous bright body colour (Fig. 2). More subtle anatomical differences can be seen in the shape and size of the eyes, wings, antennae and jumping hindlegs (Fig. 2). More subtle anatomical differences can be seen in the shape and size of the eyes, wings, antennae and jumping hindlegs (Fig. 2). More subtle anatomical differences can be seen in the shape and size of the eyes, wings, antennae and jumping hindlegs (Fig. 2).

Fig. 1. Photomontage of a solitarious (left) and gregarious (right) last instar Schistocerca gregaria. Crowding induces gregarisation and is instigated via (1) mechanical stimulation of the hindlegs, and/or (2) the combined sight and smell of other locusts. Gregarisation results in altered morphology, physiology, behaviour and colour (only the last of these is depicted; the right half of the image has been modified to artificially match the colours of gregarious locusts). Original image courtesy of Tom Fayle, modified by Fabian Ernst.

differences. Solitarious locusts normally avoid each other, but increased population density can rapidly trigger attraction to other locusts, resulting in aggregation behaviour that can lead to the generation of devastating migratory swarms (Pener and Simpson, 2009; Rogers et al., 2003; Uvarov, 1966). Gregarious morphs exhibit a wider dietary range, display increased locomotory activity, and will fly predominantly during the daytime, in contrast to isolated locusts, which generally fly at night (Pener, 1991; Uvarov, 1977).

Related to behavioural changes imposed by group living and foraging, phase transition in locusts has been found to alter brain structure and functioning. Long-term gregarious desert locusts have a smaller body size, but their brain is substantially larger than that of long-term solitarious locusts (Ott and Rogers, 2010). In addition, Ott and Rogers reported that the relative distribution of brain regions differs between the two phases. Solitarious locusts invest more in lower level sensory processing, reflected by their relatively large primary olfactory and visual neuropils. In contrast, the larger brains of gregarious locusts are more dedicated to the integration of sensory cues in higher level processing regions, which is thought to support their lifestyle as generalist foragers in dense, migratory swarms where competition among group members is high (Ott and Rogers, 2010). Other changes in brain functioning situate at the level of circuit activity and function (Ayali et al., 2004; Blackburn et al., 2010; Burrows et al., 2011; Fuchs et al., 2003). For example, gregarious L. migratoria show reduced habitation of interneuron activity in response to approaching objects. Phase transition also affects associative learning in the desert locust (Simões et al., 2013). Long-term solitarious locusts learn more quickly to associate an odour with an aversive food source, an effect that can be overridden by crowding, enabling locusts to adopt a different feeding strategy. Furthermore, brain functioning is affected by changes in the neurochemistry of the locust’s nervous system upon
phenotypic transformation (Anstey et al., 2009; Rogers et al., 2004; Verlinden et al., 2009).

Solitarious locusts can switch to gregarious behaviour in only a few hours (Ellis, 1953; Ellis, 1962), whereas other changes, in colour, morphology and reproductive physiology, alter on a much slower time scale (Pener and Simpson, 2009; Pener and Yerushalmi, 1998; Roessingh et al., 1993; Simpson et al., 1999). Reversible transition between the two phases occurs gradually and through intermediate forms, spanning multiple generations. Rapid behavioural changes can rely on short-term neuronal plasticity to alter circuit activity and function, whereas morphological responses such as in brain structure or muscle morphology depend on long-term remodelling that can span several generations (Burrows et al., 2011; Simpson and Miller, 2007; Tanaka and Maeno, 2006). This suggests that epigenetic signals accumulate when reinforced, or fade away with time when they are not reinstalled, which would account for the slow change of some phase characters (Burggren, 2015; Jablonka and Raz, 2009) (Fig. 3).

**Making the switch: initiation of phase transition**

To date, no single factor has been identified that can induce the entire set of alterations in a locust during phase transition. However, during the past two decades, a number of factors have been identified that are involved.

**Visual, olfactory and/or mechanosensory information**

Two distinct sensory pathways are involved in the onset and continuation of aggregating behaviour: the cerebral pathway, prompted by the combined visual and olfactory stimuli, and the thoracic pathway, induced by tactile information (Burrows et al., 2011). Stimulating the hindleg, but not other parts of the body, is sufficient for gregarisation behaviour to occur (Ellis, 1959; Simpson et al., 2001). In the Australian plague locust, *Chortoicetes terminifera*, stimulation of the antennae, but not the hindlegs, is sufficient to elicit a phase transition (Cullen et al., 2010). Not only tactile stimulation but also the combined sight and smell of conspecific or heterospecific locusts can induce gregariousness in...
locust nympha as well (Leo Lester et al., 2005; Roessingh et al., 1998). However, it has been reported that certain contact chemicals by themselves can induce gregarisation behaviour (Heifetz et al., 1997; Heifetz et al., 1998; Heifetz et al., 1996). For a review on the roles of semiochemicals in locusts, we refer the reader to Hassanali and co-workers (Hassanali et al., 2005). In contrast, Tanaka and Nishide report that the sight of moving animals is sufficient to induce darkening in desert locusts, whereas odour had barely any effect (Tanaka and Nishide, 2012).

Tactile information through the antennae reflecting the degree of crowding experienced by the mother directly influences the colour of hatchlings in S. gregaria and L. migratoria, arguing for a maternal factor (Maeno et al., 2011). It is likely that an alkylated l-DOPA analogue recently isolated from egg foam, which was found to induce gregarious behaviour in nympha hatched from treated eggs deposited by solitary females, fulfils this role (Islam, 2013; Miller et al., 2008).

**Neuroendocrine control of phase transition: corazonin, juvenile hormone, serotonin and dopamine**

Although the concentration of several potential neurochemicals differs between solitary and gregarious locusts, only serotonin shows a dramatic transient increase within hours of crowding (Rogers et al., 2004). Moreover, injection of serotonin (and its analogues) is sufficient to induce gregariousness in both S. gregaria and L. migratoria, whereas injecting serotonin antagonists inhibits tactile stimulation-induced phase transition (Anstey et al., 2009; Ma et al., 2011). These observations establish serotonin as an important regulatory agent in phase transition, an effect that is probably mediated through protein kinase A signalling (Ott et al., 2012). We note that other researchers report no influence on attraction/avoidance behaviour by serotonin (Tanaka and Nishide, 2013) or on darkening of hatchlings (Maeno et al., 2011) in S. gregaria. However, in the latter experiments, serotonin was injected through the abdominal sternites in the thorax instead of directly into the thoracic ganglia, which could account for the difference in results.

Surprisingly, injection of serotonin or serotonin receptor (5-HTR) agonists into the head cavity of gregarious L. migratoria nympha shifts their behaviour towards solitaryness (Guo et al., 2013a), suggesting serotonin is involved in both the gregarisation and solitarisation decision. These results correspond well with earlier observations that serotonin displays a peak in the optic lobes (located in the head cavity) shortly after isolation, while a similar peak is seen in the thoracic ganglia when these individuals are re-aggregated (Rogers et al., 2004).

Pharmacological and transcriptional silencing experiments identified dopamine, in addition to serotonin, as another potent gregarisation factor in the migratory locust, driving both behavioural transition and melanin deposition (Ma et al., 2011).

Tanaka has postulated that juvenile hormone (JH) in conjunction with corazonin can account for body colour polyphenism, in both the desert and migratory locust (Tanaka, 2006). Corazonin, an undecapeptide released from the corpora cardiaca, causes darkening of locust body colour (Vandersmissen et al., 2006), whereas JH and JH analogues induce the green body colour in solitary forms (Hasegawa and Tanaka, 1994). Besides instigating dark body colouring, corazonin induces a more convex pronotum in isolated-reared locusts (Tanaka et al., 2002), changes body aspect ratios towards those of gregarious animals (Breuer et al., 2003; Hoste et al., 2002; Maeno et al., 2004; Tanaka et al., 2002) and reduces the number of antennal sensilla to that of crow-reared locusts (Maeno and Tanaka, 2004; Yamamoto-Kihara et al., 2004). However, corazonin cannot induce gregarious behaviour (Hoste et al., 2003) or the darkening of offspring of solitary individuals (Tanaka, 2001).

**Comparing solitarious and gregarious phases: molecular differences derived from -omics data**

The recent publication of the first draft of the 6.5 Gbp genome sequence of the migratory locust (Wang et al., 2014) marks a crucial milestone that should greatly aid researchers in their quest to unravel the molecular foundations underlying locust phase change. Earlier -omics approaches have been instrumental in gaining new insights into the physiological transformation that differentiates the solitary locust from its gregarious phase, and will be summarised in this section.

**Proteomics and peptidomics**

*Schistocerca gregaria* pacifastin-like precursors (SGPP1–4), peptides with still unknown functions, display a strong phase-dependent transcription in the brain and fat body, suggesting a physiological role in phase transition (Simonet et al., 2005; Simonet et al., 2004). Similarly, *L. migratoria* LMPP2 has a differential expression between the two phases (Kang et al., 2004). Neuroparsin precursors (NPP) from the desert locust (*Sg-NPP1–4*) similarly show a very distinct phase-dependent expression (Claeys et al., 2006b; Claey s et al., 2005). Interestingly, NPP transcription is induced upon injection with JH or 20-hydroxyecdysone (Claeyes et al., 2006a).

In *L. migratoria*, higher levels of neuruparsin A (NP-A) and ovary maturing parsin (Lom-OMP) were found in the corpora cardiaca of crowded locusts. In contrast, the concentration of adipokinetic hormone (AKH) precursor-related peptide (APRP) was decreased (Ayali et al., 1996). APRPs are by-products of AKH synthesis with an as-yet unknown functional role (Baggerman et al., 2002; Hatle and Springer, 1999). Neuropeptide profiles of the corpora cardiaca and haemolymph confirmed phase-differentiating levels of AKH I and II and APRPs in *S. gregaria*, but strong sex- and age-dependent concentration differences were also observed (Clynen et al., 2002).

The most enigmatic factor is the phase-related peptide (Clynen et al., 2002), a 6 kDa peptide present in much higher concentrations in the haemolymph of gregarious locusts (up to 0.1 mmol l\(^{-1}\)) compared with solitary locusts. Interestingly, upon isolation of crowd-reared locusts, this peptide shows a progressive decrease in concentration spanning multiple generations. This finding represents another clear manifestation of the epigenetic aspects of locust phase differentiation and maintenance. Higher concentrations of this peptide are also found in the eggs of gregarious than in the eggs of solitary *S. gregaria* (Rahman et al., 2003a; Rahman et al., 2002).

**Metabolomics**

Comparison of the metabolic profiles of haemolymph between solitary and crowd-reared *S. gregaria* using \(^1\)H NMR spectroscopy revealed over 20 differentially abundant metabolites (Lenz et al., 2001). Most notably, crowd-reared insects showed much lower levels of lipid and trehalose, as well as the polyamine putrescine, a breakdown product of amino acids.

More recently, HPLC-GC/MS-based metabolic profiles of solitary and gregarious *L. migratoria* haemolymph were analysed over the time course of phase transition (Wu et al., 2012). Over 45% (319) of the detected metabolites differed between the two phases, with multiple lipids, carbohydrates, amino acids and carnitines showing
the most prominent changes. Interestingly, the genetic or pharmacological manipulation of acylcarnitine and acetylcarnitine levels induces both metabolic and behavioural changes associated with phase transition. As carnitine is responsible for the transport of fatty acids into the mitochondrial matrix prior to their degradation, this could reflect increased energy requirements in gregarious locusts. Alterations in energy metabolism, membrane fluidity and lipid-mediated cell signalling are all potential avenues by which changes in lipid metabolism could be implicated in the phase transition.

Transcriptomics

Rahman and co-workers used differential display RT-PCR (Rahman et al., 2003b) in an early attempt to investigate differential expression in phase transition. They found one gene with higher expression in gregarious desert locusts (resembling *Drosophila* SPARC) and one unidentified gene with higher expression in solitary locusts. The subsequent construction of expressed sequence tag (EST) databases has greatly facilitated locust research. In 2004, a high-coverage EST data set (76,012 ESTs) of the whole body, head, hindlegs and midgut tissue of *L. migratoria* hoppers was generated (Kang et al., 2004). A general repression of (mostly anabolic, biosynthetic and muscle-specific) gene expression was observed in the hindleg, midgut and head tissue of gregarious compared with solitary animals, an observation consistent with their weaker leaping ability. The relative expression pattern in the head of gregarious animals was characterised by the strong upregulation of a functionally diverse set of genes. Most notably, very strong activation of the JHPH (juvenile hormone binding protein, hexamerins, prophenoloxidase and haemocyanins) gene family was observed in the head of gregarious locusts. These findings suggest extensive expression changes in nerve cells that could reflect how hormonal signals govern the phase state transition. The same group also described elevated transcript levels of six heat-shock protein (HSP) genes in gregarious locusts compared with solitary locusts (Wang et al., 2007). This activation of HSP expression probably reflects a stress response triggered by increased population density. This increase in the expression of genes related to stress is consistent with the findings of Badisco and co-workers; they detected 214 differentially expressed genes (Badisco et al., 2011b) in an EST database (34,672 ESTs) generated from the central nervous system (CNS) of both phases of the desert locust (Badisco et al., 2011a). The upregulation of HSP genes and immunity genes in gregarious locusts seems to offer protection from the acute detrimental effects of crowding. At the same time, a reduction in the expression of genes associated with energy metabolism and protein synthesis occurs. This reduction of metabolism and biosynthesis in gregarious locusts was also found in migratory locusts by comparing the transcriptome (72,977 sequences) of developing solitary and gregarious locusts (Chen et al., 2010). Taken together, these data clearly suggest that gregarious locusts are associated with a physiological stress state (Boerjan et al., 2010). Conversely, expression of antioxidant genes is repressed in gregarious desert locusts, suggesting increased susceptibility to oxidative stress (Badisco et al., 2011b). In gregarious migratory locusts, a cluster of synaptic transport components, neurotransmitter and neuromodulator receptors, neurotransmitter synthetases and G protein-coupled receptors (GPCRs) was also prominently upregulated, pointing to an important neuromodulatory aspect in phase transition (Chen et al., 2010). In addition, it was found that fourth instar locusts displayed the most divergence between polymorphic states. This was confirmed by micro-array studies, which illustrated that the top pathways affected in gregarious nymphs at fourth stage were involved in general metabolism and catecholamine biosynthesis (Ma et al., 2011).

Another micro-array study on fourth instar migratory locusts during phase transition revealed differential expression in chemo-sensory proteins and *takeout* proteins (Guo et al., 2011). Genes of these families were found to be highly expressed in peripheral tissues (sensilla, antennae, labial palps, wings and hindlegs) but not internal tissues. RNA interference (RNAi) against chemo-sensory proteins increased repulsion in gregarious fourth instar nymphs, whereas knockdown of one *takeout* gene (*LmigT01*) increased attraction behaviour in solitary fourth instar nymphs, probably by altering peripheral olfactory sensitivity.

Large-scale transcriptome analysis revealed 105 retro-elements in the migratory locust, some of which show a differential expression pattern between the solitarious and gregarious phase at the fifth instar and in adults (Jiang et al., 2012). The developmental and tissue-specific expression pattern of a single class I transposon element was also highly different in gregarious compared with solitarious locusts (Guo et al., 2010). These observations have prompted the Kang lab to propose a regulatory role for these transposon elements in the phase transition of migratory locusts (Jiang et al., 2012). Interestingly, the degree of genome methylation constitutes an important factor determining the transcriptional activity of (retro)transposons and suggests one possible role for genome methylation in the modulation of phenotypic plasticity in insects (Slotkin and Martienssen, 2007).

A recent transcriptome analysis of brain tissue of locusts experiencing short-term (64 h) gregariisation and solitarisation revealed a staggering 4893 differentially expressed genes (28% of the transcriptome) in the two processes (Wang et al., 2014). Increased expression of genes in synaptic transmission, carbohydrate metabolism and nucleosome assembly seems to point towards increased neuronal activity during locust crowding, while at the same time the lowered expression of genes in redox biology suggests suppression of antioxidative responses in the CNS. The same authors also identified 45 genes where alternative splicing results in differential isoform expression between phases. A common theme in the transcriptome, alternative splicing and methylome datasets (see below) is the differential expression (and methylation) of cytoskeletal/microtubular genes, probably reflecting the neuronal plasticity accompanying the behavioural changes upon phase transition.

Epigenetic mechanisms in invertebrates

**Histones in insect epigenetics**

Next to DNA methylation and demethylation processes, reversible post-translational modifications (PTMs) of histone proteins are the best-studied elements of epigenetic mechanisms. More than 160 histone modifications, e.g. methylation, acetylation and phosphorylation (Bannister and Kouzarides, 2011; Suganuma and Workman, 2011; Tan et al., 2011), alter chromatin structure and density and hence the accessibility of DNA, which influences transcription rates ('histone code') (Jenuwein and Allis, 2001; Strahl and Allis, 2000). Histone modifications and chromatin states have been intensively studied in *D. melanogaster* (Filion et al., 2010; Kharchenko et al., 2011; Nègre et al., 2011). Less is known about the role of histone PTMs in other insects. However, Nanty and colleagues showed that patterns of histone PTMs are largely conserved between invertebrate species and can therefore be predicted for different taxa (Nanty et al., 2011). Indeed, DNA methylation and histone modifications seem to work together, if not
is no link established between hypomethylation and body length. However, it should be noted that 5-azacytidine and other nucleoside analogues are not specific DNA methylation inhibitors and also affect other pathways (Gnyszka et al., 2013; Poirier et al., 2014).

Further evidence for epigenetic inheritance in *Drosophila* and *Daphnia* is reviewed elsewhere (Youngson and Whitelaw, 2008), but the mechanisms of this 'soft' inheritance are generally not well understood.

### Evidence for epigenetics in locusts

#### DNA methylation

To our knowledge, the earliest study on DNA methylation in locusts dates from 1951 (Wyatt, 1951), and reported that 0.96% of all cytosines are methylated in *L. migratoria*. Surprisingly, these early findings were followed by a 60 year gap. In 2011, we showed that, compared with that of other insects, *S. gregaria* DNA is relatively highly methylated (1.3–1.9% of total cytosines, depending on the tissue) (Boerjan et al., 2011). The *Schistocerca* transcriptome contains transcripts for some of the enzymes belonging to the epigenetic machinery, including a methyl binding protein (MBD), a histone acetyl transferase (HAT), a histone deacetylase (HDAC) and homologues of Dnmt1 and Dnmt2 (Boerjan et al., 2011; Falckenhayn et al., 2013). Expression levels of Dnmt2 in the metathoracic ganglion change during crowding. Next-generation shotgun bisulphite sequencing to identify the *S. gregaria* methylene confirmed 1.3–1.4% cytosine methylation, 90% of which are in a CpG (cytosine followed by guanine) context (Falckenhayn et al., 2013). The locust genome is more highly methylated than most known insect genomes (but less so than the genomes of other Orthoptera: *L. migratoria*, 1.6% (Wang et al., 2014); *Grylloptalpa fossor*, 3% (Sarkar et al., 1992); *Chorthippus parallelus*, 4.06±0.68% (Lechner et al., 2013)). As the genome of *S. gregaria* has not been sequenced so far, the sequences were mapped against an EST database (Badisco et al., 2011a). Of those that could be mapped, 3.2% and 3.1% of cytosines for brain and metathoracic ganglia, respectively, are methylated, 97% of which are in a CpG context. This more than twofold higher methylation level in comparison with the overall methylation pattern suggests that methylation is targeted to exons. In contrast to the honeybee (Lyko et al., 2010) and the silkworm (Xiang et al., 2010; Zemach et al., 2010), but similar to the stick insect (Krauss et al., 2009) and *L. migratoria* (Robinson et al., 2011; Wang et al., 2014), repetitive elements (ribosomal DNA, transposons) are also methylated. Genes with a low CpG observed/expected ratio are assumed to have been (historically) methylated in the germline, as methylated cytosines tend to mutate to tyrosines, causing depletion of cytosines over evolutionary time scales (Bird, 1980; Duncan and Miller, 1980). Indeed, these genes are more methylated in *L. migratoria*: 20% of the contigs were over 95% methylated and another 20% were more than 65% methylated, an unusually high methylation rate and distinct from other invertebrates. However, gene methylation did not correlate with gene expression levels in six of the investigated genes, suggesting that there is no straightforward link between methylation levels and gene expression.

Recently, the early report of DNA methylation in *L. migratoria* (Wyatt, 1951) was confirmed by Robinson and co-workers, who found that 1.3% of the total cytosines are methylated (Robinson et al., 2011), which is in the same range as *S. gregaria*. Similar to *S. gregaria*, DNA methylation in *L. migratoria* is not restricted to gene bodies, but also targeted to repetitive elements (Robinson et al., 2011). In addition, Robinson and colleagues found transcripts for MBD 2/3, Dnmt2 and two copies of Dnmt1 (Robinson et al., 2011).
Interestingly, genes differentially expressed between the two phases in *L. migratoria* show signs of CpG depletion. The hypermutability of methylated cytosines leads to the formation of thymines via deamination. Thus, a depletion of CpGs over time occurs if highly methylated sequences in the genome were affected. The CpG O/E value is the ratio between observed and expected CpGs within a sequence and is a signature of historical DNA methylation in the germline.

In 2014 the genome sequence of *L. migratoria* was reported (Wang et al., 2014) and confirmed the findings of Robinson and co-workers (Robinson et al., 2011). The *L. migratoria* genome encodes an apparently functional methylation system, containing two copies of Dnmt1 and a single copy of Dnmt2 and 3. Besides the genes reported by Wang and co-workers, we found evidence of additional genes involved in epigenetic mechanisms in the published genome. BLAST searches suggest the presence of at least six HDACs, two HATs, five histone methyltransferases (HMTs), two histone demethylases (HDMs) and one MBD (supplementary material Table S2, Figs S1–S6).

A comparative methyleme analysis of brain tissues between fourth instar solitary and gregarious *L. migratoria* (Wang et al., 2014) revealed lower and more fluctuating levels of CpG methylation in the coding regions of the genome compared with the genome as a whole. The ratios of observed/expected CpG levels show a bimodal distribution curve [as in *A. mellifera* and *S. gregaria* (Elango et al., 2009; Falckenhayn et al., 2013; Foret et al., 2009; Wang and Leung, 2009)] and suggest historical germline methylation, particularly in the coding regions of the genome. As in *S. gregaria* (Falckenhayn et al., 2013), repetitive elements are highly methylated; introns are more methylated than exons; and 90 genes are differentially methylated (at least four differentially methylated CpG sites) in gregarious versus solitary locusts, including genes involved in cytoskeleton formation. Wang and colleagues suggest that these genes might be involved in synaptic plasticity and, for the phase transition, point to a crucial role of microtubule dynamics control in locust brains.

**Histone modifications**

The role of histone modifications has been less well studied in locusts. Using immunoassays for *S. gregaria*, we found that histone H3 contains phosphorylation (at serine 10 and 28, and threonine 3 and 11, respectively), tri-methylation and acetylation (both at lysine 9 and 27) (Boerjan et al., 2013). Preliminary data suggest that brains of gregarious *S. gregaria* contain more phosphorylated histone H3 than do those of solitary *S. gregaria*.

**ncRNA**

In migratory locusts, burst expression of retro-elements has been observed in the egg stage, which is thought to be involved in locust development and has been proposed as a regulatory mechanism in phase transition (Guo et al., 2010). The involvement of small ncRNAs in *L. migratoria* phase transition was investigated by Wei et al., who compared small ncRNA abundance between the gregarious and solitary phase states (Wei et al., 2009). The two phases differed strongly in both length distribution and type of small RNAs. Gregarious animals had higher expression of small RNAs below 22 nucleotides, whereas the opposite was true for small RNAs above 22 nucleotides. Gregarious animals also had double the amount of miRNAs, whereas the solitary state expressed higher levels of endo-siRNAs and piRNA-like small RNAs. Moreover, miRNA-133 has been shown to inhibit behavioural aggregation by controlling dopamine synthesis in locusts (Yang et al., 2014). All this is in strong support of an epigenetic basis for phase polymorphism.

In summary, increasing evidence points to a pivotal role of epigenetics in controlling phase transitions in locusts, yet definite evidence, i.e. more than correlational, is still lacking.

**Future research directions**

Until recently, evidence for methylation in locusts relied on paper chromatography and photo-spectroscopy (Wyatt, 1951), mass spectrometry (Boerjan et al., 2011; Lechner et al., 2013) and methylation-specific restriction enzyme assays (Robinson et al., 2011). While these methods have their specific value, single base resolution methyleme analyses allow detailed analysis of hypomethylated changes in DNA methylation status between the solitary and gregarious phase (Robinson et al., 2011). Recently, these analyses have been performed in *L. migratoria* (Wang et al., 2014). Parts of a *Schistocerca* methylome have also been characterised (Falckenhayn et al., 2013) but, unfortunately, differences between phases were not investigated. Methyleme studies to date have been based on whole organisms or brain tissue. However, it is likely that some epigenetic differences might be concealed by these approaches, as they might be specifically directed to particular brain regions or even cells (Bonasio, 2012).

**Pharmaceutical manipulation of global DNA methylation status**

Strong evidence for a role of epigenetics in locust phase transition would be the experimental switch between phases in the treated animal and/or its offspring by altering parts of the epigenetic setting, be it DNA methylation, histone modification, ncRNAs, nucleosome positioning, or other. The epigenetic machinery can be experimentally manipulated, e.g. by blocking involved enzymes (such as Dnmts or MBDS), or by downregulating the synthesis of these enzymes. In recent years, it has even become possible to manipulate DNA methylation and histone modifications at specific sites (see below).

Gene regulation by DNA methylation is a complex matter (not to mention the intricate interactions with histones, their modifications and ncRNAs), and transitions from one state to another are usually characterised by a dynamic alteration in the methylation pattern. While some genes (or promoters) are being methylated, others are demethylated, and as such the total content of methylated cytosines could be more or less constant. For instance, most cancers are associated with a specific methylation pattern, where tumour suppressor genes are hypermethylated (i.e. inactivated), while oncogenes are hypomethylated (i.e. activated). It is evident that drugs or other experimental methods that aim at a general lower or higher level of DNA methylation are rather crude tools and will probably not be able to mimic such a delicate balance. Indeed, some cancer forms are associated with global hypomethylation, which also might promote metastasis (reviewed in Szfy, 2009). A complete demethylation is usually lethal, and Dnmts are often essential during development (e.g. Zwier et al., 2012).

Drugs like zebularine (Zhou et al., 2002), azacytidine, and decitabine (reviewed in Christman, 2002) are nucleoside analogues that prevent de novo and maintenance methylation by binding covalently to Dnmts when incorporated into DNA. As these drugs do not actively remove methyl groups, they are only effective where DNA replication takes place, limiting their usage to dividing cells. Several other compounds that are not incorporated in DNA have been described as Dnmt inhibitors or DNA demethylating agents (reviewed in Szfy, 2009), but their mode of action is not yet understood. Mechanisms and enzymes that actively demethylate...
DNA have been described in mammals and plants, the most prominent being TET (ten–eleven translocation) proteins (reviewed in Kohli and Zhang, 2013; Piccolo and Fisher, 2014; Wu and Zhang, 2014). Some of these demethylating pathways do not require DNA replication and would therefore be attractive to manipulate post-mitotic cells.

It should be noted that DNA methylation in insects is not necessarily associated with silencing of gene expression; in fact, DNA methylation in insects is highly correlated with steady gene expression and reduced variability in transcript levels (Foret et al., 2009; Lyko et al., 2010; Zemach et al., 2010).

Pharmaceutical manipulation of global histone modifications
Besides targeting DNA methylation, another promising avenue would be the manipulation of histone modifications. In humans, several classes of drugs have been developed for treating cancer and physiological disorders that interact with histone modifying enzymes such as HMT, HDM, HAT or HDAC (reviewed in Grayson et al., 2010; Bock et al., 2010). Some of these enzymes are widely conserved, and could thus serve as targets for pharmaceuticals. Interestingly, HDAC inhibitors will not affect the whole genome, as both HDACs and HATs appear to be specific for certain sequences (reviewed in Szyf, 2009), which offers the opportunity to target a subset of genes. Some HDAC inhibitors also induce active DNA demethylation independent of DNA replication, which circumvents the disadvantages of drugs based on nucleoside analogues that are restricted to replicating DNA.

Similarly, HMT inhibitors will prevent the methylation of specific histones, which in turn prevent DNA methylation. In contrast, inhibition of histone demethylases would shift the balance between methylation and demethylation towards higher methylated histones. Histone (de)methylating enzymes are also specific, with several dozen specimens in several families described to date (reviewed in Højfeldt et al., 2013). So far, however, there are no reports of active demethylation processes in insects. While some drugs have been shown to work in humans and Drosophila alike (e.g. Greiner et al., 2005; but see Cherblanc et al., 2013), for some it remains to be evaluated whether they are also effective in other insects.

Analyses of DNA and histone modifications
Recent technologies allow mapping of different DNA modifications (methylation, hydroxymethylation, formylation, carboxylation) at single-base resolution (Booth et al., 2012; Raiber et al., 2012; Yu et al., 2012; Song et al., 2013), and we are only beginning to grasp the significance of these unusual nucleobases.

Various methods are available to map DNA methylation, each with their specific advantages and shortcomings (Bock et al., 2010; Harris et al., 2010; Laird, 2010; Mensaert et al., 2014; Nagarajan et al., 2013; Umer and Herceg, 2013). Wang and colleagues (Wang et al., 2014) used reduced representation bisulphite sequencing (RRBS) to map methylated CpG in the brains of solitary and gregarious locusts. This approach allows single-base resolution and absolute quantification of (hydroxy)methylation [albeit with imperfect quantification (Harris et al., 2010)], but is biased towards CpG-rich regions (Bock et al., 2010; Harris et al., 2010) and by the choice of the restriction enzyme used (Deng et al., 2009) and size selection of fragments (Bock et al., 2010). Both RRBS and whole-genome bisulphite sequencing (WGBS) may contain artifacts due to the bisulphite conversion process and subsequent PCR bias.

Studies of DNA modifications should be complemented by analyses of histone modifications where possible, e.g. by ChIP-Seq (chromatin immunoprecipitation followed by sequencing of the immunoprecipitated DNA fragments) (Nagarajan et al., 2013; Rivera and Ren, 2013). A comparison of the epigenomes of the two phases between L. migratoria and S. gregaria will reveal the extent of common mechanisms as well as potentially specific modifications. Epigenetic changes that are shared between the two species should be prioritised in investigating their role in phase transition. It will be interesting to unravel how the different epigenetic pathways in these two species interact to bring about the phenomenon of phase transition.

Genome and epigenome editing
More specific tools to manipulate particular epigenetic signatures will have to be developed, e.g. ‘site-specific epimutagenesis’, as suggested previously (Bonasio, 2012). One strategy might be to target genes with a role in epigenetics using RNAi. Because of the robust systemic RNAi mechanism in locusts (Wynant et al., 2012), knockdown of target genes throughout the insect can easily be achieved by injecting dsRNA into the haemolymph. While RNAi has been successfully used for several years in locusts, it has one major drawback: it rarely induces a complete loss of function of the targeted gene. Innovative site-specific genomic engineering tools are currently being explored. For instance, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) or CRISPR-Cas (clustered regularly interspaced short palindromic repeat–CRISPR-associated proteins) could be used to specifically introduce mutations (indels, insertions and deletions) in target genes with roles in epigenetics, such as HATs and HDACs. CRISPR-Cas systems could even be used to edit several genes at once (Jao et al., 2013; Li et al., 2013; Wang et al., 2013). As HATs and HDACs are substrate specific, this would only affect a subset of histones. Alternatively, methylated DNA sequences of interest might be specifically excised and replaced by unmethylated DNA, or vice versa (Ramalingam et al., 2013).

Even more intriguing was the suggestion that methylating and demethylating enzymes as well as histone modifying enzymes (e.g. HMT, HAT) might be directed to specific DNA sequences of interest using transcription activator-like effector (TALE) or CRISPR-Cas (Bonasio, 2012; Gaj et al., 2013; Rivera and Ren, 2013). This would allow specific (de)methylation of DNA and histones, thereby enabling manipulation of epigenomes with high precision. This can be accomplished by using a catalytically inactive (or ‘dead’) version of Cas9, called dCas9, fused to an effector domain that will carry out its function (Sander and Joung, 2014). While the CRISPR-Cas system has already been shown to work in diverse organisms, such as bacteria, plants, vertebrates and insects, it has not yet been employed in locusts.

Recently, several groups succeeded in using TALE to selectively demethylate DNA (Maeder et al., 2013) and histones (Mendenhall et al., 2013). Konermann and colleagues (Konermann et al., 2013) were able to switch histone deacetylation and methylation on and off in a time- and space-specific manner using light-inducible transcriptional effectors (LITEs) in a combination of TALEs and optogenetic methods. This method is highly versatile, as it allows different DNA sequence recognition systems (e.g. zinc finger, TALE) to be combined with various ‘switches’ (e.g. light or chemical induced) and effectors (e.g. HDACs, HMTs). These manipulations would allow functional analyses of specific sites in the epigenome.

However, to date, locusts have not been genetically or epigenetically modified. One of the challenges will be to deliver the modified enzymes into cells of interest (Gaj et al., 2013). Interestingly, some ZFNs can be delivered directly as proteins across...
the cell membrane (Gaj et al., 2012), but this is unlikely to work for the complex LITEs, which have been delivered by viral vectors (Konermann et al., 2013). In D. melanogaster, injection of cas9 – as DNA, RNA or protein – into the embryo allowed genetic manipulation that was transmitted to the offspring through the germline in 25–100% of cases, depending on the method and genes targeted (e.g. Bassett and Liu, 2014; Lee et al., 2014; Port et al., 2014). Given the ready use of RNAI in locusts (Wynant et al., 2012), it is possible that a similar approach will also work in locusts. If successful, this would open unprecedented opportunities for functional genomics and epigenomics, circumventing the drawbacks of RNAI and pharmaceutical approaches.

Conclusions
Future research should aim to reveal the mechanisms of epigenetic control of locust phase transition. Additional insights are to be expected from comparisons with the Australian plague locust, Chortoicetes terminifera, where phase characteristics can change within 72 h in both directions and change abruptly between generations. Given the advent of new and exciting techniques that allow the specific manipulation of genes and proteins, and the analysis of single cells, these are excellent times to study the molecular mechanics of phase polyphenism in locusts.

Acknowledgements
We apologise to those colleagues whose work we could not cite because of space limitations. We acknowledge the constructive suggestions of two anonymous reviewers that greatly helped to improve the article. We are grateful for the contributions of Liesbet Temmerman and Isabel Beets to the manuscript. Thanks are due to Tom Fayle for permission to use some of his images, to Roger Jonckers for taking care of our locusts and to Marijke Christiaens for assistance with the figures.

Competing interests
The authors declare no competing financial interests.

Author contributions
U.R.E., M.B.V.H., G.D., A.D.L. and L.S. drafted the manuscript. U.R.E. made for taking care of our locusts and to Marijke Christiaens for assistance with the figures.

Supporting information
Supplementary material available online at http://jeb.biologists.org/lookup/suppl/doi:10.1242/jeb.107078/-/DC1

References:


Fig. S1: Putative sequences of Histone Deacetylases found in the genome of Locusta migratoria. Sequences were derived from concatenated BLAST hits.

1.1

Drm Rpd3:  
1   MQSHSKKVCCYYDSDIGNYQQGHPKPKHR1RMTMN1LLNGLYRKLMEIYRPVHATADDEMYFKFSDYEYVRLRSLIRPD1M5EYNKMQRFNVEGDCPV 100
Lom Rpd3:  
---DIGNYYQQQGHPKPKHR1RMTMN1LLNGLYRKLMEIYRPVHATADDEMYFKFSDYEYVRLRSLIRPD1M5EYNKMQRFNVEGDCPV

Drm Rpd3:  
101 FDGLYECQLSAGSSVAVKNKQASEICINWGGGLN1AKKEASVGFCYVNDVIG1LLELLYRQVR1YIDIDVHGDVVEEAFYTTYRVMSTFPHKYG 200
Lom Rpd3:  
FDGLYECQLSAGSSVAVKNKQASEICINWGGGLN1AKKEASVGFCYVNDVIG1LLELLYRQVR1YIDIDVHGDVVEEAFYTTYRVMSTFPHKYG

Drm Rpd3:  
201 EFPGTGDRLDIGRKYAV1N1PLRDMDDAYE1IFPIKSVKMETFPQA1AVLQCGASS1TDGLGLCFLTVRHKCGCE1FVKYMLPMLMVGGGG 300
Lom Rpd3:  
EFPGTGDRLDIGRKYAV1N1PLRDMDDAYE1IFPIKSVKMETFPQA1AVLQCGASS1TDGLGLCFLTVRHKCGCE1FVKYMLPMLMVGGGG

Drm Rpd3:  
301 YTRINVSRCWTSE1ALV1NEALPYDSEYFPGPDFKLHISPSN1MTQN1TE1Y1K1NKRL1F1NLPHAPVG1Q1A1P1EDA1N1ID1S1DDE1ED1KD1 400
Lom Rpd3:  
YTRINVSRCWTSE1ALV1NEALPYDSEYFPGPDFKLHISPSN1MTQN1TE1Y1K1NKRL1F1NLPHAPVG1Q1A1P1EDA1N1ID1S1DDE1ED1KD1

Drm Rpd3:  
401 D11L1PQQ1SDK1R1IV1P1ENA1K1G1A1K1EN1NI 521
Lom Rpd3:  
ERE1L1T1Q1A1L1D1K1I1AP1D1N1Y11S1E1E1E1G1R1G1DR1Q1K

Drm Rpd3:  
501 AGSSGSGS1SG1AG1G1A1G1KEN11 521
Lom Rpd3:  

1.2

Apm HDAC3:  
1   MSYNKKVS1Y1F1N1PDV1GNPHYGPGPHPKPKHR1L1A1V1H1S1L1V1L1N1G1L1H1K1M1Q1Y1R1P1Y1R1A1S1T1H1M1C1R1P1H1S1D1Y1E1V1F1E1L1Y1Q1V1T1Q1P1N1L1Q1Y11K1L1S1H1F1N1V1G1D1C1P1V1F1E1 100
Lom HDAC3:  
----K1V1S1Y1F1N1PDV1GNPHYGPGPHPKPKHR1L1A1V1H1S1L1V1L1N1G1L1H1K1M1Q1Y1R1P1Y1R1A1S1T1H1M1C1R1P1H1S1D1Y1E1V1F1E1L1Y1Q1V1T1Q1P1N1L1Q1Y11K1L1S1H1F1N1V1G1D1C1P1V1F1E1

Apm HDAC3:  
101 GL1FD1FC1SM1TY1G1S1M1AY1G1A1G1K1L1N1C1D1IA1N1W1S1G1G1L1H1K1F1K1E1A1S1G1F1C1Y1I1D1V1I1A1L1E1L1Y1K1H1A1R1V1L1Y1D1D1V1H1G1D1G1V1Q1E1A1P1Y1L1T1D1R1V1M1S1T1F1H1K1G1Y1N1 200
Lom HDAC3:  
GL1FD1FC1SM1TY1G1S1M1AY1G1A1G1K1L1N1C1D1IA1N1W1S1G1G1L1H1K1F1K1E1A1S1G1F1C1Y1I1D1V1I1A1L1E1L1Y1K1H1A1R1V1L1Y1D1D1V1H1G1D1G1V1Q1E1A1P1Y1L1T1D1R1V1M1S1T1F1H1K1G1Y1N1

Apm HDAC3:  
201 FFFGT1GD1M1Y1GE1A1G1E1S1R1Y1S1V1N1V1P1L1K1E1G1D1D1T1S1Y1V1Q1V1F1P1V1I1S1H1M1E1Q1F1P1T1A1V1L1Q1C1G1S1L1A1N1D1R1L1C1F1S1L1T1K1G1H1E1C1V1F1V1R1D1L1V1P1L111V1G1G1Y 300
Lom HDAC3:  
FFFGT1GD1M1Y1GE1A1G1E1S1R1Y1S1V1N1V1P1L1K1E1G1D1D1T1S1Y1V1Q1V1F1P1V1I1S1H1M1E1Q1F1P1T1A1V1L1Q1C1G1S1L1A1N1D1R1L1C1F1S1L1T1K1G1H1E1C1V1F1V1R1D1L1V1P1L111V1G1G1Y

Apm HDAC3:  
301 TLR1N1V1C1R1W1T1E1S1T1L1V1D1E1G11S1N1L1E1N1L1Y1E1L1Y1E1F1A1P1D1F1T1L1P1D1V1V1T1R1Q1D1N1A1S1K1N1S1Q1Y1LE1T1I1T1K1H1V1Y1D1N1K1M1Q1H1S1F1S1V1Q1H1D1V1P1D1F1A1P1D1S1Q1E1G1H1E1D1P1 400
Lom HDAC3:  
TLRN1V1C1R1W1T1E1S1T1L1V1D1E1G11S1N1L1E1N1L1Y1E1L1Y1E1F1A1P1D1F1T1L1P1D1V1V1T1R1Q1D1N1A1S1K1N1S1Q1Y1LE1T1I1T1K1H1V1Y1D1N1K1M1Q1H1S1F1S1V1Q1H1D1V1P1D1F1A1P1D1S1Q1E1G1H1E1D1P1

Apm HDAC3:  
401 DR1N1V1Q1D1T1K11E1P1A1N1Y1S1Y1G1E1K1D1Q1D1K1M1V1S1E1T1 433
Lom HDAC3:  
Dir1V1S1A1E1D1R1E1P1N1K1N1F1D1G1D1E1Q1D1K1

The Journal of Experimental Biology | Supplementary Material
Drm HDAC3:  
1 MTDRKVSFYVADNGVHFGHAPMPQKPRQAVTHSLVMYGLHKKMIYRPYKASAQDMLRFHSDEYIAYLQQPTQNIQCNSVAYTKYLAHFSVGECDP 100
Drm HDAC3:  
101 VFDGLFDPCAMYTMQSGLEAQKLNHNSCINIWSSGGLHAKFEAGFSCYFNDIVIGILLEKLYHPRLVLYPIDVHDGDQVEAFYLDTRVMTASFHKY 200
Lom HDAC3-2: 
201 GNYFFPGTDMYIEIAGAESGYYSNVMPLKEIDQGYEQFVFPIIIAISMDFYRTPAIVLQGADSLAGDRLGCSLSTKGDCEVFKVEKLNVPLTVVG 300
Lom HDAC3-2: 
201 GNYFFPGTDMYIEIAGAESGYYSNVMPLKEIDQGYEQFVFPIIIAISMDFYRTPAIVLQGADSLAGDRLGCSLSTKGDCEVFKVEKLNVPLTVVG 300
Lom HDAC3-2: 
201 GNYFFPGTDMYIEIAGAESGYYSNVMPLKEIDQGYEQFVFPIIIAISMDFYRTPAIVLQGADSLAGDRLGCSLSTKGDCEVFKVEKLNVPLTVVG 300

1.3
Nav HDAC5:  
1 MQQLWEOQQKLEEQQREKIERKLLRALHKKDHDHASAVSTEVKQRLQSFLVNKQREAAANAGAVPGTPGTYRTNLQSQSDSSSTTAHPYRPMQLQEKF 100
Nav HDAC5:  
101 GDDFPLRKTASEPNMLKRLKQVRVMERNMAASRNPSLMPARKDDLSSHLLSRSLALSNSGPSPSPTVGNQSSUPTAGGSSGGSPTIPSEEEA 200
Nav HDAC5:  
201 AAYGGQLSSASQQGSLDLSLFSSPSMPNISLGRPHVPSSSTSGSKLATSEAEVAARLAMGPMLQGLPTFLPSLTVDEQPMHKMQSLE 300
Nav HDAC5:  
301 QATTASPPPTQVHAYTPTITQDVASHARHLKAGHRPLRTSAPLGLQGLMLAPFTHETYDELAEQKLEQORTQHNYQRIQRVTNLVG 400
Nav HDAC5:  
401 SRSQSSAQGQQLLDEAPESESSEVIDLTGKDNSTQGGSASGSSIESIESISIKQQRDREFQLFQQQDLRRONLQHQUATDASYPGRSQAQPELSIALRS 500
Lom HDAC5: 
401 SRSQSSAQGQQLLDEAPESESSEVIDLTGKDNSTQGGSASGSSIESIESISIKQQRDREFQLFQQQDLRRONLQHQUATDASYPGRSQAQPELSIALRS 500
Lom HDAC5: 
401 SRSQSSAQGQQLLDEAPESESSEVIDLTGKDNSTQGGSASGSSIESIESISIKQQRDREFQLFQQQDLRRONLQHQUATDASYPGRSQAQPELSIALRS 500
Lom HDAC5: 
501 SPLVHLGQPAAVAADTYSRPRQDPTTGLAYDLMKLHACACGEMRSHEPHHGRQLQSWARLSETGLQCRDVRISRKRATLLEIQCTHEAHALLFGTN 600
Nav HDAC5:  
501 SPLVHLGQPAAVAADTYSRPRQDPTTGLAYDLMKLHACACGEMRSHEPHHGRQLQSWARLSETGLQCRDVRISRKRATLLEIQCTHEAHALLFGTN 600
Nav HDAC5:  
601 PLTRHKLDMSMQ----PVFLRLSCGIGVDSDTTWDHTAPAAMAGVCVIELAFKAAAGDIKNGFAVRRPQGHARANAMKCFQNAVIAAQLL 700
Nav HDAC5:  
601 PLTRHKLDMSMQ----PVFLRLSCGIGVDSDTTWDHTAPAAMAGVCVIELAFKAAAGDIKNGFAVRRPQGHARANAMKCFQNAVIAAQLL 700
Nav HDAC5:  
701 QKLRLRKLHLLDLNVDHPHNGGQMYQFDFRVRRLYLSIHRRHDXNFPPFGTECGAGMEGMRNVNVSGLNPAGDEYAILAFRTVMPIAKFDPDLP 800
Nav HDAC5:  
701 QKLRLRKLHLLDLNVDHPHNGGQMYQFDFRVRRLYLSIHRRHDXNFPPFGTECGAGMEGMRNVNVSGLNPAGDEYAILAFRTVMPIAKFDPDLP 800
Nav HDAC5:  
701 QKLRLRKLHLLDLNVDHPHNGGQMYQFDFRVRRLYLSIHRRHDXNFPPFGTECGAGMEGMRNVNVSGLNPAGDEYAILAFRTVMPIAKFDPDLP 800
Nav HDAC5:  
801 VLVSAGFDDAHGHPPLGGYKVSAPCGRMTQMLMLTLAGKCVVIALEGYYDLTACDSQACECVARLLEHDELPDIRADPQNAKLDTLQKIAIQMS 900
Nav HDAC5:  
801 VLVSAGFDDAHGHPPLGGYKVSAPCGRMTQMLMLTLAGKCVVIALEGYYDLTACDSQACECVARLLEHDELPDIRADPQNAKLDTLQKIAIQMS 900
Nav HDAC5:  
901 HWFCVHDLAHTVDMSAIEASQKEQHDGTDVTSAMASLSMQQQNNLASTPEHSESVSEEPMEDEAK 961
Nav HDAC5:  
901 HWFCVHDLAHTVDMSAIEASQKEQHDGTDVTSAMASLSMQQQNNLASTPEHSESVSEEPMEDEAK 961
Drm HDAC6: 1  MLKNQCGSVEQFTDIFQNANSKSLVRKPTALIYDESMSSQHCLWLDEKHEYECPRFTRLRCEELNTERCLERPERSATKDEILRLHTEEHFERLKE
Lom HDAC6: 101 TSGDRDEDEELSSYDIYHFSTELFELSILASGTSIETLVDHVLHMARKGMAIFRFPGGHRMKAENYGCFNNVALATQHDLVHQLQRIRILY
Drm HDAC6: 201 HHGGTQRFVYNDPRVYFSHFRHSWPWHLHELSDYAIAGSGAATYNFNVPLATGNTGDYEAFIQQLLPVALFQPELVASGADDLGCPEG
Lom HDAC6: 301 EMEVTPACYPHLLNLPLRDLAIRVAVVLQEGCYCDSLAEAGALTLLRSLQRPPCCPILVETPVLPRAEILAQALLSCIAVHHRWRCQLOQQTYVC
Drm HDAC6: 101 WSMNVACIAEVPLMAFVPNRYIQRSLYLRPMRFQAAGSILAGKLALDYGWAINLGGFHHCCSRYGGFPGCPYADFLLIVLRLFEQFPFR
Lom HDAC6: 201 QGNGHERDFSGDNSVLYILDYNKGIYPNDGEGAIRRAVQLEHFTDAEYLS---RNLEEALREHPDFIVNGTDILDGRDLALLAVSSQ---
Drm HDAC6: 301 LVFSTFRLALGPIVVMILSSGGLKADAVGIQTSVLNLQGLRN

1.5
Drm HDACX: 1 MKRLTERVHHEESESQEEVVVWKRDDASAGKLPIVFSNYAVFGLRHLHFDAAAGKHIHKLLCAQLQDDGSFYEPTLKDQLRHHTREYLKSLR
Lom HDACX: 101 WSMNVACIAEVPLMAFVPNRYIQRSYLRPMRFQAAGSILAGKLALDYGWAINLGGFHHCCSRYGGFPGCPYADFLLIVLRLFEQFPFR
Lom HDACX: 201 QNGHERDFSGDNSVLYILDYNKGIYPNDGEGAIRRAVQLEHFTDAEYLS---RNLEEALREHPDFIVNGTDILDGRDLALLAVSSQ---
Drm HDACX: 301 LVFSTFRLALGPIVVMILSSGGLKADAVGIQTSVLNLQGLRN
The Journal of Experimental Biology | Supplementary Material
Fig. S2: Putative sequences of Histone Acetyltransferases found in the genome of Locusta migratoria. Sequences were derived from concatenated BLAST hits.

2.1

Drm Mof: 1 MSEAELEQTPSAQGVQQIPQEEHEPEQEPQTDAYTIGGPPRTVPVEDAALASLSLDSLQSGDSQAEQSBLDLSSLGQVAEEAAASEEPAPPRQHRD1PS1EDST 100
Drm Mof: 101 PASSTTSSTSSSRRSSRRDDYDSVEAEBAPPEPPEPQQQQQQEEKEKODQQGQSKIPVLEAQAPQPAQQPQKVVEVQQTIETDEPSSSDTVICADINPLY 200
Drm Mof: 201 SGSS11DFFMDPAPNNAITEEVTIVTLPALHKTQGQLPLAAAPPPPAPAAAEEQVEPETASPTDGEPEPAPPVAVLYPTYIRYRMEQPTGQLPTRLAPRD 300
Drm Mof: 301 PRQRMMPPPPPVVLIPQTVSVISANVEIASDDEETSSSDDEDEEEDDADLMNHEMTSIRETVTTGDPMLQKIDISNPDLYIFIRREDGTQVRQQLQG 400
Drm Mof: 401 RTTENAAAPPEYEVYYVLYGILMLGDRGHRISDNADLGITVLPAPPLDPQPQSTREMLAQAAAASASSERQPKAKNYDYLSTCENDSYDSR 500
Drm Mof: 501 KNTRYQRKRYDEINHVAQSAEKHESITIKMKIKLVQAYYQSGRAEHQWMCELQCYLKLAMLQKPRGDANNELK1PQKRRP 600
Lom Mof: --IIEFQYETIQTWYSSYPQYEYARISLFLCECFILKTSGAVRHDQKCTWPHPP

2.2

Drm Gcn5: 1 MSGGPSITIKSPQIDNNTGNAAGQAAATAGGAGSAGAQPQPHGAGASAGGVSAPAEQTRQSLQIQRSKRQKFMNTLPQKLAQMSACQ 100
Drm Gcn5: 101 SEGCTGTWTQPNGHRIDEEQGCFEPEBBRSTCQHLSIHAGLDNEISSNELLGLIAIDMLEMNLgsubQVEEDETKYQLYQFLLRQCVLZA 200
Lom Gcn5: -------------------THVSHLQPLMEQVEQLQLGMVVDENFMSMEQREDDADTEKVYYQFQLLHKLCTL 300
Drm Gcn5: 201 QOAVIYRGPGQDETPCTIKTVALKLVFLYFQIKHLSPLQMTQVEQFLQHLNLYFENFSESTRGLTHDASNNAYRNLTVCFVFACNLSRQCE 400
Lom Gcn5: TKPSVEPGQLQEFERPRAQATLTIENLYGQKHEL---EYQMYLAKMLHLHMLNFWTFSPVRHHTVEAIRSYAVNTRYRLVFCHVPASCDSLPHFE 500
Drm Gcn5: -------------------RTRLPPQAVIPHRDETAKQEEARGVISFHVQGNSLFMRPVSKSTMLWLVGLQNVFS 600
Lom Gcn5: HQLPRMPKHYISRLVHKLMLLHLQKQK---QKVGKQGFS

The Journal of Experimental Biology | Supplementary Material
Drm Gcn5:       601 KDIKLARPVAYGIKEYDSATLMHCELHPSVNTQFIAVIRQBQVEKLIAQRHVNEVQKVRFGLTCFKEGLPVEISIPGLREIGWKPQNRPSRSSP
Lom Gcn5:            KDIKNRAQYQGIKDTEFATLMHCELNPKIAEFSAVRRQKEIVKIIKIEAQEEQVQKVFGLTCEFRGVRSPIESIPGIRETGWRPAARATRVS

Drm Gcn5:       701 LEESTDPEKLATSFASVLPVRQHHTAAVPFRPVTAEEPDYYDHIKYPMKLMGERLKLGYYQTRRLFMADMARIIFSNCDFYNPDTEYRCAVSLER
Lom Gcn5:            TEETADPDALHATLKNVLNTVQAHKSAPMEPDTEAPDYYKVKEPMKLMGTLKIEHNYWSORCQDENFTLCYVYYYKPDFGVVMAQTLKE

Drm Gcn5:       801 YFQTKRRELGLWKD 813
Lom Gcn5:            LFLTKV--------
The Journal of Experimental Biology | Supplementary Material

Fig. S3: Putative sequences of DNA methyltransferases found in the genome of Locusta migratoria. Sequences were derived from concatenated BLAST hits.

3.1

Apm Dnmt1a: 1 MKSN5SFKQT31IDMFSQQLAKKQMNADATKED1QKLGNQIKTENNILQKCEQKNKENHPKTDIEYDIELKHEIEKKAKLTDVNLPTQIKINIFHNC 100
Apm Dnmt1a: 101 EYCRQKLNI---E1KFYFGHPNGAVDEEIADLPICLFLENGTEFSIHEDDRPQNKLTYFSYDKNHGLCSFDTLGIEKENVMLYSFGYIAKEYEASPEGG 200
Apm Dnmt1a: 101 EYCRQKLNI---E1KFYFGHPNGAVDEEIADLPICLFLENGTEFSIHEDDRPQNKLTYFSYDKNHGLCSFDTLGIEKENVMLYSFGYIAKEYEASPEGG 200
Lom Dnmt1: 101 EYCRQKLNI---E1KFYFGHPNGAVDEEIADLPICLFLENGTEFSIHEDDRPQNKLTYFSYDKNHGLCSFDTLGIEKENVMLYSFGYIAKEYEASPEGG 200
Lom Dnmt1: 101 EYCRQKLNI---E1KFYFGHPNGAVDEEIADLPICLFLENGTEFSIHEDDRPQNKLTYFSYDKNHGLCSFDTLGIEKENVMLYSFGYIAKEYEASPEGG 200
**3.2**

Drm Dnmt2:     1 MVFRVLELFSGGGMHYAFNQYAOIADLGQVIAALDNTVANAVYAHNYGNSLKVTRNIQSLSVKGTEKLIQANMLMSPPCQHPTRQGLQRDEKCTSDALTG 100
Lom Dnmt2:                                     GEVVAAAVINTVANMYANFPQKLLQKNIESLTAEEINKMIAEMLSSPCPFPTRNLKQKDRTSPEFLHG 100
Drm Dnmt2:        101 LCGILPECOELEYILVENKGDFSSQARQFNIESELSGFHWFHREPEILTQPVYNYCTVYCIARKGADFPAGGKIEEMPGAAIQNOGLQIAEIV 200
Lom Dnmt2:            FLKLLPELTENLKYILLENVKGFEKSDARDKFVNILQDNSAFQFLSSPCQFHIPESTPPLIAKQ------AAAAAGENPETNTDKTCSLFFHIL 200
Drm Dnmt2:        301 TPREVARLMSFPENFEFPETTVQRRKLYRLIGSINVKGVGERLKITIK 345
Lom Dnmt2:            TPLEIAKLMCFPQDSTFVKKQRYVRLLGSINHVSILKLL---
Fig. S4: Putative sequences of Methyl-CpG-binding protein found in the genome of Locusta migratoria. Sequences were derived from concatenated BLAST hits.

4.1

Drm MBD-like: 1 MQMNPSVTIERKVDSCPWPQDREVRSGNSSANDNSSATASSNNNNKVDVFYYTPTKRAEGPKQDIALPDFQPGKMPHCALPSISLY 100
Lom MBD-like: ---TMALEKRYCEALPKGWQREWIRKSGLSA----------------------GKVDVYYPPP-GLRLPSAVAAVPP--PGGFPFCQLPAPTSLS

Drm MBD-like: 101 RCSAMPLPISGGNGATSGSAANARKFARSQGGNGAAGAAGAAPATATATATATASAPST--ANRQQQQIELSRALTVDVSLVPQRTASIFK 200
Lom MBD-like: NFSAMPLPIS-----------------------------AAPPAATAAATTTTTATAAGQAAGRQQQHYGRALRNDASLVPQRTASIFK

Drm MBD-like: 201 QPVTVIRNHKQDPAKNEPKGKETREKQOFWKEKKLREHDGNDISLPKTIRTVGNQVTQLSVATAHLHMLNAGQGQHSTKADLTKNA 300
Lom MBD-like: QPVTV---HKMSEQKVDKFHKFQKEPWQFWEKEREGLCLDGTEFDAMELPRALPVGVBEDTTLSQVATAHLHSAQPVTQGTSRTVLKNP

Drm MBD-like: 301 MAFMNPEQPLMVAIIREDIRKQEDRVGVARLQDALKT 339
Lom MBD-like: GVFLNPEQPLVQAVTIADIREKQEEVASARKKLQEALK-

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
Fig. S5: Putative sequences of Histone Methyltransferases found in the genome of Locusta migratoria. Sequences were derived from concatenated BLAST hits.

5.1

Drm Set:           1 MSSVPKRAKLGDAPAGDNTSAAAGNEESEEAQLDACNEIDALNEKSEEILKVEQYKNFLKPCYEDRRSLVKPLINPNWVTLSNFINHPQVSGILDEE 100
Lom Set:           -----KQRKSEGSAAE----TEDRRDYNETQALEIDACNEIDALNEKSEEILKVEQYKNFLKPCYEDRRSLVKPLINPNWVTLSNFINHPQVSGILDEE

5.2

Peh MLL:           1 MEIVSEDESEDSQPPKTDPSLPKSKSAAMEAALGLDDEVGEIKKIRKKERENKKEKLKSGFAGIDRKLDSRSVNTETKSIQNK 100
Peh MLL:           101 NADKNNYKDVTVNDHKSEKFNDRTNEKLKNRQKMEIINGDHYSEKVDSLLVDNYKSYGSKVGLKDEALEIDVEDEPSQIMEHSYSLPRER 200
Peh MLL:           201 DFSNSNPSASEEIIKKQEAFLTDHDLTTSKSLTPKSLDNFKGTTTVKEIKSKKSEKLKSHLLYTLPEKVPVYKEKRLILEMVTLEVFLTKIDA 300
Peh MLL:           301 EIDYILKSYSDMLADESESEDLSQPPKTDPSLPKSKSAAMEAALGLDDEVGEIKKIRKKERENKKEKLKSGFAGIDRKLDSRSVNTETKSIQNK 400
Peh MLL:           401 DATKCGNLARFINHSCNPNYCIAKITYEIGKQIVYSKDKIADVITYDKFIEEIEKPCLCQAACGKLYLN 574

5.3

Peh E(z):           1 MNSTKVPEWKVRKSEYIKRQKRKRAYKDEIEKAEWIRNWDEHNHNVIDQLYCESEKQWQAKYDPDPHVDVCRAEVTSYNGIPSQPKVPCVINAPTI 100
Peh E(z):           101 PTMYTWAPTQQNFMVEDETVLHNIPYMGDEVLDKDGKFIEELIKNYDGVHGKDPSFMDDAIFREVLSHVAEASAAPTATKETLASKQ 200
Peh E(z):           201 GDPSSNSFSESEIIEKIKQEAEFTTDLTSSKSLTPKSLDNFKGTTTVKEIKSKKSEKLKSHLLYTLPEKVPVYKEKRLILEMVTLEVFLTKIDA 300
Peh E(z):           301 CTPNIDGIKAESVSRERTMHSFHTLCRRCFKYDCFLHRLQGHAGPNLQKRRYPELKFPAEPCNCSYMIDMKEKLADSKFTSPIDSCNASESDNS 400
Peh E(z):           401 SNSQFSNKDFHENSIOKNGLTVNSAASEINIAIMAMGNMNSTQCQVWTADQAYLRVLHVKYLKNYCAIAHMMNLITCQRCVYEFAQKEDAEFSFEDLRQ 500
Peh E(z):           501 QACGADQFKLTIKCRWKVQGLHRLKHMMPALQAMGAGFQEGAKQNFISEYECGIEI5QGQDARAGVRKGYDYMCSLFLLNDFVVTATRKGKINRF 700
Peh E(z):           601 VTCGAQEQFDIKISCHKNVSVQRGLK-----------------------------QACGADQFKLTIKCRWKVQGLHRLKHMMPALQAMGAGFQEGAKQNFISEYECGIEI5QGQDARAGVRKGYDYMCSLFLLNDFVVTATRKGKINRF 800
Fig. S6: Putative sequences of Histone Demethylases found in the genome of Locusta migratoria. Sequences were derived from concatenated BLAST hits.

6.1

Drm Su(var)3-3:  1 MKPTQFGGSSSKMTEPIEYTLISDDDSGEPTPKRNVHPSSALSAPNPGQKHKHPDESNPATARSDERTSRNRPKVYDSNRPSPGSDTASNDSGS 100
Drm Su(var)3-3:  101 A6GMNPQQAERSOSQTRKSEANATTTSSVGPSGAPNSRPNQNGSDKRSDRAGTPTVLSQEGAVQSQRFNLKNKTPNNEACFPFDISRGSLGHLVYNIR 200
Lom Su(var)3-3:  1 ---------------AGLEGAAAFQSRPFQDQLTEANECFPDFVAQGPPQTKQVFHLR
6.2

Apm Su(var)3-9:  1 MGGEEGLGVTGFQPNLYQKLQDSKLVLQSLTSNQSAQATIGQHAVGKTOIAVSTVQTVFNELENRTENITKIDLTDREADSTRQGEQANLAY 100
Lom Su(var)3-9:  101 TKNMRGKLPAHSGVQOPPTKNDQKLDFQIPQENNYNDNYF-YSKNSNGSNSNDVLFQKNNKEEQETIINKNSDLNQELSMTNENKSE- 200
Table S1: Examples for differences in morphology, physiology, behaviour, and molecular changes between solitarious and gregarious locusts of *Schistocerca gregaria* FORSSKÅL 1775 and *Locusta migratoria* LINNAEUS 1758.

<table>
<thead>
<tr>
<th></th>
<th>Schistocerca gregaria</th>
<th>Locusta migratoria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Solitarious</strong></td>
<td><strong>Gregarious</strong></td>
<td><strong>Solitarious</strong></td>
</tr>
<tr>
<td><strong>Gregarious</strong></td>
<td></td>
<td><strong>Gregarious</strong></td>
</tr>
<tr>
<td><strong>1. Morphological differences</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large females, small males</td>
<td>Smaller females, large males</td>
<td>Large females, small males</td>
</tr>
<tr>
<td>Long wings</td>
<td>Short wings</td>
<td>Long wings</td>
</tr>
<tr>
<td>Convex pronotum</td>
<td>More concave pronotum</td>
<td>Convex pronotum</td>
</tr>
<tr>
<td>High F/C ratio$^1$</td>
<td>Low F/C ratio</td>
<td>High F/C ratio</td>
</tr>
<tr>
<td>Low E/F ratio$^2$</td>
<td>High E/F ratio</td>
<td>Low E/F ratio</td>
</tr>
<tr>
<td>More sensilla on antennae</td>
<td>Less sensilla on antennae</td>
<td>More sensilla on antennae</td>
</tr>
<tr>
<td>More ovarioles</td>
<td>Less ovarioles</td>
<td>More ovarioles</td>
</tr>
<tr>
<td>More, smaller eggs</td>
<td>Fewer, bigger eggs</td>
<td>More, smaller eggs</td>
</tr>
<tr>
<td>Light hatchlings</td>
<td>Dark hatchlings</td>
<td>Light hatchlings</td>
</tr>
<tr>
<td><strong>Beige adults</strong></td>
<td><strong>Yellow adults</strong></td>
<td><strong>Green adults</strong></td>
</tr>
<tr>
<td><strong>Darkbrown females/yellow males</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Green/beige nymphs</strong></td>
<td><strong>Black and yellow nymphs</strong></td>
<td><strong>Green/lightbrown nymphs</strong></td>
</tr>
<tr>
<td><strong>5-6 nymphal stages</strong></td>
<td><strong>5 nymphal stages</strong></td>
<td><strong>5-7 nymphal stages</strong></td>
</tr>
<tr>
<td><strong>5 nymphal stages</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long hind legs</td>
<td>Shorter hind legs</td>
<td></td>
</tr>
</tbody>
</table>

$^1$Note: F stands for female, C for conocephalus sex.$^2$Note: E stands for eye, F for femur.
<table>
<thead>
<tr>
<th>Trait</th>
<th>Trait</th>
<th>Trait</th>
<th>Trait</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-6 eye stripes</td>
<td>5 eye stripes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large eyes</td>
<td>Small eyes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small brain</td>
<td>Larger brain (30%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long antennae</td>
<td>Short antennae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More sensilla on hind legs</td>
<td>Less sensilla on hind legs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More frontal hairs</td>
<td>Less frontal hairs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less touch receptors</td>
<td>More touch receptors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shorter sternal hairs</td>
<td>Longer sternal hairs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less spermatophores</td>
<td>More spermatophores</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 2. Physiological differences

<table>
<thead>
<tr>
<th>Trait</th>
<th>Trait</th>
<th>Trait</th>
<th>Trait</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long-lived</td>
<td>Short-lived</td>
<td>Long-lived</td>
<td>Short-lived</td>
</tr>
<tr>
<td>Long maturation times</td>
<td>Short maturation times</td>
<td>Long maturation times</td>
<td>Short maturation times</td>
</tr>
<tr>
<td>Lower O₂ consumption</td>
<td>Higher O₂ consumption</td>
<td>Lower O₂ consumption</td>
<td>Higher O₂ consumption</td>
</tr>
<tr>
<td>Lower heartbeat frequency</td>
<td>Higher heartbeat frequency</td>
<td>Lower heartbeat frequency</td>
<td>Higher heartbeat frequency</td>
</tr>
<tr>
<td>Lower fat content</td>
<td>Higher fat content</td>
<td>Lower fat content</td>
<td>Higher fat content</td>
</tr>
<tr>
<td>Lower carbohydrate content</td>
<td>Higher carbohydrate</td>
<td>Lower carbohydrate content</td>
<td>Higher carbohydrate</td>
</tr>
<tr>
<td>Lower spontaneous activity of TCD in the light</td>
<td>Higher spontaneous activity of TCD in the light</td>
<td>Lower spontaneous activity of TCD in the light</td>
<td>Higher spontaneous activity of TCD in the light</td>
</tr>
<tr>
<td>No difference between activity of TCG in light and dark</td>
<td>Higher spontaneous activity of TCG in the dark</td>
<td>Lower spontaneous activity of TCG in the dark</td>
<td>Higher spontaneous activity of TCG in the dark</td>
</tr>
</tbody>
</table>
### 3. Molecular differences

<table>
<thead>
<tr>
<th><strong>Lower SGPP1-4 expression</strong>&lt;sup&gt;5&lt;/sup&gt;</th>
<th><strong>Higher SGPP1-4 expression</strong></th>
<th><strong>Higher LMPP2 expression</strong>&lt;sup&gt;6&lt;/sup&gt;</th>
<th><strong>Lower LMPP2 expression</strong></th>
<th><strong>No difference in octopamine levels</strong></th>
<th><strong>Higher octopamine levels</strong></th>
<th><strong>Lower octopamine levels</strong></th>
<th><strong>Higher JH titer</strong>&lt;sup&gt;7&lt;/sup&gt; (debated)</th>
<th><strong>Higher JH titer</strong> (debated)</th>
<th><strong>Higher overall JH biosynthesis</strong></th>
<th><strong>Lower overall JH biosynthesis</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lower JH titer</strong>&lt;sup&gt;7&lt;/sup&gt; (debated)</td>
<td><strong>Higher JH titer</strong> (debated)</td>
<td><strong>Higher overall JH biosynthesis</strong></td>
<td><strong>Lower overall JH biosynthesis</strong></td>
<td><strong>Lower JH titer</strong>&lt;sup&gt;7&lt;/sup&gt; (debated)</td>
<td><strong>Higher JH titer</strong> (debated)</td>
<td><strong>Higher overall JH biosynthesis</strong></td>
<td><strong>Lower overall JH biosynthesis</strong></td>
<td><strong>Lower JH titer</strong>&lt;sup&gt;7&lt;/sup&gt; (debated)</td>
<td><strong>Higher JH titer</strong> (debated)</td>
<td><strong>Higher overall JH biosynthesis</strong></td>
</tr>
<tr>
<td>Low ecdysteroid level in eggs</td>
<td>High ecdysteroid level in eggs</td>
<td>Low ecdysteroid levels in haemolymph</td>
<td>High ecdysteroid levels in haemolymph</td>
<td>No phenylacetonitrile production</td>
<td>Phenylacetonitrile production</td>
<td>Low phenylacetonitrile production</td>
<td>Low phenylacetonitrile production</td>
<td>Low yellow protein (YP) levels</td>
<td>High YP levels in adult males</td>
<td>High YP levels in adult males</td>
</tr>
<tr>
<td>Low carotenoid/biliverdin ratio</td>
<td>High carotenoid/biliverdin ratio</td>
<td>Low carotenoid/biliverdin ratio</td>
<td>High carotenoid/biliverdin ratio</td>
<td>Low phase-related peptide (PRP) level</td>
<td>High PRP level</td>
<td>Low PRP level</td>
<td>High PRP level</td>
<td>Low levels of PRP in eggs</td>
<td>High levels of PRP in eggs</td>
<td>High levels of PRP in eggs</td>
</tr>
<tr>
<td>Slower corazonin breakdown</td>
<td>Faster corazonin breakdown</td>
<td>Faster corazonin breakdown</td>
<td>Faster corazonin breakdown</td>
<td>High expression of NPP1-2&lt;sup&gt;8&lt;/sup&gt;</td>
<td>Low expression of NPP1-2</td>
<td>Low expression of NPP1-2</td>
<td>Low expression of NPP1-2</td>
<td>Low expression of NPP3-4&lt;sup&gt;8&lt;/sup&gt;</td>
<td>High expression of NPP3-4</td>
<td>High expression of NPP3-4</td>
</tr>
<tr>
<td>Higher levels of 8 neurochemicals: Arg, Asp, Glu, Gly, taurine, GABA, dopamine, and serotonin</td>
<td>Higher levels of 4 neurochemicals: acetylcholine, citrulline, N-acetyldopamine, and tyramine</td>
<td>Higher levels of 4 neurochemicals: acetylcholine, citrulline, N-acetyldopamine, and tyramine</td>
<td>Higher levels of 4 neurochemicals: acetylcholine, citrulline, N-acetyldopamine, and tyramine</td>
<td>Higher levels of 8 neurochemicals: Arg, Asp, Glu, Gly, taurine, GABA, dopamine, and serotonin</td>
<td>Higher levels of 4 neurochemicals: acetylcholine, citrulline, N-acetyldopamine, and tyramine</td>
<td>Higher levels of 4 neurochemicals: acetylcholine, citrulline, N-acetyldopamine, and tyramine</td>
<td>Higher levels of 4 neurochemicals: acetylcholine, citrulline, N-acetyldopamine, and tyramine</td>
<td>Higher levels of 8 neurochemicals: Arg, Asp, Glu, Gly, taurine, GABA, dopamine, and serotonin</td>
<td>Higher levels of 4 neurochemicals: acetylcholine, citrulline, N-acetyldopamine, and tyramine</td>
<td>Higher levels of 4 neurochemicals: acetylcholine, citrulline, N-acetyldopamine, and tyramine</td>
</tr>
</tbody>
</table>
Higher levels of 155 metabolites, especially phosphatidylethanolamines (PEs) and diacylglycerols (DGs)

Differential expression in 214 genes (long-term solitarious), 114 genes upregulated, including genes involved in energy metabolism, protein synthesis, and antioxidant genes

Higher levels of 164 metabolites, especially carnitines, lysophosphatidylcholines, several carbohydrates, amino acids, and some DGs

Differential expression in 214 genes (long-term gregarious), 100 genes upregulated, including heat-shock and immunity genes

Upregulation of genes involved in energy metabolism and protein synthesis

Upregulation of 6 heat-shock protein (HSP) genes

Lower levels of Neuroparsin A

Lower OMP levels in mature adults

Higher levels of APRP

Lower levels of APRP

Lower levels of dopamine

Higher levels of dopamine

Lower haemolymph levels of lipid, trehalose, and putrescine

Differential expression in 532 genes (2-3 generations), including upregulation of mostly anabolic, biosynthetic, and muscle-specific gene expression

Differential expression in 532 genes (2-3 generations), including very strong upregulation of the JHPH gene family in the head

Upregulation of a cluster of synaptic transport components, neurotransmitter and neuro-modulator receptors, neurotransmitter synthetases, and G protein coupled receptors
<table>
<thead>
<tr>
<th>Upregulation of <em>take out</em> in antennae at 4(^{th}) instar</th>
<th>Upregulation of genes for CSPs (chemosensory protein) at 4(^{th}) instar</th>
<th>Upregulation of genes of general metabolism and catecholamine biosynthesis at 4(^{th}) instar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differential expression in 4893 genes in the brain (short-term, 64h), including upregulation of genes in redox biology</td>
<td>Differential expression in 4893 genes in the brain (short-term, 64h), including upregulation of genes in synaptic transmission, carbohydrate metabolism and nucleosome assembly</td>
<td></td>
</tr>
<tr>
<td>Higher abundance of longer (&gt;26-29 nt(^{12})) small RNAs</td>
<td>Lower abundance of longer (&gt;26-29 nt) small RNAs</td>
<td>Lower expression levels of micro-RNA 133</td>
</tr>
<tr>
<td>Higher expression levels of micro-RNA 133</td>
<td>Lower expression levels of micro-RNA 133</td>
<td>Lower expression levels of micro-RNA 133</td>
</tr>
<tr>
<td>Upregulation of transposon element I in antenna and labial palp at 4(^{th}) instar</td>
<td>Upregulation of transposon element I in brain and wings at 4(^{th}) instar</td>
<td>45 alternative spliced isoforms between gregarious and solitary animals</td>
</tr>
<tr>
<td>45 alternative spliced isoforms between gregarious and solitary animals</td>
<td>90 genes differentially methylated</td>
<td></td>
</tr>
</tbody>
</table>

## 4. Behavioural differences

| Flight at night | Flight by day | Flight at night | Flight by day |
| Similar findings for both species are listed first, followed by differences (italic) and findings that have only be described in one of the species. Abbreviations: 1 F/C (length of femur/maximum head width), 2 E/F (length of fore-wing/length of femur), 3 TCD (tritocerebral dwarf), 4 TCG (tritocerebral giant), 5 SGPP (Schistocerca gregaria pacifastin-related precursor), 6 LMPP (Locusta migratoria pacifastin-related precursor), 7 JH (juvenile hormone), 8 NPP (neuropsin precursor), 9 OMP (ovary maturating parsin), 10 APRP (adipokinetic hormone (AKH) precursor-related peptide), 11 JHPH gene family (juvenile hormone binding protein, hexamerins, prophenoloxidase, and hemocyanins), 12 nt (nucleotides) |
**Table S2: List of genes found.** Evidence of the existence of genes with putative roles in epigenetics in the genome of *Locusta migratoria*. Sequences of known genes were used to search the recently released genome sequence. Hits with significant similarity were then reblast to the genome database of NCBI. When this resulted in hits of the same (orthologous) gene, the hit in the *Locusta* genome was assumed to be the orthologue of the initial query sequence.

<table>
<thead>
<tr>
<th>Group</th>
<th>Gene</th>
<th>Fig reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Histone deacetylases (HDAC):</td>
<td><em>Rpd3</em></td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td><em>HDAC3</em>-like</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td><em>HDAC5</em>-like</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td><em>HDAC6</em>-like</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td><em>HDAC-X</em></td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td><em>Sirt4</em></td>
<td>1.6</td>
</tr>
<tr>
<td>2. Histone acetyltransferases (HAT):</td>
<td><em>Mof</em></td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td><em>Gcn5/Pcaf/Kat2A</em></td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td><em>Taf1</em>: not found</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>NCOA2</em>: not found</td>
<td></td>
</tr>
<tr>
<td>3. DNA methyltransferases (Dnmt):</td>
<td><em>Dnmt 1</em> (two copies) (Robinson et al., 2011)</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>Isoform 1, partial: CCA62565.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Isoform 2, partial: CCA62566.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Dnmt 2</em> (Robinson et al., 2011)</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>Partial sequence: CCA62567.1</td>
<td></td>
</tr>
<tr>
<td>4. Methyl-CpG-binding domain (MBD):</td>
<td><em>MBD</em>-like</td>
<td>4.1</td>
</tr>
<tr>
<td>5. Histone methyltransferases (HMT):</td>
<td><em>Set</em></td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td><em>MLL</em></td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td><em>Enhancer of zeste</em></td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td><em>Buzidau</em></td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td><em>Smyd4</em></td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td><em>G9a</em>: not found</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>K9H9</em>: not found</td>
<td></td>
</tr>
<tr>
<td>6. Histone demethylases (HDM):</td>
<td><em>Su(var)3-3</em></td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td><em>Su(var)3-9</em></td>
<td>6.2</td>
</tr>
</tbody>
</table>