Genetic control of reproductive hierarchies in honeybees (*Apis mellifera* sp.).

Dissertation (kumulativ)

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der

Naturwissenschaftlichen Fakultät I - Biowissenschaften
der Martin-Luther-Universität Halle-Wittenberg

von

Frau Antje Jarosch

t geb. am 06.01.1981 in Hennigsdorf

Gutachter:

1. Prof. Dr. R.F.A. Moritz
2. Prof. Dr. Y. LeConte
3. Prof. Dr. Klaus Hartfelder

Halle (Saale), Datum der Verteidigung: 22.03.2012
1 INTRODUCTION........................................................................................................... 1-31

1.1 Reproductive monopolies and reproductive conflicts in social insects
1.2 Worker reproduction in social insects
1.3 Phenotypic plasticity in honeybees
1.4 Intraspecific reproductive parasitism in honeybees
1.5 Genetic determination of reproductive hierarchies in honeybees
1.6 Aims of this work
1.7 References

2. Systemic RNA-interference in the honeybee Apis mellifera: tissue dependent uptake of fluorescent siRNA after intra-abdominal application observed by laser-scanning microscopy ................................................................. 32-50

3. RNA interference in honeybees: off-target effects caused by dsRNA .......... 51-69

4. Molecular characterization of a genomic region which controls reproductive hierarchies in honeybees (Apis mellifera sp.) ............................................................. 70-86

5. Alternative splicing of a single transcription factor drives selfish reproductive behavior in honeybee workers (Apis mellifera) .................................................. 87-112

6. Summary ................................................................................................................. 113-120

6.1 Functional gene studies of the th locus
6.2 gemini – functional and evolutionary aspects
6.3 References

7. Zusammenfassung .................................................................................................... 121-126

7.1 Funktionelle Genstudien am th-Locus
7.2 gemini-funktionelle und evolutionäre Aspekte

8. Appendix ................................................................................................................ 127-132

8.1 Declaration on the contributions to the manuscripts/papers on which this thesis is based
8.2 Acknowledgements
8.3 Curriculum vitae
8.4 Publications
8.5 Erklärung
1. Introduction

1.1. Reproductive monopolies and reproductive conflicts in social insects

Monopoly, which stems from the Greek words *monos* (alone or single) and *polein* (to sell), is economically defined as the existence of a specific individual or an enterprise, which is the only supplier of a particular product or service (Friedman 2002). It is expected to be coercive when the monopoly blocks competing individuals by using unfair competitive practices.

Reproductive monopoly in eusocial societies seems to exactly fit this definition. Such communities are characterised by a reproductive division of labour, where one or few females monopolise both mating and reproduction, whereas workers refrain from reproduction and instead perform tasks that are necessary for the maintenance and growth of the colony (Wilson 1971). Furthermore, eusocial insects are distinguished by an overlap of generations, where the offspring assists parents in the brood care. A highly developed eusocial organisation is described in the insect taxa ants, bees, wasps and termites. Furthermore, some aphids and thrips, an ambrosia beetle, some shrimps and two species of mole rats are described to exhibit eusociality (Queller and Strassmann 2003). Altogether eusociality independently evolved 12 times in arthropods and seven times in male haploid hymenoptera (Wilson and Hölldobler 2005).

Similar to economy, monopolisation of reproduction entails conflicts between competing partners. Reproductive conflicts arise between competitors, which selfishly want to increase their own fitness at the expense of their counterparts. Regarding the Darwinian theory of natural selection (1859), worker sterility had long been regarded as a discrepancy and it was only in 1964 when Hamilton offered an elegant explanation to overcome the apparent contradiction of worker altruism. In his commonly accepted “inclusive fitness theory”, workers may, apart from reproducing themselves, increase their own fitness by an increased reproductive output of related individuals which, due to kin, are genetically related. Thus workers engage in rearing the off-spring of her mother-queen in order to gain a fitness advantage. Therefore, kin selection is the evolutionary force stabilizing the eusocial structures of insect colonies, whereas individual direct selection is negligible.
1.2. Worker reproduction in social insects

In monogynous species where the queen mates just once worker reproduction leads to a conflict over male production between queens and workers, as workers are more closely related to their own offspring \((r=0.5)\) and to the offspring of her sisters \((r=0.35)\), than to her brothers (queen-laid offspring; \(r=0.25\)). This relatedness skew between worker and queen laid male offspring shifts in polyandrous species like the honeybee, where the queen mates up to 18 times on her nuptial flights (Estoup et al. 1994; Neumann and Moritz 2000). In case of *Apis dorsata* Fabricius which has the highest level of polyandry recorded for any social insect the mating frequency even ranges between 27 and 89 (Moritz et al., 1995; Oldroyd et al., 1996; Wattanachaiyingcharoen et al. 2002). In such polyandrous species workers are more closely related to there own sons \((r=0.5)\) compared to the sons of their sisters \((r=0.125)\). As workers are conclusively more closely related to the sons of the queen \((r=0.25)\) than to their nephews a female secures her own genetic interests more effectively by either raising her own sons or the off-spring of the queen, then the off-spring of her sisters.

Complete absence of worker reproduction due to a total reduction of reproductive organs is just present in a few ant species including the genera *Solenopsis*, *Pheidole*, *Tetramorium*, and *Eciton* (Wilson 1971; Oster and Wilson 1978; Fletcher and Ross 1985). In other eusocial insect species workers still exhibit rudimentary reproductive organs. Although mating is not possible (Bourke 1988) workers are capable of laying unfertilized, haploid eggs via arrhenotokous parthenogenesis which develop into males (Ruttner and Hesse 1981; Winston 1987; Page and Erickson 1988; Visscher 1989). But, as such worker reproduction is in conflict with the queen’s and the other workers’ interests, it very rarely occurs. In queenright honeybee colonies less then 0.01 % of all workers have functional ovaries (Visscher 1989, 1996; Ratnieks 1993). These low rates are due to the pheromonal control released by the mandibular and the dufours glands of the queen (Butler et al., 1962; Slessor et al., 1988; Plettner et al. 1993; Winston and Slessor 1998; Katzav-Gozansky et al. 2001) and the pheromonal blend of queen laid brood (Arnold et al. 1994). These queen specific pheromones as well as chemical cues mediate the suppression of worker ovarian development and may communicate the reproductive potential of the queen (Leconte and Hefetz 2008). Nevertheless, very few selfish reproductive workers are able to lay about 7 % off the colony’s male offspring (Visscher 1996). But just 1 out of 1000 worker laid eggs develops into adulthood (Visscher 1989). These low egg hatching rates are due to the second restrictive
mechanism against worker reproduction referred to as worker policing (Ratnieks 1988). Worker laid eggs lack the pheromonal blend of queen laid eggs (Jay 1970; Ratnieks 1995; Mohammedi et al. 1998, Martin et al. 2002a). Missing this discriminative blend, worker laid eggs get eaten and disappear within one hour after laying (Ratnieks and Visscher 1989). This behavioural trait referred to as worker policing also includes aggression towards reproductive workers (Visscher and Dukas 1995). The phenomenon of restrictive behaviour against worker reproduction has also been observed in some wasp species (Foster and Ratnieks 2001 and citations within) and in ants (Monnin and Ratnieks 2001 and citations within).

Nevertheless, under queenless conditions when the queen restrictions against worker fertility get lost, a higher proportion of honeybee workers (up to 10 %) are able to activate their ovaries. Thus the amount of worker laid male eggs increases dramatically (Velthuis 1970; Page and Erickson 1988). But not all workers participate equally in egg-laying activity. Traits related to worker reproduction show a high degree of heritability ($h^2$; trophalactic dominance 0.32; ovary activation 0.27; amount of 9-ODA 0.89), suggesting a strong genetic determination of behavioural, physiological and biochemical cues determining reproductive hierarchies in honeybee colonies (Moritz and Hillesheim 1985). These findings might be reflected by different subfamilies. In fact, some subfamilies engage to a higher extend in drone production than others do (Page and Erickson 1988; Moritz et al. 1996; Martin et al. 2004), suggesting a strong intracolonial selection.

1.3. Phenotypic plasticity in honeybees

Developmental plasticity can be divided into two related classes. (1) Within the reaction norms of a given genotype an organism's environment can elicit a continuous phenotypic variation. (2) Environmental variations result in discrete phenotypic classes that result in polyphenisms (see review Evans and Wheeler 2001). In comparison to other animal groups many insect species show a broad range of polyphenisms especially due to environmental variation during larval development. One classical example involves the dimorphic oak caterpillars, whose morphologies differ in response to diet and time of the year (Greene 1996). Moreover, caste systems of social insects offer a variety of opportunities to study polyphenisms and the related phenomena of phenotypic plasticity. In honeybees the worker and queen castes show an alternative polyphenism with both forms undergoing a complete metamorphosis. Differences between both
1. Introduction

Involve aging, behaviour, physiology and anatomy of different organs, but differences in the reproductive development are the most striking ones. Beside the queen’s capability to mate with a maximum of 90 drones on their nuptial flights (see chapter 1.2.) where the sperm is stored in their well developed spermatheca (Snodgrass and Erickson 1992), they are furthermore able to lay many thousands of fertilized as well as unfertilized eggs over several years (Ribbands 1953). Queen ovaries each consists of 160-180 ovarioles whereas worker ovaries contain 12 ovarioles at most (Snodgrass 1956). These morphological differences are established early during larvae development. The reduction of the number of ovarioles in workers occurs during the prepupal stage, whereas the retrogression of the spermatheca takes place during the pupal stadium (Zander et al. 1916). But in contrast to several ants species where worker reproduction is completely absent due to the complete reduction of ovaries (Wilson 1971; Oster and Wilson 1978; Fletcher and Ross 1985), worker ovaries in honeybees are not rudimentary. They produce oocytes that are normally resorbed (Wigglesworth 1954) but can be activated under queen- and broodless conditions (Ribbands 1953). Moreover, queens possess larger mandibular glands. They are in charge of producing the queen pheromone 9-Oxo-2-decenoid-Acid (9-ODA) a major pheromone to suppress worker ovary activation (Butler 1957).

Since honeybee workers and queens derive from totipotent diploid larva, dimorphism in different castes of the honeybee is not a consequence of different genotypes (Shuel and Dixon 1960), but exclusively depends on the nutritional control during larval development. Whereas honeybee worker larvae receive 5 mg royal jelly in their diet, queen destined larvae are fed an excessive amount of royal jelly of up to 300 mg (reviewed in Rembold 1964). Moreover, caste dichotomy is not just a matter of quantity but quality, i.e. food composition. Whereas larvae of both caste destinations get unrestricted quantities of a highly nutritious diet in the first phase, worker destined larvae get a diet with low protein and a high carbohydrate concentration because the nurse bees add honey during the second developmental phase which spans the third and fourth day of larval development. The feeding of queen larvae proceeds without a nutritional change during development (Haydak 1970 and citations within). After decades of research on the actual compound which qualitatively controls the diphenic caste differentiation (e.g. Haydak 1970 and citations within; Rembold et al. 1974; Bíliková et al. 2002; Furusawa et al. 2008) a 57 kDA protein (Royalactin) has been suggested to be the nutritional switch which decides whether totipotent larvae become
queens or workers (Kamakura 2011). Larvae fed with denatured royal jelly with 0.5-2 % w/w Royalactin (Rol) added showed a shortened developmental time, increased weight at adult emergence and an increased ovary size. Moreover, this diet induced larvae to develop into queens as effectively as royal jelly.

Figure 1. Queen destined larvae feeding on royal jelly in opened queen cells (photograph from Wikipedia; taken by Waugsberg).

However, the larvae fed with the denaturated royal jelly only also showed increased weight compared to natural workers and developed into intercaste individuals, suggesting that more than a single protein compound is essential for caste determination. The restricted diet also led to increased titres of juvenile hormone (JH). JH is known to increase due to the differential nutrition at the fourth larval instar and causes the development into a queen (Goewie 1978; Weaver 1957; Asencot and Lensky 1976; Wirtz and Beetsma 1972; Bloch et al. 2002). In the absence of sufficient JH larvae develop into workers (Wirtz 1973). Worker destined larvae treated with exogenous JH during the late fourth and fifth instars show gyne characters (Wirtz 1973; Copijn et al. 1979; Goewie 1978). This suggests, that during the JH sensitive period, either the queen or worker developmental program is initiated, respectively that the developmental pathways of both castes diverge at this point, leading to the morphological and physiological traits that distinguish them (de Wilde and Beetsma 1982; Rachinsky and Hartfelder 1990). Moreover, Rol also activates the Map-kinase pathway down-stream of the epidermal growth factor which is responsible for ecdysteroid 20-hydroxyecdysone synthesis (Kamakura 2011). In interplay with JH the ecdysteroid makisterone A was
1. Introduction

found to be an additional critical factor for caste development and metamorphosis in the last larval instar (Feldlaufer et al. 1985; Rachinsky and Engels 1995).

1.4. Intraspecific reproductive parasitism in honeybees

Research providing more detailed understanding of the genetic mechanisms controlling conflicts and cooperation in eusocial insects systems is dependent on exceptional, mutant individuals. There are two exceptional systems in A. mellifera where workers disunite and disturb the often-cited social harmony.

Oldroyd and colleagues (1994) report on a specific behavioural trait in honeybees – anarchy. In human societies anarchy is defined as the breakdown of the normal social order including the avoidance of police control. Such anarchy in honeybees very rarely occurs. Worldwide just four colonies were reported to have an exceptionally high number of workers that have developed ovaries and produce large amounts of haploid eggs that develop into males even in the presence of a successful queen (Barron et al. 2001). Genetic analyses revealed, that those workers originate from just a few patrilines (Oldroyd et al. 1994; Montague and Oldroyd 1998; Châline et al. 2002), suggesting a genetic component for this behaviour. This assumption was supported by successful breeding of an anarchistic line (selected anarchists=SA) which stably exhibits 10-40 % of workers with anarchistic characteristics (Oldroyd and Osborne 1999). Conducted experiments to clear the proximate mechanisms which enable workers to become anarchists revealed SA larvae to be less effective at preventing ovary activation than wild-type (WT) larvae (Oldroyd et al. 2001). In contrast, mandibular pheromones of SA queens do not differ from those of WT queens in both the chemical composition and the attractiveness for fellow workers (Hoover et al. 2005 A and B). For successful rearing of worker produced males anarchistic workers must have evolved strategies to avoid worker policing. Indeed, SA workers are less successful in discriminating between queen and worker laid eggs than WT workers (Oldroyd and Ratnieks 2000). Additionally, SA workers produce eggs that are more often accepted than those laid by queenless WT workers. But eggs of anarchistic queens are removed at higher rates as WT queen laid eggs (Beekman et al. 2007). Moreover, the selfish activation of ovaries and the production of eggs with lower removal rates are genetically unrelated (Oldroyd and Osborne 1999). Two independent mutations, very rare and probably costly when they occur separately and the costly removal of queen laid eggs might explain why anarchy is so rare in honeybees (Beekman and Oldroyd 2008).
Figure 2. Laying *Apis mellifera capensis* worker surrounded by a court of *Apis mellifera scutellata* workers (photograph by Stephan Härtel).

Cape honeybees (*Apis mellifera capensis*) represent another unique exception that provides a keyhole to study the mechanisms that normally keep workers functionally sterile. Cape honeybees are native to the fynbos biome of the southern tip of South-Africa (Hepburn and Crewe 1991). Workers of this subspecies have the unique ability to produce diploid female offspring (Onions 1912; Anderson 1963) by thelytokous parthenogenesis (Verma and Ruttner 1983). The restoration of egg-diploidy is achieved by a fusion of the central haploid nuclei after meiosis, a process termed automixis with central fusion (Verma and Ruttner 1983). This mechanism increases homozygosity of loci situated between the chiasmata and the telomeres. In honeybees such an increased homozygosity represents a potential for lethality due to the sex determining system, where the zygosity of a single locus decides on the sex of the zygotes (Whiting 1943; Beye et al. 2003). However, meiosis of Cape honeybee workers shows an extremely low frequency or even absence of crossing over (Moritz and Haberl 1994; Baudry et al. 2004). Thus the thelytokous way of reproduction causes relatedness by unity between laying workers and their offspring (r=1) which is much higher than among supersisters (r=0.75). This altered kin structure favours selection of reproductive workers (Greef 1996). These workers increase their individual reproductive success by producing clonal female off-spring which in turn might replace the queen (Jordan et al. 2008a) or requeen queenless foreign *A. m. capensis* colonies (Moritz et al. 2011; Holmes et al. 2010).
unique trait of thelytokous production of female clonal offspring favoured selection for intracolonial competitiveness among subfamilies to produce the next generation of queens (Moritz et al. 1996; Beekman and Oldroyd 2008). This could have led to an evolutionary arms race that resulted in different traits associated with reproduction that are now characteristic for *A. m. capensis*. Laying Cape honeybee workers show considerable longevity of five months or even more (Velthuis et al. 1990; Tribe and Allsopp 2001) which is five times longer than the life expectancy of non-reproductive workers (Winston 1987). A variable proportion of *A. m. capensis* possess a spermatheca, the organ where sperm is stored after the queens’ nuptial flights in all other subspecies (Hepburn and Crewe 1991; Phiancharoeng et al. 2010). Moreover, Capensis workers solicit larger quantities of larval food when fed by other subspecies resulting in workers with a higher reproductive potential (Calis et al. 2002; Allsopp et al. 2003). Indeed, Capensis workers are faster in activating their ovaries and in producing off-spring (reviewed by Neumann and Hepburn 2002 and Neumann and Moritz 2002) and are able to reproduce in the presence of a laying queen much more frequent than other subspecies (e.g. Pettey 1922; Pirk et al. 2002; Jordan et al. 2008a; Beekman et al. 2009). Furthermore, reproductive Cape honeybee workers produce pheromonal compounds normally only found in queens (Plettner et al. 1993; Crewe and Velthuis 1980), even under queenright conditions (Reece 2002; Moritz et al. 2002). Studies on geographical variations in the ratio of queen- to worker specific compounds in mandibular glands in Cape honeybee workers revealed a dominance of the precursor of the major compound of the queen pheromone even in the presence of the queen (Zheng et al. 2010). This suggests that Cape honeybee workers are primed for reproduction and, in combination with their ability to lay unfertilized diploid eggs, predisposed for social parasitism (Beekman and Oldroyd 2008; Zheng et al. 2010). This predisposition is enhanced by several behavioural traits. Cape bee workers disperse into foreign colonies more frequently then other subspecies (Neumann et al. 2001). Moreover, once in a foreign colony Cape honeybee workers are preferentially fed (Beekman et al. 2000) and are more often found in areas away from the queen (Moritz et al. 2001, 2002; Neumann et al. 2003a).

Indeed, such intraspecific social parasitism can be observed in *A. m. scutellata* a subspecies which does not have the thelytokous mode of parthenogenesis (Radloff and Hepburn 2000). Both species are separated by a hybrid zone with Capensis confined to the southernmost part and Scutellata located throughout the rest of South-Africa and
countries to its north (Radloff and Hepburn 2000). Migratory beekeeping which
introduced Cape bees into the endemic regions of Scutellata has been associated with a
sharp decline of Scutellata colonies in apiculture known as the “Capensis calamity” of
1990 (Allsopp 1992, 1993). This was due to a massive invasion of parasitizing Capensis
workers which turned out to be clonal descendents of a single parasitic worker (Baudry
et al. 2004). Parasitic thelytokous workers invade queenright foreign colonies and
establish themselves as pseudo-queens (Sakagami 1958; Moritz et al. 1996). As soon as
many parasites are within the colony, the host queen gets lost probably due to lethal
fights with parasitic pseudoqueens (Moritz et al. 2003). The majority of A. m. scutellata
workers are replaced by parasitic workers. A regular brood rearing and foraging ceases
(Martin et al. 2002b) which eventually results in the death of the host colony (Neumann
and Hepburn 2002).

Parasitizing is possible because the host workers do not recognize the absence of
the queen since the parasitic workers produce queen-like amounts of the queen
pheromonal blend in the mandibular glands (Ruttner et al. 1976; Hemmling 1979;
Crewe and Velthuis 1980; Wossler 2002; Simon et al. 2001; Dietemann et al. 2007), the
tergal glands (Wossler and Crewe 1999 a; b) and the Dufours gland (Sole et al. 2002;
Martin and Jones 2004). This queen like pheromonal bouquet suppresses queen rearing
and ovary activation in other workers (Hepburn et al. 1988) and releases retinue
behaviour (Anderson 1968). Furthermore, A. m. capensis pseudoqueens suppress the
production of queen-like pheromones in other workers (Moritz et al. 2000). Moreover,
policing is less frequent than in normal honeybee colonies (Moritz et al. 1999; Martin et
al. 2002c; Pirk et al. 2003) leading to a lower egg removal rate of eggs laid by Capensis
workers comparable to those eggs laid by Capensis queens (Calis et al. 2003). This
suggests that eggs laid by Capensis workers are indistinguishable from those laid by
Capensis queens (Calis et al. 2003).

1.5. Genetic determination of reproductive hierarchies in
honeybees

Insect workers all express genes that reduce their individual reproductive success
by suppressing evolutionary favoured behaviours of parental care and egg-production
and by increasing alloparental care and sterility instead. The identification of genes that
regulate this shift and the characterisation of its regulation during development and in
different environments would substantially enlarge the understanding of the defining
characteristic of eusociality. Indeed, considerable research in ants (Sameshima et al.
2004; Tian et al. 2004; Anderson et al. 2008 and citations within; Gotzek and Ross 2009
and citations within), wasps (Sumner et al. 2006; Hoffman and Goodisman 2007; Toth
et al. 2007; Hunt et al. 2010), termites (Scharf et al. 2003; 2005; Zhou et al. 2006 a and
b; Lienard et al. 2006; Korb et al. 2009; Weil et al. 2009) and bees (reviewed in Smith et
al. 2008; Hepperle and Hartfelder 2001; Pereboom et al. 2005; Hartfelder et al. 2006; de
Azevedo and Hartfelder 2008; Humann and Hartfelder 2011) has focussed on
identifying genes that are differentially expressed between castes. But the key
regulatory genes for ovary activation are apparently difficult to detect by such studies,
because caste differences are all-embracing the physiology, anatomy and behaviour that
expression differences are expected all across the genome. Thus transcriptome data will
primarily provide a phenotype at a level more closely linked to the genome but still at a
level of complexity preventing causal analyses of gene cascades determining caste.

Indeed, a huge set of honeybee genes show different expression patterns between
queens and workers. In microarray studies on brains of same-aged virgin queens,
reproductive and non-reproductive workers more than 2000 genes were differentially
expressed between queens and workers (Grozinger et al. 2007). A much smaller
proportion of genes showed different transcript abundances between the evaluated
worker groups (221), whereas the gene expression in reproductive workers became
more queen-like. This differential gene expression between reproductive and non-
reproductive workers may be driven by a different exposure to the queen’s mandibular
pheromonal (QMP) bouquet. Studies revealed that exposure with QMP resulted in
extensive changes in the expression of several hundreds of genes in the brains of young
workers (Grozinger et al. 2003).

According to the caste specific phenotypical differences, several classes of genes
are predicted to be differentially expressed between the worker and the queen castes.
Nutritional control of queen development and hence a different metabolism of queen
destined larvae as well as queen longevity suggest physiometabolic genes associated
with oxidative stress responses, growth and development to be more abundant in the
queen caste. In fact, this has been shown for larvae of worker honeybees (Corona et al.
1999; Evans and Wheeler 1999, 2001; Cristino et al. 2006; Barchuk et al. 2007) and
bumble bees (Pereboom et al. 2005) and in adult stingless bees (Judice et al. 2004;
2006) and honeybees (Corona et al. 2005; Grozinger et al. 2007). Interestingly
microarray based expression patterns suggested that workers and younger, bipotential
larvae share a similar gene expression pattern, whereas queen destined larvae express different genes reflecting predicted hormonal effects (Evans and Wheeler 2000).

A second group of genes associated with insulin signalling was proposed to be associated with the phenotypic plasticity of the queen. Insulin signalling regulates growth, longevity and reproduction in multiple species (reviewed in Wu and Brown 2006). Differential expression of putative insulin-pathway associated genes has been identified in honeybee worker and queen larvae (Wheeler et al., 2006; Patel et al. 2007; de Azevedo and Hartfelder 2008; Wolschin et al. 2011) and adults (Corona et al. 2007). Nevertheless, diets containing the putative caste developmental protein Royalactin (Rol) in bees with decreased transcripts of the insulin receptor did not show an altered final adult size, a decreased developmental time or a change of the ovary size (Kamakura 2011). This suggests that insulin signalling is not involved in the signalling cascades leading to caste differentiation, at least not with Rol as an effector molecule. But irrespective of the actual signalling cascades, epigenetic modification is very clearly involved in translating the nutritional status into different expression of genes (Kucharski et al. 2008; Elango et al. 2009; Lyko et al. 2010).

Nevertheless, apart from attempts to knock-down genes in order to learn about their functionality (Patel et al. 2007; Kamakura et al. 2011; Wolschin et al. 2011) all of the gene expression studies do not pass the constraints of correlation. An alternative attempt (Amdam et al. 2004) to characterise putative gene networks associated with reproductive potential of workers was based on the reproductive groundplan hypothesis (RGPH). The original RGPH stated by West-Eberhard (1996) proposed that ovarian development, brood care, and foraging are uncoupled in eusocial insects, such that ovary development is maintained in queens, brood care is performed in young workers (which can also be competent in activating their ovaries) and foraging is performed by older workers, which have the lowest capacity to become reproductively active. Thus queens and nurse bees should be physiologically more equal than foragers. The altered RGPH (Amdam et al. 2004) suggests variance in worker behaviour to be emerged from pre-existing mechanisms associated with female reproduction. Amdam et al (2004) postulate regulatory gene and hormone networks that pleiotropically control both reproductive characters and foraging. Artificial selection for colonies with high and low pollen foraging behaviour was used reveal the ancestral gene networks that link ovarian and yolk protein physiology to behaviour (Amdam and Page 2010). Indeed colonies selected to forage more pollen (Page et al. 1991; 1995; 1998) have workers with larger ovaries.
1. Introduction

and the pollen foraging bias correlates with ovary size (Amdam et al. 2004, 2006). Furthermore, loci associated with foraging behaviour of honeybees (Hunt et al. 1995; Page et al. 2000; Rueppel et al. 2004 a,b; 2009) were analysed by behavioural and physiological phenotyping and by candidate gene expression studies in backcross colonies selected for a higher foraging rate. Out of this study two candidate genes (PDK1 and HR46) were suggested to be genetically linked to foraging behaviour and to ovary size simultaneously (Wang et al. 2009). However, studies on anarchistic bees, which served as an alternative route to test the altered RGP hypothesis, suggest this link between foraging and reproduction may be rather strain specific. According to RGPH anarchistic worker bees must forage for pollen and nectar with a low sugar concentration and forage earlier in life. However, unfortunately anarchistic workers have a delayed onset of foraging and show no preference for nectar and pollen (Oldroyd and Beekman 2008). In the light of these divergent results, it appears that the correlations between the studied traits suggested by RGPH may not present a general pattern in honeybees (Oldroyd and Beekman 2008).

A more promising attempt to unravel genes controlling worker sterility involves rare mutant worker phenotypes which conduct intraspecific parasitism. Gene expression studies in heads and abdomens of anarchistic worker bees revealed two genes coding for major royal jelly proteins 2 and 7, and one gene coding for a protein which is involved in a hereditable nutritional disease in humans to be significantly higher expressed in heads of wild-type bees (Thompson et al. 2006). In abdomens of reproductive anarchist workers genes for the yolk protein vitellogenin, venom peptides and a member of the AdoHycase superfamily, among others were up-regulated (Thompson et al. 2008). The very same strain also provided the opportunity to identify candidate genes involved in regulating worker sterility by Quantitative Trait Loci (QTL) mapping (Oxley et al. 2008). A selected anarchist line was used to create a backcross queenright colony that segregated for high and low levels of ovary activation and were successfully used to identify four QTLs that together explained approximately 25 % of the phenotypic variance for ovary activation. They have been suggested to influence a workers’ propensity to activate her ovaries in the absence of pheromonal suppression due to a suggested yet-unmapped Anarchy mutation.

Just like the thelytokous reproduction in the solitary parasitic wasp Lysiphlebus fabarum where the restoration of egg-diploidy is genetically equivalent to a central fusion automixis in Cape bees (Sandrock and Vorburger 2011), thelytoky in Cape
1. Introduction

Honeybees was shown to be inherited by a single allele (Ruttner 1988). Thelytoky as well as other reproductive traits (the production of queenlike amounts of 9-ODA and the early onset of egg-laying) all of them characterising thelytokous pseudoqueens have been shown to be under control of a single recessive locus (Lattorff et al., 2005), which is located on chromosome 13 and spans a region of 11.4 cM (Lattorff et al. 2007).

![Figure 1.3](image)

**Figure 1.3.** Location and candidate genes within the mapped \(th\) region on chromosome 13 (based on Amel 4.5). Associated microsatellite markers are given in the middle. Homologies and putative functions of the 15 candidate genes are given in the table. Italic marking stands for the former functional classification of the respective gene in Amel 4.0 (Consortium HGS 2006). Bold marking represents the most promising candidates. (Based on figure 1 from chapter 4).

An almost complete co-segregation between the production of 9-ODA and the early onset of egg-laying with the type of parthenogenesis strongly suggests the same locus to pleiotropically control all of these reproductive traits. The \(th\) locus comprises 15 annotated genes (Amel 4.5; NCBI). The most promising two candidate genes code for transcription factors (ATF-2 (XM_393896; NCBI) and CP2 (XM_001121158; NCBI)).
XM_393898; NCBI), which are feasible to control several traits simultaneously (Courey 2008). Another interesting candidate gene is locus410417 (XM_393897; NCBI) which codes for a receptor protein tyrosin phosphatase (RPTP). The RPTPs form a subfamily of the classical protein-tyrosine phosphatases (PTPs) that exclusively dephosphorylate phosphotyrosine (pTyr) in target proteins (den Hertog 2004). One of such PTPs (PTP-1B) has a major role in modulating insulin sensitivity (Elchebly et al. 1999) in mice, probably by dephosphorylating the β-subunit of the insulin receptor (IR, Lammers 1997). IR homologues in Drosophila melanogaster are central regulators of growth, metabolism, longevity and reproduction (Garafalo 2002 and citations within) and may therefore be essential in the control of reproductive hierarchies in honeybees as well.

1.6. Aims of this work

Queens and workers, although substantially different in their anatomy and physiology, share the same genome. The same holds true for reproductive and non-reproductive workers. This thesis reports on using a previously mapped region on chromosome 13 in the A. mellifera genome comprising the putative th-locus in a set of 15 candidate genes that was shown to pleiotropically control several reproductive traits (early onset of egg-laying; queen-like production of the queen pheromone 9-ODA) as well as the mode of parthenogenesis (thelytoky). Because the reproductive traits tested also ensure queen-dominance and reproductive success of workers in normal arrhenotokous colonies, the th locus region may reveal a major genetic switch to regulate altruism in a eusocial colony. I here embark on narrowing down the relevant genes among the 15 candidates by gene expression analyses in different castes and both, arrhenotokous and thelytokous laying workers. This will be underlined by knock-down studies by RNA-Interference. After identification of the most promising candidate gene its actual mechanism and the thelytoky causing allele shall be examined by sequencing in order to gain insights into a putative molecular motif that provides the very basis for eusociality - altruistic worker sterility.

1.7. References


1. Introduction


1. Introduction


1. Introduction


1. Introduction


1. Introduction


1. Introduction


2. Systemic RNA-interference in the honeybee *Apis mellifera*: Tissue dependent uptake of fluorescent siRNA after intra-abdominal application observed by laser-scanning microscopy

A. Jarosch\textsuperscript{a}, R.F.A. Moritz\textsuperscript{a,b}

\textsuperscript{a} Institut für Biologie, Martin-Luther-Universität Halle-Wittenberg, Hoher Weg 4, 06120 Halle (Saale), Germany

\textsuperscript{b} Department of Zoology, University of Pretoria, Pretoria 0002, South Africa

*Journal of Insect Physiology* 57, 851–857 (2011)
doi:10.1016/j.jinsphys.2011.03.013

Received 1 December 2010; revised 10 March 2011; Accepted 15 March 2011.
Abstract

RNA interference has been successfully used in adult honeybees, but there are only few reports about abdominal application of dsRNA/siRNA which have reached more distant tissues than the fat body. We studied systemic RNAi in honeybees by injecting fluorescent siRNA of the ubiquitously expressed honeybee homologue of the Glycerol-3-Phosphate Dehydrogenase (amGpdh) into the abdomens of adult bees and followed them by laser scanning microscopy and qPCR. The fat body was the sole tissue emitting fluorescence and showing a decreased gene expression, whereas the siRNA had apparently not reached the other tissues. Therefore, we conclude that certain genes in other tissues than the fat body cannot be easily reached by injecting siRNA into the body cavity. In particular, the lack of amGpdh knock down in ovaries after amGpdh dsRNA injection, supports that in some cases it may be particularly difficult to interfere with gene expression in ovaries by intra-abdominal injection. In these cases alternative inhibition techniques may be required to achieve an organismic non-lethal disruption of transcription.

Keywords: RNA-interference, Honeybees, Laser scanning microscopy, Fluorescent siRNA, Tissue composition
1. Introduction

The honeybee, *Apis mellifera*, has been established as a highly valuable model organism for numerous biological disciplines. Especially with the availability of its complete genome (Honey Bee genome Sequencing Consortium, 2006) genetic tools are increasingly adapted to facilitate the study of the genetic basis of eusociality. One of the most promising techniques – RNA interference (RNAi) – has been introduced to study gene functions related to caste differentiation (Kucharski et al., 2008), sex determination (Beye et al., 2003), immune response (Aronstein and Saldivar, 2005; Schlüns and Crozier, 2007) and aging (Seehuus et al., 2006). This technique is especially important for genetic model systems like the honeybee, where the recovery of mutants is not feasible.

RNAi is a cellular mechanism leading to the knock-down of gene expression by mRNA degradation triggered by target specific double-stranded RNA (dsRNA) (Fire et al., 1998). Naturally, RNAi is a physiological mechanism which serves as a defence against mutation causing mobile DNA elements or viruses (Zamore, 2001; Obbard et al., 2009). Since the discovery of its mechanism RNAi has become a dominant reverse genetic method for the study of gene functions. Moreover, RNAi plays an increasing role in therapeutics and in pest control (Maori et al., 2009; Liu et al., 2010). Initially established for *Caenorhabditis elegans*, RNAi has been successfully implemented in various insect systems including the mosquito *Anopheles gambiae* (Blandin et al., 2002; Boisson et al., 2006), the wasp *Nasonia vitripennis* (Lynch and Desplan, 2006), the large milkweed bug *Oncopeltus fasciatus* (Liu and Kaufman, 2004), the silkworm *Bombyx mori* (Tabunoki, 2004; Ohnishi et al., 2006) and the honeybee *A. mellifera* (Beye et al., 2002) among many others.

In organisms showing a systemic RNAi response, the application of dsRNA triggers the silencing of homologous mRNA far away from the application site. This is of particular value when dissecting the molecular mechanisms of genes whose expression persists over a long time and spans the whole body. In insects systemic RNAi has for example been successfully applied in grasshoppers (Dong and Friedrich, 2005) and termites (Zhou et al., 2008; Korb et al., 2009), whereas other insect systems like the fruit fly *Drosophila melanogaster* (Rognant et al., 2003) fail to show a robust systemic RNAi response. Molecular mechanisms of systemic RNAi in insects have been excessively studied in *Tribolium castaneum*, another model organism with a completely sequenced genome (Tribolium Genome Sequencing Consortium, 2008). The injection of
dsRNA into *Tribolium* larvae causes RNAi effects throughout the entire body persisting up to the adult stage (Tomoyasu and Denell, 2004). However, genes which are required for a systemic response in *Tribolium* appear to be species specific. Genes facilitating systemic RNAi here are different to those facilitating systemic RNAi in *C. elegans*, the best evaluated RNAi model organism (Tomoyasu et al., 2008). For example SID-1, the best described membrane-spanning protein responsible for dsRNA uptake in *C. elegans* (Winston et al., 2002; Feinberg and Hunter, 2003), is not involved in *Tribolium* suggesting another putative role of SID-1 in insects.

In honeybees, *A. mellifera*, only few studies report about successful systemic RNAi. Several studies report that the application of dsRNA to eggs and larvae whether by injection (Aronstein and Saldivar, 2005; Beye et al., 2002; Maleszka et al., 2007) or by ingestion (Aronstein et al., 2006; Patel et al., 2007; Kucharski et al., 2008; Nunes and Simões, 2009; Liu et al., 2010) cause a knockdown of gene expression even in the adult. Very few studies report on successful and persisting manipulation of adult bees (Amdam et al., 2003; Farooqui, 2004; Seehuus et al., 2006; Schlüns and Crozier, 2007; Maori et al., 2009; Paldi et al., 2010).

Nevertheless, the uptake of RNAi triggering molecules into specific tissues has not been shown so far for the honeybee. Radioactive labelled dsRNA molecules disappear from the haemolymph within few hours (Gatehouse et al., 2004), but the fate of the dsRNA remains unclear. Therefore we studied the fate of fluorescently labelled short-interfering RNA molecules in selected tissues after the application into the body cavity. In particular, we evaluated the uptake of siRNA and dsRNA into the ovaries, since they are the central organ for the female caste in the honeybee. A successful and stable manipulation of gene expression in the ovaries would greatly facilitate the study of caste differentiation and reproductive dominance in the honeybee colony.

2. Materials and methods

2.1. Production of dsRNA

We cloned the *amGpdh* part chosen for down-regulation into pGem-T easy vectors (Promega) to obtain templates for dsRNA production. The *amGpdh* fragment was amplified by standard PCRs using amGPDHI: 5’-GCT GGT TTC ATC GAT GGT TT-3’ and amGPDHII: 5’-ACG ATT TCG ACC ACC GTA AC-3’. In order to avoid a mixture of different PCR products due to the presence of different *Apis Gpdh* isoforms, we chose two primers showing no sequence homology to any other gene (NCBI blast
adapted for short sequence input) in a region lacking conserved domains (dsGpdh position within the *amGpdh* gene: 636–816). Furthermore, we confirmed the product identity by direct sequencing. One of the obtained *amGpdh* clones as well as the *GFP* encoding sequence of the pGFP vector (GenBank ID: U17997; Clontech) were used for PCRs producing the dsRNA templates. These PCR were adapted to the BiothermTM DNA Polymerase (Genecraft), using 0.2 mM dNTPs, 0.3 µM of T7-promotor added primer (T7-promotor underlined: GFPI: 5´-TAA TAC GAC TCA CTA TAG GGC GAT TTC ATG GCC AAC ACT TGT CC-3´; GFPII: 5´-TAA TAC GAC TCA CTA TAG GGC GAT CAA GAA GGA CCA TGT GGT C-3´; GpdhT7I: 5´-TAA TAC GAC TCA CTA TAG GGC GAT CAA GAA GGA CCA TGT GGT C-3´; GpdhT7II: 5´-TAA TAC GAC TCA CTA TAG GGC GAT CAA GAA GGA CCA TGT GGT C-3´) and 5 U Taq Polymerase in a total reaction volume of 100 µl. PCR protocols consisted of 5 min DNA denaturation and Taq activation at 95 °C, followed by 40 cycles of 30 s at 95 °C, 30 s at 56 °C for *GFP* and 54 °C for *amGpdh* and 1 min at 72 °C. A final extension of 20 min at 72 °C completed the protocol. The resulting PCR-products were purified with the QIAquick PCR Purification Kit (Qiagen). Subsequently, dsRNA from *GFP* and *amGpdh* was derived using the T7 RibomaxTM Express RNAi System (Promega) with an extended transcription time of 5 h at 32 °C. The resulting dsRNA was purified by a Qiazol- Chloroform-treatment and the pellet resolved in Nuclease free water. The dsRNA quality was verified in 1.8% agarose gels and its concentration quantified photometrically. dsRNA concentrations were adjusted to 5 µg/µl by diluting with insect ringer (see Section 2.3) right before the injection. The resulting dsRNA molecules had a length of 503 bp (*GFP*) and 180 bp (*amGpdh*).

### 2.2. Design and production of fluorescent siRNA

We designed siRNA of *amGpdh* with the Block-iT™ RNAi Designer (Invitrogen). The derived siRNA was produced and tested for in vivo studies by Invitrogen. It was labelled with Alexa-488 and had a length of 21 bp. (siRNA sequence: 5´-Alexa 488–GCA GAU CUA AUU GCA ACU U [dT] [dT] 3´). siRNA specificity was checked by BLAST analysis (NCBI), especially focusing on different *Gpdh* isoforms found in *A. mellifera*. Additionally, we designed the siRNA in the very same region of the previously used dsRNA. Hence the siRNA was in a region of *amGpdh* lacking conserved domains (position within *amGpdh*: 777–796).
2. Systemic RNAi in honeybees

2.3. Injection and incubation

The age defined honeybee workers used for RNA injections were obtained from *A. m. carnica* colonies kept at the apiary of the Martin Luther-University in Halle/Saale. Brood combs were incubated at 34 °C and 60% humidity and freshly emerged workers collected daily. 20 newly emerged workers each were injected with either 5 µg dsGFP RNA or 5 µg dsGpdh RNA between the 5th and 6th abdominal segment using a 10 µl microsyringe (Hamilton). For siRNA injections workers were injected with 3 µg or 6 µg siRNA. For the microscopy experiment and the gene knock down analysis, three and respectively ten replicates were injected. Negative controls were injected with insect ringer (n = 9; 54 mM NaCl; 24 mM KCl; 7 mM CaCl₂ x 2H₂O). Injected workers were kept on wax plates until they recovered and bees not showing haemolymph leakage were kept together with 25 nurse bees at 34 °C with food and water ad libitum. The worker tissues were prepared after 24 h.

*A. m. carnica* queens were reared with standard apicultural techniques. They were treated twice to CO₂ on two consecutive days to induce ovary activation (Mackensen, 1947). They were held in cages with about 50 workers and injected with 3 µg siRNA after five days.

2.4. Tissue preparation and microscopy

All bees were sacrificed in liquid nitrogen and stored at -80 °C until preparation. Ovaries and fat bodies were prepared from the abdomen, flight muscles from the thorax, brains and hypopharyngeal glands from the head. Tissues destined for RNA preparation were put in RNAlater (Ambion) and stored at -80 °C until the RNA preparation; tissues destined for microscopy were placed on slides flooded with glycerol and McIlvaine’s buffer (0.1 mol/l citric acid; 0.2 mol/l Na₂HPO₄; pH 7.0) with a ratio of 1:2. They were covered with a cover slip and sealed with nail varnish. Preparation under dark conditions and storage of the samples for less then two hours avoided a decrease of the fluorescence. Additionally, samples were kept in the dark and on dry ice until examination under an inverted Carl Zeiss LSM 510 microscope with a Plan-Neofluar objective. Scans were analysed using the LSM510/4.2. Image browser software package (Zeiss).
2.5. RNA preparation and real-time measurements of amGpdh gene expression

For analysing the transcript level of *amGpdh* in different tissues of *A. mellifera carnica*, newly emerged workers were held in cages right after emergence and incubated with food and water ad libitum at 34 °C. After 24 h they were sacrificed in liquid nitrogen. Tissues from these workers as well as tissues from injected workers were manually homogenised using plastic pestles. RNA extraction followed the standard Trizol (Invitrogen) protocol (Chomczynski and Sacchi, 1987) with subsequent DNAse (Promega) digestion. RNA quality and quantity was assessed by photometric measurements. Equal amounts of RNA were immediately reverse transcribed with M-MLV H-Point Mutant Reverse Transcriptase (Promega) using oligo-dT Primer (0.5 µg/µl; Promega) according to the manufacturer’s instructions. For gene expression studies Sybr-Green assays were run using 5 µl iQ SYBR Green Supermix (Biorad), 1 µl template, 1 µl Primer (1 µM) in a 10 µl reaction volume. The real-time PCR cycle profile started with a 3 min incubation phase at 95 °C for Taq activation, followed by 39 cycles of 15 s at 95 °C and 30 s at 54 °C for annealing, 30 s at 72 °C for extension and data collection. Each sample was run in duplicate with a subsequent melting curve analysis between 50 °C and 90 °C, reading the fluorescence at 1 °C increments. The purity of the PCR products was additionally checked on 1.8% agarose gels. C(t) values were calculated by the Opticon Monitor 3 software (Biorad) using a single standard deviation over cycle range. Primers for *rp49* (NCBI ID: NM_001011587) were: rp49I 5´-TCGTCACCAGAGTGATCGTT-3´; rp49II 5’-CCATGAGCAATTT CAGCACA-3´. Primers for *amGpdh* were GpdhI 5´-GGATCAGGAAATTGGGGTTC-3´; GpdhII 5´-CGGAAGCTTATGTCCTGGAA-3´. Primer specificity was confirmed by sequencing the PCR products for every tissue. Primer positions for *amGpdh* amplification were outside the RNAi targeted mRNA sequence (position of the PCR product within *amGpdh*: 30–213).

2.6. Data analysis

The honeybee ortholog of the ribosomal protein 49 (*rp49*) was used as a stable house-keeping gene to calculate the relative gene expression of *amGpdh* (Grozinger et al., 2003). The PCR efficiency for every sample was calculated from the linear phase of fluorescence increase due to target duplication to control for different PCR efficiencies within different samples and different genes (Pecced and Jacob, 1996; Pfaffl, 2001a). Relative gene expressions were calculated according to Pfaffl (2001b).
3. Results

3.1. amGpdh expression in different tissues

To assess if the honeybee homologue of the Glycerol-3-Phosphate Dehydrogenase amGpdh is suitable for analysing systemic RNAi in the honeybee, we quantified its relative gene expression in all of the evaluated tissues. Although amGpdh is higher and more variable expressed in fat body cells of different workers, it is ubiquitously expressed in all screened tissues (Fig. 1). Therefore we chose this gene to study the tissue dependent RNAi knock down after the application of gene specific dsRNA as well as siRNA.

![Fig. 1. Relative gene expression of amGpdh in different tissues of the honeybee Apis mellifera. Missing significant differences (p = 0.8; one-way ANOVA of log transformed data) between transcript levels show amGpdh to be ubiquitously expressed throughout the evaluated tissues.](image)

3.2. dsRNA does not affect gene expression of amGpdh in honeybee ovaries

The application of amGpdh-specific dsRNA into the abdomen caused no knock down of amGpdh in the ovaries. All test groups showed similar amounts of amGpdh transcripts irrespective of the treatment (Fig. 2; p = 0.35 one-way ANOVA of log-transformed data). In contrast, the gene expression of amGpdh was significantly down regulated after the treatment with dsGPDH in the fat body (p ≤ 0.001 one-way ANOVA)
of log-transformed data). Even the injection of the inject ringer control led to a significant reduction of amGpdh transcripts in the fat body compared to the untreated bees.

Fig. 2. Relative gene expression of amGpdh in honeybee fat bodies and ovaries either untreated, injected with insect ringer or dsGpdh. Expression data are standardized to the expression of amGpdh in untreated bees. Different letters indicate significant differences between the groups (Newman–Keuls post hoc tests).

3.3. Laser-scanning microscopy of honeybees treated with fluorescent siRNA

Since the follicular membrane surrounding the oocytes may be a principle obstacle preventing the uptake of a 180 bp dsRNA, we injected short 21 nt siRNA with a fluorescent label into the abdomen to trace the fate of these labelled siRNAs with laser scanning microscopy.

Control ringer injected workers did not show fluorescence in the fat body apart from the background fluorescence emitted by the tissue itself. The treatment with fluorescent siRNA on the other hand yielded very strong signals in the fat body trophocytes in every siRNA treated worker, irrespective of the injected siRNA concentrations (Fig. 3). In contrast, the same treatment had no effect on the fluorescence signal in ovaries of sterile workers (Fig. 4). The same was true for every other tested
tissue. Irrespective of the injected siRNA concentration, no fluorescence signal was detected, neither in the flight muscles, the brain, nor in the pharyngeal glands.

We also followed the fate of siRNA injected into *A. mellifera carnica* queens to test for effects caused by differences in ovary size and activation between queens and workers. Like in workers, the fat body of queens was the sole tissue with a fluorescence signal after injection (Fig. 5). Detectable levels of fluorescent siRNA had not reached any of the other tested tissues, including the ovaries.

**Fig. 3.** Images of fat bodies treated with either 3 µg fluorescence labelled siRNA (A - C) or insect ringer (D – F). (C and F) Single confocal sections (A and D) are aligned with the corresponding differential interference contrast image (B and E). White bars represent 100 µm (magnification: 20x).
Fig. 4. Images of ovaries treated with either 3 µg fluorescence labelled siRNA (A – C) or insect ringer (D – F). (C and F) Single confocal sections (A and D) are aligned with the corresponding differential interference contrast image (B and E). White bars represent 100 µm (magnification: 10x).

Fig. 5. Fat body (A) and ovaries (B) of a siRNA (3 µg) treated queen. White bars represent 100 µm (magnification: 20x).
3.4. Relative gene expression of amGPDH in honeybees treated with fluorescent siGPDH

The amGpdh qPCR data confirmed the results from the laser scanning microscope. The fat body was the sole tissue showing a significant decrease in gene expression after treating adult honeybees with amGpdh specific siRNA (p = 0.02; MWU test; siRNA treatment compared to ringer injected honeybees). None of the other tissues responded to the treatment (Fig. 6), although the thoracic muscles show an overall decreased expression level of 70%. This decrease was, due to a high variance within the control group, beyond the probability value of 5% (p = 0.33; MWU test; siRNA treatment compared to ringer injected honeybees).

![Fig. 6. Relative gene expression of amGPDH in honeybees treated with amGPDH specific siRNA. Gene expressions are normalised to ringer injected bees, whose amGPDH expression was set to one. The asterisk indicates a significantly reduced gene expression compared to the ringer injected control group (p ≤ 0.05; MWU test).](image_url)

4. Discussion

RNA-interference is an important tool for manipulating transcriptional pathways and exploring functional genetics. In adult honeybees it has been used following injection protocols established in other insects. In this study we aimed at a better
understanding of the RNAi mode of function in honeybees, when using this application technique. Hence, we followed the way and uptake of injected fluorescent siRNA of *amGpdh* throughout the insect body. We show that this gene is ubiquitously expressed and therefore very well suitable for basic studies like this. Moreover, we are able to show that the designed siRNA has the power to reduce the *amGpdh* transcript level at least in fat body cells. Therefore, we expect *amGpdh* and its siRNA counterpart to be effective to reach any tissue apart from the very well studied RNAi model tissue fat body.

According to our expectations we observed a strong siRNA uptake into the trophocytes of the fat body, which supports experiments manipulating gene expression in adult honeybees by RNAi here (Amad et al., 2003; Seehuus et al., 2006). In contrast, we could not detect any fluorescence due to *amGpdh* siRNA uptake in any other tissue. This might be contradictory to studies stating that intra-abdominal injection of dsRNA or siRNA leads to gene knock down in tissues far away from the injection site (Gatehouse et al., 2004; Schlüs and Crozier, 2007). One might argue that we just missed some fluorescence, since the siRNA concentration within the haemolymph becomes diluted over distance well below the detection limit of the laser scanning microscope. These low concentrations might otherwise be sufficient to cause a knock down of gene expression when entering the cells, as only a few dsRNA molecules per cell are required to silence thousands of target mRNA molecules (Kennerdell and Carthew, 1998). To overcome this technical shortcoming we pursued two different strategies. On the one hand we chose two rather high siRNA concentrations (3 and 6 µg) to ensure that even under high dilution any uptake of fluorescence labelled siRNA remains visible. In addition, qPCR measurements, which are much more sensitive than laser scanning microscopy, support the lack of *amGpdh* knock down in any tissue other than the fat body.

Concerning the reproductive organs of honeybees, we failed to knock down gene expression with *amGpdh* specific dsRNA or siRNA and found no fluorescent siRNAs in ovary tissue (irrespective of their status). This suggests that it might be difficult to use injection as a routine application technique for RNAi in honeybee ovaries. A reason for that might be a different composition of the ovary and fat body tissue. Since the honeybee fat body is a storage, detoxification and excretory tissue (de Oliveira and Cruz-Landim, 2003) it is designed to have a high uptake of most diverse substances from the haemolymph. It contains suites of passive transporters and pores through which
the RNAi causing molecules may pass. In contrast, honeybee follicles are covered by two cellular layers. The ovariole sheath, a relic of the peritoneal sheath, completely covers all ovarioles (King et al., 1968). It may function similar to the insect neurilemma, which surrounds the brain and serves as a blood-brain barrier system by controlling the diffusion of water soluble ions and molecules (Abbott et al., 1986). In addition to the ovariole sheath, a single layer of follicular epithelial cells surrounds both oocytes and nurse cells (Engels, 1968). This follicular layer largely prevents any passive uptake of substances and only tiny, pore–like canals within the follicular layer of vitellogenic follicles allow transport into honeybee ovaries (Fleig, 1995). Moreover, active transport mechanisms facilitating yolk uptake (Engels, 1973; Fleig, 1995) into the perivitelline space between oocytes and follicular layer are known. Given our negative results of siRNA/dsRNA transport into ovaries and the high specificity of this active transport system, it seems unlikely that it is involved in the uptake of RNAi molecules from the hemolymph. Finally, other dsRNA transport mechanisms known from honeybees (amSID-1; Aronstein et al., 2006) and other organisms (ABC transporter, Timmons, 2007 or receptor mediated endocytosis, Ulvila et al., 2006; Saleh et al., 2006) are apparently also not involved or at least not effective in transporting amGpdh siRNA/dsRNA into the ovaries.

Summarizing our results, we find that the injection of a siRNA/ dsRNA specific for a ubiquitously expressed gene does not result in any gene knock down or even fluorescence increase in any tissue apart from the fat body. Although we cannot exclude, that other genes respond differently to the same treatment, our data suggest that the tissue composition has dramatic influence on the experimental outcome when using abdominal injection as application method. Conclusively, there is urgent need to improve RNAi application techniques in the honeybee model system, for example including chemically modified siRNA (review by Tseng et al., 2009). Such lipophilic molecules may overcome kinetic and physical barriers within an organism. Only if we can reach specific tissues with a robust and reliable gene knock down tool, it will be possible to unfold the full potential of the eusocial honeybee model system.

Acknowledgements

We thank the Developmental Genetics group of Prof. Reuter at the Martin-Luther-University Halle-Wittenberg and especially Manuela Gabler for handling the LSM. This study was financed by the DFG (RFAM).
References
2. Systemic RNAi in honeybees


Maori, E., Paldi, N., Shafir, S., Kalev, H., Tsur, E., Glick, E., et al., 2009. IAPV, a bee affecting virus associated with colony collapse disorder can be silenced by dsRNA ingestion. *Insect Molecular Biology* 18, 55–60.


2. Systemic RNAi in honeybees


3. RNA interference in honeybees: off-target effects caused by dsRNA

Antje Jarosch¹, Robin F. A. Moritz¹,²

¹Institut für Biologie, Martin-Luther-Universität Halle-Wittenberg, Hoher Weg 4, 06120 Halle (Saale), Germany
²Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa

Apidologie online first (2011)
DOI:10.1007/s13592-011-0092-y

Received 22 December 2010 – Revised 15 June 2011 – Accepted 8 July 2011
Abstract

RNA interference involves the targeted knockdown of mRNA triggered by complementary dsRNA molecules applied to an experimental organism. Although this technique has been successfully used in honeybees (*Apis mellifera*), it remains unclear whether the application of dsRNA leads to unintended expression knockdown in unspecific, non-targeted genes. Therefore, we studied the gene expression of four non-target genes coding for proteins that are involved in different physiological processes after treatment with three dsRNAs in two abdominal tissues. We found unspecific gene downregulation depending on both the dsRNA used and the different tissues. Hence, RNAi experiments in the honeybee require rigid controls and carefully selected dsRNA sequences to avoid misinterpretation of RNAi-derived phenotypes.

**Keywords:** RNA interference, honeybees, off-target effects, real-time PCR
1. Introduction

After the honeybee (*Apis mellifera*) became a model organism for the study of the genetic basis of eusociality, it was important to have a well established, specific system to knock down genes. Presently, knockout mutants cannot be produced in *Apis*; thus, RNA interference (RNAi) appeared as a powerful tool for such functional gene studies by inducing loss-of-function phenotypes through target complementary short double-stranded RNA (dsRNA) molecules. Since its discovery in *Caenorhabditis elegans* (Fire et al. 1998), RNAi has become the predominant reverse genetic method in a variety of non-model organisms. Moreover, as honeybees are among the few recognized beneficial insects with a large economical and ecological impact, the use of RNAi is increasingly used as a tool for fighting pests and pathogens in apiculture (Maori et al. 2009; Paldi et al. 2010). In light of the high potential power of RNAi for understanding honeybee genetics, it is surprising how few studies have been conducted using RNAi in *A. mellifera*. Some of them manipulated eggs or applied dsRNA to larvae either by feeding or injections (Aronstein and Saldivar 2005; Beye et al. 2003; Aronstein et al. 2006; Patel et al. 2007; Kucharski et al. 2008; Nunes and Simões 2009; Maori et al. 2009), whereas there are only very few reports on the successful manipulation of adult individuals (Amdam et al. 2003; Farooqui et al. 2004; Seehuus et al. 2006; Schlüns and Crozier 2007; Gatehouse et al. 2004; Müßig et al. 2010; Mustard et al. 2010).

Apart from the study of Müßig and colleagues, who use a combination of siRNAs and dsRNAs, all studies used target-specific dsRNA rather than siRNAs, the 21–23-nucleotide (nt) molecules processed out of longer dsRNAs, to manipulate gene function. However, dsRNAs have repeatedly been shown to cause off-target effects in higher animals. Studies in mammalian cells have shown that RNAi can cause the degradation of untargeted mRNAs by crosshybridization regions to the processed siRNAs (Jackson et al. 2003; Scacheri et al. 2004) or by siRNAs acting as miRNAs (Jackson et al. 2006). Additionally, dsRNAs may also alter gene expressions in a sequence-independent manner, such as activating antiviral mechanisms (Kumar and Carmichael 1998). Hence, the introduction of exogenous dsRNA molecules into mammalian cells often results in a global, nonspecific suppression of gene expression. This is achieved by the activation of two independent RNAi-activated pathways: the dsRNA recognition protein PKR (dsRNA-dependent protein kinase; Nanduri et al. 1998) and the 2′,5′-oligoadenylate synthetase. Both pathways lead to a general inhibition of protein synthesis (Sledz and Williams 2004). Double-stranded RNA also initiates a signalling
3. Off-target effects caused by dsRNA cascade leading to the production of interferons (Williams 1999). Cytokines, which represent the first line of defence against viral infections, trigger the up-regulation of interferon-stimulated genes and consequently lead to altered protein synthesis. Such sequence dependent off-target effects as well as sequence independent reactions towards dsRNA were also found in higher non-mammalian vertebrates (Oates et al. 2000; Zhao et al. 2001) and insects (Kulkarni et al. 2006), suggesting that the phenomenon of off-target RNAi reactions is not restricted to mammals.

To assess whether such off-target effects also occur in adult honeybees treated with dsRNAs, we analysed the gene expression of four nontarget genes in two different abdominal tissues, the fat body and the ovaries, to compare whether different tissues treated with the same dsRNA show tissue-specific responses. The selected tissues are of prime interest for understanding honeybee biology because they are closely linked to the control of reproduction (ovaries) and are central to the honeybee’s immune system (fat body). Furthermore, we chose one dsRNA (dsGFP) that has no known honeybee homologue and two dsRNAs (dsGPDH and dsVG) from the honeybee transcriptome. In particular, the dsVG sequence used in this study has been shown to successfully knock down its targeted gene in the honeybee fat body (Amdam et al. 2003).

To quantify the impact on expression levels, four non-target genes that lacked similarities with any of the injected dsRNAs were chosen: (1) amSID-I: This is the honeybee homologue of the SID-I transmembrane channel protein. It is involved in dsRNA internalization in C. elegans and humans and facilitates systemic RNAi responses (Winston et al. 2002; Feinberg and Hunter 2003). This gene is particularly suited because Aronstein et al. (2006) report on a correlation between amSID-1 expression and the application of dsRNA in adult honeybees. (2) amATF-2: This gene shares homologies with the mammalian ATF-2 transcription factor. Among others, genes targeted by amATF-2 regulate transcription factors and proteins engaged in stress and DNA damage response (Bhoumik et al. 2007). (3) amDHAP-AT: Dihydroxy acetone phosphate acyl transferase is involved in lipid metabolism, facilitating the production of triacylglycerides (TAG). TAGs are used in eukaryotes as energy storages and repository of essential and non-essential fatty acids (Coleman and Lee 2004). (4) amCPR: NADPH-dependent cytochrome P450 reductase belongs to cytochrome P450 enzymes. These enzymes are involved in the detoxification of xenobiotics and are therefore commonly used as stress biomarkers. In insects, endogenous functions of these enzymes include the metabolism of ecdysteroids, juvenile hormones and pheromones.
3. Off-target effects caused by dsRNA

(Feyereisen 1999). As these four selected genes code for proteins that cover very different physiological functions, they are particularly suited to screen a variety of different gene cascades for unspecific RNAi effects in the organism.

2. Materials and Methods

2.1. BLAST analyses of dsRNA sequences

All three selected dsRNA sequences were compared with the honeybee genome during the design process using the Basic Local Alignment Tool. None of the dsRNAs shared sequence similarities with any of the evaluated non-target genes or contain any 20-bp segments identical to any known bee sequence. As dsRNAs are processed by the dicer complex into a cocktail of siRNAs 19–21 nt in length, the absence of 20-nt stretches of homology minimizes the possibility of off-target effects.

2.2. Production of dsRNA

To generate templates for dsRNA production, we cloned the amVG and the amGPDH part into pGem-T easy vectors (Promega). The respective fragments were obtained by standard PCRs using approximately 100-ng genomic DNA obtained by chloroform–phenol extraction (e.g. Maniatis et al. 1982; for primers, see Table I). As there are several *Apis GPDH* isoforms, there is the danger of getting a mixture of different PCR products for amGPDH. Therefore, we chose two primers in a region lacking the conserved domains (dsGPDH position within the amGPDH gene, 636–816). Furthermore, we checked the product identity by direct sequencing. In the case of amVG, we used primers from a well-established protocol (Amdam et al. 2003). The obtained vectors containing amGPDH and amVG, as well as the pGFP vector (GenBank ID: U17997, Clontech) were cloned into JM109 competent cells according to the manufacturer’s instructions (Promega). Plasmids were purified after Del Sal et al. (1988). One of the obtained amGPDH and amVG clones, as well as one clone carrying the GFP encoding sequence, was used for PCRs producing the dsRNA templates. PCRs were adapted to the Biotherm™ DNA Polymerase (Genecraft) using 0.2 mM dNTPs, 0.3 µM of T7 promoter-added primer (see Table I), 1.5 mM MgCl₂ and 5 U Taq polymerase in a total reaction volume of 100 µL. PCR protocols consisted of 5-min DNA denaturation and Taq activation at 95°C, followed by 40 cycles of 30 s at 95°C, 30 s at 56°C for GFP and 54°C for amGpdh and amVG, and 1 min at 72°C. A final extension of 20 min at 72°C completed the protocol. The resulting PCR products were
purified with the QIAquick PCR Purification Kit (Qiagen). Subsequently, dsRNA from GFP, amVG and amGpdh was derived using the T7 Ribomax™ Express RNAi System (Promega) with an extended transcription time of 5 h at 32°C. The resulting dsRNA was purified by a Qiazol chloroform treatment and the pellet resolved in nuclease-free water. The dsRNA quality was verified in 1.8% agarose gels and its concentration photometrically quantified. dsRNA concentrations were adjusted to 5 µg/µL by diluting with insect ringer (see Section 2.3.) right before the injection.

2.3. Injection and incubation

Brood combs from one Apis mellifera carnica colony from the apiary of the Martin-Luther-University (Halle/Saale) were incubated at 34°C and 60% humidity. Newly emerged workers were anaesthetized by cooling on ice and subsequently injected with 5 µg of each dsRNA with a microsyringe (Hamilton, 10 µL) between the fifth and sixth abdominal segments following established protocols (Amdam et al. 2003). Negative controls were injected with insect ringer (54 mM NaCl, 24 mM KCl, 7 mM CaCl$_2\times$2H$_2$O). Both groups were marked with coloured tags. Injected bees were kept on wax plates until they recovered. Bees not showing haemolymph leakage were kept for 24 h at 34°C with food and water ad libitum together with 25 untreated worker bees. After 24 h, the bees were shock-frozen in liquid nitrogen and stored at −80°C until tissue preparation.

2.4. RNA preparation and real-time measurements

Ovaries and fat bodies were dissected on cooled wax plates using RNAlater (Ambion) in order to avoid RNA degradation. Tissues were manually homogenised using plastic pestles. RNA extraction followed the standard Trizol (Invitrogen) protocol (Chomczynski and Sacchi 1987) with subsequent DNase (Promega) digestion. RNA quality and quantity were assessed by photometry. Aliquots containing 1 µg RNA were immediately reverse-transcribed with M-MLV H-Point Mutant Reverse Transcriptase (Promega) using oligo-dT Primer (0.5 µg/µL, Promega) according to the manufacturer’s instructions. Sybr Green assays consisting of 5 µL iQ SYBR Green Supermix (Biorad), 1 µL template and 1 µL of each Primer (1 µM) in a 10-µL reaction volume were run for gene expression studies. Each sample was run in duplicate. The real-time PCR cycling profile consisted of 3-min incubation at 95°C, followed by 39 cycles of 15 s at 95°C and 30 s at 54°C for annealing and 30 s at 72°C for extension and data collection. The
following melting curve analysis was performed between 50°C and 90°C, reading the fluorescence at 1°C increments. The purity of the PCR products was additionally checked on 1.8% agarose gels. C(t) values were calculated by the Opticon Monitor 3 software (Biorad) using a single standard deviation over cycle range after baseline subtraction using the Global Minimum Trend option.

2.5. Data analyses and statistics

Whenever replicate samples differed in C(t) values larger than 0.5, the samples were rerun to obtain more reliable estimates for the average C(t) values. For calculating the respective relative gene expressions (RGE), the honeybee ortholog of the ribosomal protein 49 (rp49) was used as a housekeeping gene (Lourenço et al. 2008). The PCR efficiency for every sample was calculated from the linear phase of fluorescence increase due to target duplication (Peccoud and Jacob 1996; Pfaffl 2001a) to control for different PCR efficiencies between different samples and different genes. Relative gene expressions were calculated according to Pfaffl (2001b) using the following equation:

\[
RGE = \frac{\text{Efficiency}_{\text{target}}^{-C(t)}}{\text{Efficiency}_{\text{rp49}}^{-C(t)}}
\]

3. Results

The injection with insect ringer, which was used for dsRNA dilution, had no detectable impact on the gene expression of the four analysed non-target genes (Figure 1) in both of the evaluated tissues. Hence, ringer-injected and untreated bees were pooled to provide the controls for further analyses.
3. Off-target effects caused by dsRNA

Figure 1. Pooled relative gene expression of four non-target genes in untreated bees compared with bees injected with honeybee ringer in two different tissues. We pooled the data for simplification as none of the individual genes showed an altered gene expression after the injection of ringer solution (t tests of log-transformed data, $P_{\text{min}}=0.24$). After pooling, the injected bees did not differ from the untreated bees in their transcript level of the evaluated genes in either tissue (repeated-measures ANOVA of log-transformed data—fat body: $P=0.642$, $F=0.649$; ovaries: $P=0.926$, $F=0.200$). N refers to the number of individual bees. Note that the expression of all four non-target genes of every bee was taken into account when calculating the pooled gene expression.

3.1. Effect of amGPDH-specific dsRNA on its target gene in different tissues

Injection of dsGPDH and dsGFP into the body cavity of adult honeybees led to a marked amGPDH gene knockdown of 81% and 79%, respectively, in the fat body tissue (Figure 2). The similarity in the extent of the amGPDH gene knockdown is surprising as, unlike dsGPDH, dsGFP does not show any sequence similarity towards amGPDH. In contrast, neither dsGPDH nor dsGFP affected the amGPDH gene expression in ovarian tissue.
3. Off-target effects caused by dsRNA

Figure 2. amGPDH knockdown in fat body and ovarian tissue after injection of dsRNA directed against amGPDH and GFP, respectively. Both the amGPDH gene expression in the fat bodies and ovaries of untreated and ringer-injected bees did not significantly deviate from each other (t tests of log-transformed data: fat body, \( P = 0.293 \); ovaries, \( P = 0.177 \)). They serve as controls and were set to 1. Asterisks indicate significant differences (*\( P \leq 0.05 \), **\( P \leq 0.01 \); t test of log-transformed data).

3.2. Effects of dsRNA sequences on the overall gene expression in abdominal tissues

The dsRNA sequences had highly variable impacts on the overall gene expression of the non-target genes in ovarian and fat body tissues. The specific dsRNAs for amVg and for GFP did not alter the overall gene expression of the four non-target genes in either tissue. In contrast, the dsRNA for the honeybee amGPDH homologue had a strong impact on the gene expression of the evaluated genes in the fat body. Injection of this dsRNA led to a transcript level decrease of 70% in the fat body compared with the gene expression in the ovaries (Figure 3). To exclude potential differences in endogenous expression of the nontarget genes between ovarian and fat body tissues that could confound the observed tissue-specific differences in gene expression, we compared the endogenous tissue-specific expression of every gene in untreated and ringer-injected bees (Figure 4). None of the genes showed a significantly different expression between the fat body and the ovaries in both experimental groups.
3. Off-target effects caused by dsRNA

Figure 3. Impact of three different dsRNA sequences on the gene expression of four non-target genes in two abdominal tissues. N number of measurements. Asterisks indicate significant differences between the gene expression in honeybee ovaries and fat body (repeated-measures ANOVA of log-transformed data: P=0.013, F=7.111). The overall gene expression in the fat body and the ovaries did not differ from the controls (repeated-measures ANOVA of log-transformed data—fat body: P=0.380, F=1.109; ovaries: P=0.330, F=1.321).
3. Off-target effects caused by dsRNA

Figure 4. Endogenous gene expression of *amSID-1*, *amATF-2*, *amCPR* and *amDHAP-AT* in ovaries and fat bodies of untreated and ringer-injected honeybees. Individuals of both groups did not show differential expression between the evaluated tissues for any tested gene (t test of log-transformed data).

3.3. Tissue-specific effects of dsRNA sequences on selected genes

In parallel to the differences in the overall gene expression across both tissues, we observed gene- and tissue-specific differences in transcript abundances after dsRNA treatment (Figure 5). Compared with the controls, injection of dsGPDH led to an increased *amATF-2* expression in the ovaries and a significantly reduced *amDHAP-AT* expression in the fat body. As the tissue-specific endogenous expression of both genes did not differ (Figure 4), it is clear that the injection of dsGPDHs led to the altered gene expression profile between both abdominal tissues.

In comparison to dsGPDH and dsGFP, dsVG injection resulted in a marked (90%) and specific knockdown of *amCPR* mRNA in ovarian tissue relative to control bees.
The downregulation of *amCPR* in the ovaries was specific for dsVG as the gene expression of *amCPR* within this experimental group significantly differed from all other dsRNA treatments. Hence, *amCPR* expression in the ovaries was significantly different from all other evaluated genes (Newman–Keuls post hoc test, \( P \leq 0.05 \)). Finally, in contrast to dsGPDH and dsVG, dsGFP did not alter transcript abundances in any of these four evaluated genes in either abdominal tissue.

**Figure 5.** Tissue-dependent transcript level of four non-target genes after the treatment with one out of three different dsRNAs. The relative gene expression values for pairwise comparisons of the injection effects on the four non-target genes in the fat body and the ovaries were normalized by setting the transcript level of the control groups to 1. Asterisks indicate significant differences between the relative expression of the non-target gene between the two tissues (*\( P \leq 0.05 \), **\( P \leq 0.01 \); t test of log-transformed data). Number sign indicates significant differences (#\( P < 0.01 \); t test of log-transformed data) between the respective treatment and the tissue specific control (untreated and ringer-injected individuals). Note the logarithmic scale of the y-axis.
4. Discussion

The injection of the dsRNA solvent did not have any detectable impact on the studied non-target genes. Therefore, our observations were not the result of a wounding or septic reaction in response to the ringer injection, but specific responses to our dsRNA treatments, either caused by sequence homologies or toxicity of the dsRNA molecules.

4.1. Tissue-specific response on dsRNA injections

Initially, we determined the relative expression of the non-target genes in the fat body and the ovaries in untreated and ringer-injected control individuals to ensure that the shifts in transcript abundance after dsRNA treatment were not just caused by the differences in the endogenous expression levels in both tissues. In the fat body, the lack of differential expression amongst the tested genes in control individuals confirms that the overall downregulation was due to the dsRNA treatment (most strongly for dsGPDH). The fat body trophocytes are known for dsRNA uptake (Amdam et al. 2003; Seehuus et al. 2006) since, among other functions, they are central to the detoxification and secretion of substances destined for exportation (de Oliveira and Cruz-Landim 2003). Therefore, the fat body contains a suite of transport mechanisms designed for the rapid uptake and release of an array of substances from the haemolymph. In comparison, ovaries and more particular follicle cells are less accessible for dsRNAs (Jarosch and Moritz 2011) as two dense cellular layers, the ovariole sheath (King et al. 1968) and the follicular epithelial cells (Engels 1968), may act as efficient barriers towards the dsRNA molecules.

4.2. Off-target gene regulation by dsRNA

We recorded four different dsRNA–off-target gene combinations showing altered transcript abundances after the treatment. dsGPDH altered the expression of *amATF-2* and *amDHAP-AT*, dsVG treatment decreased the expression of *amCPR* and injection of dsGFP decreased the expression of *amGPDH*. Clearly, every dsRNA evaluated in this study had an effect on a single gene, and one (dsGPDH) affected two different non-target genes. One of those genes, *amDHAP-AT*, is metabolically related to *amGPDH*. Both the target and the non-target genes code for proteins involved in lipid metabolism. The *Apis* homologue of GPDH bridges glycolysis and both, the production and degradation of triacylglycerides. Furthermore, it serves as a cytosolic partner in the
glycerol-3-phosphate shuttle (Brisson et al. 2001). *amDHAP-AT* facilitates the production of triacylglycerides, which are used as energy stores and a repository of essential and non-essential fatty acids (Coleman and Lee 2004). The downregulation of *amGPDH* and, therefore, the inhibition of glycolysis may have led to a parallel downregulation of *amDHAP-AT*. Especially the downregulation of *amDHAP-AT* within the fat body, the tissue where lipids are stored (de Oliveira and Cruz-Landim 2003), suggests a co-regulation of both enzymes in order to cope with the altered energy budget of the cells. Nevertheless, none of the other dsRNA–gene combinations in this study is physiologically related in a similar manner to *amGPDH* and *amDHAP-AT*, suggesting that they represent true off-target gene regulation.

**4.3. Mechanistic reasons for off-target effects**

Since the downregulation of *amGPDH* in the fat body cells by dsGPDH was accompanied by several non-target effects, the specificity of both the knockdown and dsGPDH remains questionable. The non-target downregulations may have been caused by sequence-specific cross-hybridizations between the processed secondary siRNAs and the genes. Nevertheless, all three dsRNAs, all specifically designed to have no sequence homology longer than 20 bp with any gene in the honeybee genome, showed at least one unspecific off-target knockdown. Although we cannot completely exclude the possibility of interactions between the secondary siRNAs and the evaluated genes, we still feel it prudent to consider the observed effects, particularly those of dsGFP and dsVG, as sequence-unspecific off-target effects.

In conclusion, we strongly recommend concentrating effort on the design of RNAi effective molecules, combining several dsRNAs for one target gene and using more stringent controls when setting up RNAi protocols in honeybees. To rigorously identify gene functions based on RNAi protocols in honeybees. To rigorously identify gene functions based on RNAi-derived phenotypes, measuring the mRNA level of RNAi targeted genes relative to a single non-target gene is clearly insufficient. As this study shows, treatments with gene-specific dsRNA can lead to nonspecific effects, which in turn may lead to false interpretations of the observed RNAi-derived phenotypes.

**Acknowledgements**

This study was financially supported by the DFG (RFAM).
3. Off-target effects caused by dsRNA

References


3. Off-target effects caused by dsRNA


3. Off-target effects caused by dsRNA

### Table I. Primer sequences and corresponding product sizes (all primers except the amVG primers were derived using Primer3; Rozen and Skaltsky 2000).

<table>
<thead>
<tr>
<th>Method</th>
<th>Gene (Accession-nr.)</th>
<th>Primer</th>
<th>Sequence (5'→3')</th>
<th>Product size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GFPI</strong></td>
<td>(M62653)</td>
<td>GFPI</td>
<td>TAATACGACTCCTATAGGGGCGATTTCC ATGGCCAACACCTGTCA</td>
<td>501</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GFPII</td>
<td>TAATACGACTCCTATAGGGGCGATCAAG AAGGACCATGGTC</td>
<td></td>
</tr>
<tr>
<td><strong>amGPDH</strong></td>
<td>(NM_001014994)</td>
<td>GPDH-T7I</td>
<td>TACGACTCCTATAGGGCGATGCTGGTT TCATCGATGTTT</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GPDH-T7II</td>
<td>TAATACGACTCCTATAGGGCGATACGA TTTGACCCACGTACCA</td>
<td></td>
</tr>
<tr>
<td><strong>amVg</strong></td>
<td>(NM_001011578)</td>
<td>GPDH-T7I</td>
<td>TACGACTCCTATAGGGCGATACGA TTTGACCCACGTACCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GPDH-T7II</td>
<td>TAATACGACTCCTATAGGGCGATACGA TTTGACCCACGTACCA</td>
<td></td>
</tr>
<tr>
<td><strong>amRp49</strong></td>
<td>(NM_001011587)</td>
<td>Rp49I</td>
<td>TCGTCACCAGAGTGATCGTT</td>
<td>243</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rp49II</td>
<td>CCATGAGCAATTTCAGCACA</td>
<td></td>
</tr>
<tr>
<td><strong>amSID-1</strong></td>
<td>(XP_395167)</td>
<td>amSID-1I</td>
<td>GCTCGGGCATTCACTGTTAATT</td>
<td>296</td>
</tr>
<tr>
<td></td>
<td></td>
<td>amSID-1II</td>
<td>ACTGCAAGAGCAATTTCAGCACA</td>
<td></td>
</tr>
<tr>
<td><strong>amATF-2</strong></td>
<td>(XP_393896)</td>
<td>amATF-2I</td>
<td>GATTGGACGGATCGAAGGGA</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td></td>
<td>amATF-2II</td>
<td>TGGATCCCTTCCTTGAGGA</td>
<td></td>
</tr>
<tr>
<td><strong>amDHAP-AT</strong></td>
<td>(XP_396018)</td>
<td>amDHAPI</td>
<td>ATTGCAAGTGGAATGGATTT</td>
<td>463</td>
</tr>
<tr>
<td></td>
<td></td>
<td>amDHAPII</td>
<td>ATTGCAAGTGGAATGGATTT</td>
<td></td>
</tr>
<tr>
<td><strong>amCPR</strong></td>
<td>(XP_001119949)</td>
<td>amCPRI</td>
<td>AATTGAAGTGCGAGGAAG</td>
<td>464</td>
</tr>
<tr>
<td></td>
<td></td>
<td>amCPRII</td>
<td>GAACATGAGTGCGTGGTT</td>
<td></td>
</tr>
<tr>
<td><strong>amGPDH</strong></td>
<td></td>
<td>GPDHIII</td>
<td>ACGGGCAAGAAAATCTCTGA</td>
<td>172</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GPDHIV</td>
<td>CCATAGGGCATCTCAGCACC</td>
<td></td>
</tr>
</tbody>
</table>
4. Molecular characterization of a genomic region which controls reproductive hierarchies in honeybees (Apis mellifera sp.)

Antje Jarosch¹, Robin F. A. Moritz¹,²

¹Institut für Biologie, Martin Luther Universität Halle-Wittenberg. Hoher Weg 4, 06099 Halle (Saale), Germany
²Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa

Submitted to Insect Biochemistry and Molecular Biology.
Abstract

In spite of the availability of the honeybee genome, the genetic switches controlling worker sterility and the reproductive monopoly of the queen still remain elusive. We here use Cape honeybee workers (*Apis mellifera capensis*) as a test system to study the genetic basis for reproductive dominance because they are able to lay unfertilized diploid eggs even in the presence of the queen. The cytological mechanism that restores egg diploidy called thelytokous parthenogenesis as well as other reproductive traits were shown to be under the control of a single locus on chromosome 13. This region with a length of about 11.4 cM comprises 15 annotated genes. The expression of these genes was assessed using sterile and reproductive workers and fully reproductive queens, to study their functional role in individuals with different reproductive capacities. Indeed, more than 40% of them were expressed in a caste specific manner. Apart from the well studied *gemini* gene which controls the egg-laying capacities of non-*capensis* workers by alternative splicing, a second candidate gene has been identified, which might be involved in the Insulin-signalling cascade that has been shown to be important for reproduction caste determination.

**Keywords:** Real-Time PCR; *th* locus; eusocial insects; caste determination
4. Molecular characterization of the $th$-locus

**Introduction**

Eusocial hymenopteran insect societies are characterized by a reproductive division of labour (Wilson 1971). Queens monopolize mating and reproduction, whereas workers refrain from reproduction and instead raise the offspring of the queen. Kin-selection has been suggested as the most plausible evolutionary force stabilizing the eusocial structures of insect colonies (Hamilton, 1964), as workers are more closely related to their half-sisters than to the offspring of them. Since workers are unable to mate, they only produce unfertilized, haploid eggs that develop into males via arrhenotokous parthenogenesis. But, as such worker reproduction is in conflict with the other workers’ interests it only scarcely occurs (Whitfield 2002). The low rates of worker oviposition are due to pheromonal control of both the queen and her brood (Slessor et al, 1988; Winston and Slessor 1998). Nevertheless, selfish reproductive workers exist and they are able to lay up to 7% off the colony’s male eggs (Visscher 1996). However, less than 0.01% worker laid eggs eventually develop into adult drones (Visscher 1989; Ratnieks and Visscher 1989), because reproductive workers are aggressively confronted (Visscher 1995) and policing workers remove and cannibalize the worker laid eggs (Ratnieks and Visscher 1989; Ratnieks 1988).

Cape honeybees (*Apis mellifera capensis*) undermine the reproductive monopoly of the queen. Workers of this subspecies are able to produce diploid female offspring (Onions 1912) by thelytokous parthenogenesis (Verma and Ruttner 1983) even in the presence of a queen. Because of an automixis with central fusion and an extremely reduced rate of crossing over (Moritz and Haberl 1994; Baudry et al. 2004) this way of reproduction causes relatedness close to unity between laying workers and their offspring ($r=1$) which is much higher than among supersisters ($r=0.75$). This altered kin structure favours selection of reproductive workers (Greef 1996), which produce clonal offspring. These clones can be raised as new queens (Moritz et al. 1996) or can re-queen queenless colonies (Holmes et al. 2010; Moritz et al. 2011). Moreover, parasitic thelytokous workers invade queenright foreign colonies, kill the resident queen and, establish themselves as pseudo-queens (Neumann and Moritz 2002). This colony takeover is possible because host workers do not recognize the absence of the queen since the parasitic workers produce queen-like amounts of the queen substance 9-Oxo-2-decenoid-Acid (9-ODA) (Ruttner et al. 1976; Hemmling 1979; Crewe and Velthuis 1980; Wossler 2002; Simon et al. 2001; Sole et al. 2002; Dietemann et al. 2007). Moreover, worker laid eggs have a lower removal rate as policing is less frequent than
in normal honeybee colonies (Moritz et al. 1999). Three of these pseudoqueen traits of cape honeybees (ovary activation, 9-ODA production, thelytoky) have been shown to be under pleiotropic control of a single recessive locus that has been mapped to chromosome 13 (Lattorff et al., 2005; 2007). Pseudoqueen workers are however also found in normal arrhenotokous colonies (Sakagami 1958). Therefore, the molecular switch controlling the thelytoky syndrome in Cape honeybees is necessarily the same as controlling reproductive castes in arrhenotokous subspecies.

The mapped region including the \textit{th} locus comprises 15 annotated genes (Lattorff et al. 2007; Figure 1). Although \textit{gemini} has been shown to be the central gene controlling ovary activation (Jarosch et al. 2011) the role of the remaining 14 genes has not been evaluated in queens and workers so far. Especially locus 410417 which encodes a receptor protein tyrosin phosphatase (RPTP) might be of interest. The RPTPs form a subfamily of the classical protein-tyrosine phosphatases (PTPs) that exclusively dephosphorylate phosphotyrosine (pTyr) in target proteins (den Hertog 2004). One of such PTPs (PTP-1B) has a major role in modulating insulin sensitivity in mice (Elchebly et al. 1999) probably by dephosphorylating the β-subunit of the insulin receptor (IR, Lammers et al. 1997). IR homologues in \textit{Drosophila melanogaster} are central regulators of growth, metabolism, longevity and reproduction (Garafalo 2002 and citations within) and are therefore candidates for caste determination in honeybees as well. Indeed, \textit{Apis} homologues of IRs as well as other genes of the insulin signalling pathway are differentially expressed between worker and queen larvae (Wheeler et al., 2006; Patel et al. 2007; Azevedo and Hartfelder 2008) and adults (Corona et al. 2007). Measurement of the gene expression of RPTP and of the remaining 14 genes annotated within the \textit{thelytoky} region in different castes and different reproductive states of arrhenotokous workers might provide additional knowledge on the genetic control of reproductive hierarchies in honeybees.
4. Molecular characterization of the \( \text{th} \)-locus

**Figure 1.** Location and candidate genes of the \( \text{th} \) region on chromosome 13 based on Amel 4.5. Characterising genetic markers are given in the middle. Homologies, conserved domains and putative functions of the 15 candidate genes associated with the thletyotous mode of parthenogenesis, the production of queen-like amounts of 9-ODA and the early onset of egg-laying are given in the table. Italic marking stands for the former functional classification of the respective gene in Amel 4.0 (Consortium HGS 2006).

**Results**

The majority of the evaluated candidate genes was differentially expressed between castes and between different reproductive states of the worker caste in the arrhenotokous subspecies \( A. \text{m. carnica} \). Only three genes (loc410420, loc410416, loc410419) were expressed independent from the evaluated phenotype (Fig. 2).
4. Molecular characterization of the \( \text{th} \)-locus

**Figure 2.** Efficiency corrected gene expression of loc410420, loc410416 and loc410419. None of the evaluated genes show a significantly altered gene expression neither in workers with different reproductive states nor in different castes. P-values are given within each figure separately (Kruskal-Wallace ANOVAs). The gene expression of reproductive workers and queens was normalized to mean values of sterile workers, which was set to one.

The transcript abundances of two putative genes (loc72550, loc408407) were correlated with different ovarian status of worker bees. Whereas locus 72550 was up-regulated (10-fold) in worker bees with developed ovaries, the mRNA level of locus 408407 was significantly decreased by more then 30 times. In both cases queens and non-reproductive workers exhibited the same transcript pattern (Fig. 3I). Six of the putative \( \text{th} \) genes were expressed in a caste specific manner. Two of them (loc725426, loc410422) were up-regulated in laying queens, whereas the transcript abundances of the other four (loc552578, loc410418, loc410423 and loc408406) were down-regulated by more then 100 times the maximum (loc408406) (Fig. 3II). The remaining three genes were regulated in the very same way. Egg-laying individuals (workers and queens) showed a decreased transcript abundance of the respective genes, compared to sterile, non-laying individuals (Fig. 3III).

Expression of locus 725247, which was formerly annotated as exonic region of locus 410418, was not detected in our samples.
Figure 3. Efficiency corrected expression of differentially expressed genes in the *thelytoky* region. The gene expression of reproductive workers and queens was normalized to sterile workers, whose mean gene expression was set to one. Different letters indicate significant differences between the evaluated groups. Capital/small letters refer to the applied statistical test (capital letters: One-way ANOVAs with log-transformed data and Newman-Keuls Post-Hoc tests; small letters: Kruskal-Wallace ANOVAs and multiple comparisons of mean ranks).

I, relative gene expressions of loc72550 and loc408407 which correlate with different worker phenotypes.

II, relative expression of six genes within the *th* locus, whose transcript abundance is associated with the phenotypic plasticity of different castes.

III, transcript abundance of three candidate genes within the *th* locus, which is correlated with different reproductive worker states as well as caste specific phenotypic differences.

Discussion

Caste specific phenotypic plasticity in honeybees is a result of differential feeding, including both the quantity as well as the quality of the food offered to queen and worker destined larvae (reviewed in Rembold 1964; Kamakura 2011). Subsequent
physiological and behavioural differences are prone to differential expression of gene sets orchestrated by juvenile hormone (JH) (reviewed by Smith et al. 2008) and epigenetic modifications (Kucharski et al. 2008; Lyko et al. 2010).

Massive caste specific gene regulation has been shown in honeybee brains. Here more than 2000 genes, which represent almost 40 % of all evaluated genes, have been shown to be differentially expressed between queens and workers (Grozinger et al. 2007). The same ratio was observed for the expression of the 14 detected transcripts in this study. Six genes (= 43 %) were differentially expressed between different castes. This includes locus 410418, which encodes the gemini transcription factor belonging to the CP2 transcription factor family and which has been shown to control ovary activation in arrhenotokous honeybee workers by alternative splicing (Jarosch et al. 2011). Moreover, simultaneous alternate splicing of two exons of gemini in Cape honeybees correlates with both, the queen like phenotype of the intrasocial parasite and the mode of worker parthenogenesis, strongly suggesting that gemini is involved in controlling the thelytoky syndrome.

Nevertheless, the expression differences of the remaining genes located within the th locus might be important in regulating ovary activation. The transcript abundance of a subset of three genes correlates with the reproductive capacity of the individual (worker non-reproductive < worker reproductive < queen) suggesting that they are somehow involved in the regulation of reproductive hierarchies. The expression pattern of these genes is similar the pattern of transcript abundance in brains of workers in getting a more queen-like expression when becoming reproductive (Grozinger et al. 2007). A particular promising candidate gene is locus 410417 which encodes a receptor type protein tyrosine phosphatase (RPTP). RPTP orthologs in other organisms are involved in insulin-signalling cascades (Elchebly et al. 1999; Lammers 1997), which in turn regulate growth, development, metabolism and reproduction (Garafalo 2002; Brogiolo et al. 2001). Mutated Drosophila insulin receptor homologues yield adults with a severe developmental delay, a female-sterile phenotype and a growth-deficiency (Chen et al. 1996; Brogiolo et al. 2001; Tatar et al. 2001). These phenotypes strongly resemble those seen in syndromes of insulin-resistance or IGF-I receptor deficiencies in higher organisms (Moller and O’Rahilly 1993; Flier 1993; Taylor and Moller 1993), suggesting an evolutionary conserved function for the insulin receptor pathway in the regulation of growth and body size from insects to humans (Chen et al. 1996).
4. Molecular characterization of the th-locus

*Drosophila* the actual level of activity of the insulin pathway reflects the provision of food (Garafalo 2002).

In honeybees, nutrition has long been known to control the caste fate of the totipotent female larvae. According to caste specific diets gene expression is suggested to be regulated by insulin/insulin-like-growth-factor-1 (IGF-1)-like signalling pathways (Wheeler et al. 2006; Corona et al. 2007; Patel et al. 2007; de Azevedo and Hartfelder 2008). Up-regulation of the RPTP gene located within the *th* locus is similar to the regulatory mode of genes coding for different insulin-like peptides and IRs in queen and worker larvae in critical stages of caste development (de Azevedo and Hartfelder 2008). Thus RPTP might be an important, upstream laying regulatory gene, which modulates subsequent IRs by dephosphorylating phosphotyrosines according to the nutritional status. This in turn might be an interesting clue towards the understanding of the molecular background controlling thelytoky in Cape honeybees. Here workers get preferentially fed once they are artificially introduced in foreign colonies (Beekman et al. 2000). This preferential feeding in combination with other behavioural and physiological pre-dispositions increases their reproductive potential (Neumann and Hepburn 2002; Zheng et al. 2010). Thus, this locus might provide the link between the environmental benefit of preferential feeding and the subsequent gene regulatory network leading to the parasitic phenotype of Cape honeybees.

**Materials and methods**

**Honeybee sampling**

*A. m. carnica* worker and queen bees were held at the apiary of the Martin-Luther-University Halle-Wittenberg. Workers, both reproductive and non-reproductive were sampled from small, hopeless queenless colonies with about 2000 workers. Every four days these colonies were checked for queen cells, which were destroyed. After massive productions of drone brood about 250 workers were randomly sampled, ovaries dissected and stored in RNAlater for subsequent molecular analyses. *A. m. carnica* queens were reared with standard apicultural techniques and held in small nuclei containing about 400 worker bees until day 4 after the onset of egg-laying. All bees, workers and queens, were sacrificed by flash-freezing in liquid nitrogen and stored at -80 °C.

**Tissue preparation and RNA extraction**
Worker ovary activation was classified according to Hess (1942), by pooling classes 1 and 2 as “inactive” and 3 to 5 as “activated”. Inactivated ovaries from three workers were pooled to allow for the amplification of rare transcripts. Worker (inactive n=30; activated n=9) and queen (n=9) ovaries were manually homogenised using plastic pestles followed by an RNA extraction using the standard Qiazol (Qiagen) protocol with subsequent Dnase (Promega) digestion. RNA quality and quantity was assessed by photometry. Equal amounts of RNA were immediately reverse transcribed with M-MLV H-Point Mutant Reverse Transcriptase (Promega) using oligo-dT Primer (0.5 µg/µl; Promega) according to the manufacturers’ instructions.

**Semiquantitative Real-Time PCR**

Semiquantitative Real-Time PCR to quantify the transcript abundance of the 15 candidate genes we conducted using Sybr-Green (Biorad) assays following the manufacturers’ instructions. Each sample was run in duplicate. The Real Time PCR cycling profile consisted of 3 min incubation at 95 °C, followed by 39 cycles of 15 sec at 95 °C and 30 sec at 55 °C for annealing, 30 sec at 72 °C for extension and data collection. Melting curve analysis was performed between 50 °C and 90 °C, reading the fluorescence at 1 °C increments. Additionally, the purity of the PCR products was visually verified by capillary gel electrophoresis (Qiaxcel; OM500). After baseline subtraction using the Global Minimum Trend option, C(t) values were calculated by the Opticon Monitor 3 software (Biorad), using a single standard deviation over cycle range.

**Data analysis**

The relative gene expressions of the th genes were calculated as the efficiency corrected gene expression of the target genes compared to the efficiency corrected gene expression of the mean of the two the house keeping genes *amGpdh* (NM001014994) and *amRp49* (AF441189.1). The PCR efficiencies of every PCR product were determined by pooling cDNA from every sample used in the calculations. After serial dilutions covering a 10⁴ fold template dilution range, gene and assay specific PCR efficiencies were determined (Pfaffl 2001). Data were log-transformed to gain a normal-distribution. Those genes still deviating from normal distribution were analysed by non-parametric Kruskal Wallace Anovas with subsequent multiple comparisons of mean ranks. Those genes showing a normal distribution were analysed by one-way ANOVAs.
4. Molecular characterization of the th-locus

with subsequent Newman-Keuls Post-Hoc tests. The type of the respectively applied statistical test is given within each figure.

Acknowledgements

We thank Johanna Neils, Ricarda Pohl, Denise Kleber and Petra Leibe for their lab- and field work assistance and Eckart Stolle for providing some of the primers used in this study. Financial support was granted by the Deutsche Forschungsgemeinschaft (RFAM) and by the National Research Foundation of South Africa.

References


4. Molecular characterization of the th-locus


4. Molecular characterization of the \textit{th}-locus


4. Molecular characterization of the th-locus


4. Molecular characterization of the \(th\)-locus

**Tables**

Table 1 Primer sequences, product sizes and PCR efficiencies of the annotated genes (Amel 4.5; Consortium HGS, 2006) situated within the \(th\)-locus.

<table>
<thead>
<tr>
<th>Gene/locus</th>
<th>Product size (bp)</th>
<th>PCR efficiency</th>
<th>Primer sequences (FW)</th>
<th>Primer sequences (RV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>amGPDH</td>
<td>183</td>
<td>2.0</td>
<td>GGATCAGGAAATGGGGTTC</td>
<td>CGGAAGCTTATGTCCTGGAA</td>
</tr>
<tr>
<td>amRp49</td>
<td>243</td>
<td>1.9</td>
<td>TCGTCACCGAGTGATCGTT</td>
<td>CCATGAGCAATTTCAGCACA</td>
</tr>
<tr>
<td>410423</td>
<td>209</td>
<td>2.26</td>
<td>GGTGCCGTTCTCACCTGCC</td>
<td>TGTTCCTCCTCGCAGGGG</td>
</tr>
<tr>
<td>408407</td>
<td>154</td>
<td>1.86</td>
<td>TGTCTGGTCGTGAGGAGGTT</td>
<td>GTGCTCTTTGACTTGGTTTC</td>
</tr>
<tr>
<td>410422</td>
<td>241</td>
<td>1.73</td>
<td>TATTTGTGGGACGTAGAACA</td>
<td>AATCGGTTAAAAATTGGAGA</td>
</tr>
<tr>
<td>408406</td>
<td>180</td>
<td>2.27</td>
<td>CCCACGCCGCCAGAAATATT</td>
<td>TGTTTTCGAGGCGCTCCC</td>
</tr>
<tr>
<td>552670</td>
<td>166</td>
<td>1.84</td>
<td>TATATAATGCTCAAGCACCAG</td>
<td>GATAACCAAAAGAAAACAT</td>
</tr>
<tr>
<td>408405</td>
<td>149</td>
<td>1.99</td>
<td>GCGACCAGCGTTCAACCGGA</td>
<td>TAGCGATCGCGAGGGCTGA</td>
</tr>
<tr>
<td>725501</td>
<td>206</td>
<td>1.85</td>
<td>TGGCCGCGAGACTCGTGAA</td>
<td>TGCCCAAGTGTTGACGCTAGC</td>
</tr>
<tr>
<td>410419</td>
<td>231</td>
<td>1.72</td>
<td>GCACCAGGTTCAAGTGGG</td>
<td>ACCGAATCTCCCATCTCTCTGGT</td>
</tr>
<tr>
<td>410420</td>
<td>316</td>
<td>1.89</td>
<td>GGTGCCAAAGGACCATTTGAAA AGACA</td>
<td>CTTTGTTTTTTCATCTGCAACAG</td>
</tr>
<tr>
<td>725426</td>
<td>202</td>
<td>2.07</td>
<td>TGGTTTTGATGCAAAACCCCTCA</td>
<td>TGCATAGTCAATCATTGCCAG</td>
</tr>
<tr>
<td>410418</td>
<td>204</td>
<td>2.18</td>
<td>TAGTCGCTTCAACGCTTTG</td>
<td>AGGCCGACTCTCAAAAGTGA</td>
</tr>
<tr>
<td>725247</td>
<td>341</td>
<td>1.85</td>
<td>GGGGGTGCTTGGGAGGTTG</td>
<td>ATAGCGCAGAGAGGCCAGG</td>
</tr>
<tr>
<td>410417</td>
<td>266</td>
<td>2.1</td>
<td>AACACGCGCAAGGCGGTTA</td>
<td>CGGCCGTTGGGAACCACG</td>
</tr>
<tr>
<td>552578</td>
<td>288</td>
<td>2.23</td>
<td>CGCGTCATCGAGGGGCTCG</td>
<td>CTAGCCGCCACAAAGGCA</td>
</tr>
<tr>
<td>410416</td>
<td>168</td>
<td>2.16</td>
<td>GATTGGACGAAATCGAGAGA</td>
<td>TGTTATCCCCCTTTGGCTTG</td>
</tr>
</tbody>
</table>
5. Alternative splicing of a single transcription factor drives selfish reproductive behavior in honeybee workers (*Apis mellifera*)

Antje Jarosch\(^a\), Eckart Stolle\(^a\), Robin M. Crewe\(^b\), and Robin F. A. Moritz\(^a,b\)

\(^a\) Institut für Biologie, Martin-Luther-Universität Halle-Wittenberg, 06099 Halle (Saale), Germany; and \(^b\) Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa

*Proceedings of the National Academy of Sciences USA* 108: 15282-15287 (2011)

doi: 10.1073/pnas.1109343108

Approved August 15, 2011 (received for review June 9, 2011)
Abstract

In eusocial insects the production of daughters is generally restricted to mated queens, and unmated workers are functionally sterile. The evolution of this worker sterility has been plausibly explained by kin selection theory [Hamilton W (1964) J Theor Biol 7:1–52], and many traits have evolved to prevent conflict over reproduction among the females in an insect colony. In honeybees (*Apis mellifera*), worker reproduction is regulated by the queen, brood pheromones, and worker policing. However, workers of the Cape honeybee, *Apis mellifera capensis*, can evade this control and establish themselves as social parasites by activating their ovaries, parthenogenetically producing diploid female offspring (thelytoky) and producing queen-like amounts of queen pheromones. All these traits have been shown to be strongly influenced by a single locus on chromosome 13 [Lattorff HMG, et al. (2007) Biol Lett 3:292–295]. We screened this region for candidate genes and found that alternative splicing of a gene homologous to the *gemini* transcription factor of *Drosophila* controls worker sterility. Knocking out the critical exon in a series of RNAi experiments resulted in rapid worker ovary activation—one of the traits characteristic of the social parasites. This genetic switch may be controlled by a short intronic splice enhancer motif of nine nucleotides attached to the alternative splice site. The lack of this motif in parasitic Cape honeybee clones suggests that the removal of nine nucleotides from the altruistic worker genome may be sufficient to turn a honeybee from an altruistic worker into a parasite.

**Keywords:** caste determination, cuticular protein 2-family, gene expression
Introduction

The evolution of a sterile worker caste in eusocial hymenoptera has been plausibly explained by inclusive fitness theory (1). Generally, workers refrain from reproduction as a result of intracolonial reproductive hierarchies. Ovary activation in honeybee workers (*Apis mellifera*) is inhibited by the pheromones from the queen and the brood (2). In addition, the multiple mating of the queen and the coexistence of many half-sibling subfamilies in the colony facilitates worker policing, where workers remove eggs laid by other workers (3), leading to <1% of worker-laid offspring (4).

These mechanisms often fail, however, to control worker reproduction of the Cape honeybee (*Apis mellifera capensis*). In this subspecies, laying workers can function as social parasites invading foreign colonies, killing the resident queen and establishing themselves as pseudoqueens (5–7). The proximate mechanisms for this parasitic life history strategy are well understood (8). Parasitic pseudoqueen workers produce a queen-like pheromonal bouquet indicating the presence of a queen to the host workers (9–13). Moreover, the diploid offspring of the parasitizing workers produce brood pheromones, suggesting the presence of a laying queen to the host workers and preventing these workers from activating their ovaries. Finally, the diploid eggs laid by the parasitic workers are not policed as predicted by evolutionary theory (1, 14). A single locus termed thelytoky (*th*) is thought to control this parasitic life history strategy (15). Workers that are homozygous for the *th* allele produce parthenogenetic diploid offspring, have rapid ovary activation, and produce queen-like pheromones. The *th* locus, which has been mapped to a region spanning 11.4 cM comprising 15 genes on chromosome 13 (16), is thus in control of the three essential characters that facilitate social parasitism. Because transcription factors control several traits within one organism, the two transcription factors in this region [*ATF2* (XM_393896) and *CP2* (XM_001121158; XM_393898)] were considered to be prime candidate genes to control the parasitic worker phenotype in *A. mellifera*. Because the *ATF2* homolog did not show differential expression between different castes (Fig. S1), and members of the *CP2* transcription factor family were differentially expressed between queens and workers in both the honeybee (Fig. S1) and the stingless bee *Melipona quadrispilota* (17), we focused on the *CP2* transcription factor as the prime candidate for the *th* locus in honeybees. This transcription factor is homologous to the *gemini* (= genitalia missing) locus in *Drosophila melanogaster* (18), which was shown to interact with the Spindle-F protein CG12114 (19). This protein has a minus end-directed microtubule activity, and
is involved in oocyte axis determination and oocyte microtubule cytoskeleton organization. This is of particular interest, as the restoration of diploidy in A. m. capensis worker eggs is caused by an abnormal spindle rotation during meiosis (20).

The functional diversity of several CP2 transcription factor family members is derived from alternative splicing leading to tissue- and stage-specific isoforms (21). We show that alternative splicing of the honeybee gemini homolog as a general mechanism for generating functional diversity of transcription factors (22) provides the genetic mechanism for the \textit{th} locus to affect reproductive dominance of honeybee workers.

\textbf{Results}

Comparing the \textit{gemini} splice patterns of laying and nonlaying thelytokous workers (\textit{A. m. capensis}) with those of laying and nonlaying arrhenotokous workers (\textit{A. m. carnica}) yielded different transcript isoforms at exon 5 and exon 7 (Fig. 1).

Exon 5 is a cassette exon with two mRNA splice isoforms—the full-length transcript and a shorter transcript with a deletion of 78 bp. The deletion of these base pairs does not affect the ORF and leads to a putative protein shortened by 26 aa. Exon 7 contains two alternative 5\' splice sites. The first site causes a deletion of 59 bp with a shift in the reading frame. The resulting isoform contains a stop codon in exon 8 and has therefore an incomplete DNA-binding domain lacking the C-terminal protein domains. Quantification of this transcript was not possible in our samples due to its low abundance. The second splice site results in a deletion of 24 bp (S7 transcript form in Fig. 1) and the loss of 8 aa in the putative DNA-binding domain spanning from exon 7 to exon 9. Both splice forms of exon 5 and exon 7 were found in combination with each other (Fig. 1).
5. Alternative splicing of \textit{gemini}

![Diagram of \textit{gemini} (introns as lines, exons as boxes) and its four pre-mRNA splice products. Arrows indicate the alternatively spliced exons. +5, isoform with exon 5; –5, isoform lacking exon 5; L, long isoform of exon 7; S, short isoform of exon 7. Exon 5 (yellow) is present in the first two isoforms but absent in the other two. The length polymorphism in exon 7 is shown in orange. Spliced intronic regions are light gray. Exon numbering refers to the annotated transcript XM_001121158.1 (National Center for Biotechnology Information; Nucleotide Database).]

Semiquantitative RT-PCRs revealed that the splice patterns of exons 5 and 7 were characteristic of the different reproductive states of the tested queens and workers (Figs. 2 and 3, Fig. S2, and Table S3), and of the different modes of parthenogenesis. In laying arrhenotokous queens that served as positive controls for fully reproductive females, both the spliced and the full transcripts were produced for both exons (Figs. 2 and 3, Fig. S2, and Table S3). Arrhenotokous \textit{A. m. carnica} workers showed a different pattern. Workers with undeveloped ovaries exclusively produced the full transcript containing exon 5, whereas laying workers almost exclusively produced the splice product lacking exon 5. Exon 7 predominantly showed the long splice variant in both reproductive states, although both isoforms were expressed at a higher level in laying workers. Thelytokous \textit{A. m. capensis} workers (homozygous at the \textit{th} locus), both laying and nonlaying, produced both transcript forms of exon 5 in a pattern similar to that of the queen. In this subspecies the different splice patterns of exon 7 rather than exon 5 matched the two reproductive states. When the splice pattern of thelytokous workers
was similar to that of nonreproductive arrhenotokous workers, they did not have activated ovaries. However, when the exon 7 splicing pathway was similar to that of the queen, workers were egg layers.

**Fig. 2.** Transcript ratios of alternate transcripts of *gemini* (mean values ± SE; log scale). (A) Exon 5-missing transcripts (−5) in relation to exon 5-containing transcripts (+5). (B) Unspliced exon 7 transcripts (L7) in relation to −24-bp deletion exon 7 transcripts (S7): arrhenotokous workers in blue (dark, nonlaying; light, egg laying), thelytokous workers in green, and arrhenotokous laying queens in black. Different letters indicate significant differences (P ≤ 0.05).
5. Alternative splicing of *gemini*

![Graph showing transcript ratio 2D scatterplot with 0.6 confidence ellipses of all *gemini* isoforms.](image)

**Fig. 3.** Transcript ratio 2D scatterplot with 0.6 confidence ellipses of all *gemini* isoforms. The ratio of exon 5-deficient transcripts in relation to exon 5-containing transcripts are plotted on the x axis, whereas the ratio of unspliced vs. spliced exon 7 transcript are plotted on the y axis. All of the values were log transformed before they were plotted. All of the tested groups apart from queen and reproductive *A. m. capensis* worker samples significantly differ from each other (one-way ANOVA $F = 29.33; \text{df} = 4; P \leq 0.001$ and Newman–Keuls post hoc tests $P \leq 0.01$). Every symbol stands for a single individual.

To rule out the possibility that the splice pattern of *gemini* was simply correlated with different reproductive states of workers, we knocked down the level of transcripts containing exon 5 in *A. m. carnica* worker bees by undertaking a series of siRNA feeding experiments to change the ratio of exon 5-containing/missing transcripts. Therefore, workers were fed daily with two siRNAs specific for the 3’ and 5’ end of exon 5 for 14 d. The abundance of transcripts containing exon 5 was reduced $>90\%$ (Fig. 4) by this treatment. Moreover, although the abundance of transcripts containing the shorter variant of exon 7 was lower in the treatment group compared with workers fed with nonsense siRNA, the ratio of exon 7-missing to exon 7-containing transcripts did not change significantly. Nevertheless, the exon 5 knockdown bees had a higher
reproductive capacity, showing ovaries with swollen ovarioles and mature eggs that were significantly more frequent than in the various controls (Fig. 5).

**Fig. 4.** Relative gene expression of all *gemini* transcripts after exon 5 knockdown by RNAi. (A and B Left) Gene expression of the transcripts containing or missing exon 5. (Right) Gene expression data on exon 7. (A) Relative gene expression of different *gemini* transcripts. The transcript expressions were normalized to untreated bees whose transcript expression was set to 1. The asterisks indicate a significantly reduced expression of transcripts containing exon 5 after the feeding of exon 5-specific siRNAs (P < 0.01; Kruskal–Wallace ANOVA and multiple comparisons of mean ranks for all groups). (B) Expression ratios of spliced vs. unspliced transcripts shown on a logarithmic scale. The asterisk indicates a significantly shifted exon 5 transcript ratio toward the shorter isoforms of exon 5 (P < 0.05; Kruskal–Wallace ANOVA and multiple comparisons of mean ranks for all groups).
Fig. 5. RNAi-mediated exon 5 knockdown and the resulting ovary activation in arrhenotokous A. m. scutellata workers from two different colonies treated with gemini-specific siRNA (n = 72), scrambled siRNA (n = 69), or sugar water (n = 79). (A) The frequency of reproductive workers in the three treatment groups. The knockdown by the combination of two transcript specific siRNAs leads to a significant increase of reproductive workers shown by different letters (P ≤ 0.05, Mann–Whitney U test tests with P values corrected for multiple testing). (B) Dissected worker ovaries were classified as developed when the ovarioles were swollen and mature eggs visible (Upper) or as undeveloped (Lower).

Sequencing the coding regions of gemini in arrhenotokous and thelytokous workers and in arrhenotokous queens yielded no differences in the overall sequence. Hence these regions themselves cannot function as the allelic forms of the th locus. Therefore, we sequenced all noncoding regions of gemini, including the promoter region and the introns flanking the alternatively spliced exons. No consistent sequence differences were found in the promoter region and in the nonflanking introns. Furthermore, sequencing the flanking intronic regions of exon 7 did not reveal any consistent sequence differences between arrhenotokous and thelytokous subspecies. Only in the downstream intron flanking cassette exon 5 (the putative molecular thelytoky switch), a consistent deletion of 9 bp (5’- GAAACGATG-3’) was found in the parasitic Cape honeybee clone (Fig. S3). We named this deletion the thelytoky associated element 1 (tae1). Both arrhenotokous subspecies, the European A. m. mellifera and the African A. m. scutellata, show the identical intronic sequence in this region, confirming that it is not just a marker that is characteristic of African honeybees.
This *tael* sequence shows a very high level of purines (77.8%) compared with the purine content of the whole intron (50.47%).

**Discussion**

In eusocial insects, caste-specific phenotypic plasticity and worker reproduction are controlled by differential gene expression (23, 24) maintained by epigenetic modifications (25, 26) and caused by differential feeding of queens with royal jelly (25, 27). This study reveals another genetic mechanism that controls reproductive dominance in honeybees: different transcript isoforms of a single transcription factor generated by alternative splicing, to control the reproductive physiology of this eusocial insect.

Characterization and quantification of honeybee *gemini* isoforms revealed that the abundance of four processed transcripts corresponded with different ovarian states of workers and also with different castes. Furthermore, knockdown experiments in *A. m. carnica* workers revealed exon 5 to be causative in regulating the ovarian development in arrhenotokous workers. This knockdown leads to an alteration in the ratio of exon 5-containing and exon 5-missing transcripts, which subsequently controls the onset of egg laying. We did not observe such a shift in the exon 7-containing/missing transcripts, suggesting that exon 7 is not involved in the control of ovary activation in arrhenotokous bees.

In addition to the control of ovary activation, the mapped *th* locus controls the production of queen-like amounts of 9-ODA, the queen substance (16). This pheromone suppresses worker reproduction in arrhenotokous subspecies and ensures the reproductive dominance of the queen. Furthermore, the very same locus is responsible for the thelytokous production of diploid worker-laid eggs (16). Because the alternative splice pattern of *gemini* is characteristic for both the queen pheromonal phenotype and the thelytokous mode of parthenogenesis in workers, we suggest that *gemini* and its altered transcripts are also important for these traits (although we did not prove this by knockout experiments as for ovary activation). As we showed that exon 5 directly affects ovarian development, we expect the shorter exon 7 in combination with a specific expression ratio of exon 5 to be responsible for both the caste-specific reproductive dominance in arrhenotokous honeybee species and the thelytokous production of diploid offspring. One possible interpretation of the data (among many others) is a simplistic two-exon regulatory model (Fig. 6) to explain this complex network. We suggest that splicing of exon 5 and a certain threshold ratio of unspliced
5. Alternative splicing of *gemini* vs. spliced transcripts directly controls the shift of ovary activation in arrhenotokous subspecies. To be consistent with our model, splicing of exon 5 alone should be insufficient to exhibit the full pseudoqueen phenotype in thelytokous Cape honeybees. The abundance of transcripts missing exon 5 differs slightly between laying and nonlaying thelytokous workers, but is well below the threshold that governs ovary activation of arrhenotokous workers; it may not sufficiently vary to serve as a switch for ovary activation in thelytokous workers. As thelytokous worker groups produce both splice variants of exon 5 qualitatively similar to that of arrhenotokous queens, we propose this queen-like level of exon 5 to determine the mode of parthenogenesis (thelytoky). Additionally, the differences in the splice products of exon 7 between laying and nonlaying thelytokous workers suggest that alternative splicing of exon 7 is involved in ovary activation in thelytokous workers. Reproductive thelytokous workers may reduce the relative abundance of the long exon 7 and 5 when exploiting a host, thereby completely resembling the queen splice pattern for both exon 5 and exon 7 (red queen line in Fig. 6).

**Fig. 6.** Double-exon regulatory model for controlling worker reproduction by alternative splicing of exon 5 (E5) and exon 7 (E7) of *gemini*. The expression of different transcript isoforms is dependent on the worker’s reproductive status and the mode of worker parthenogenesis (solid line, arrhenotoky; dotted line, thelytoky). The red line represents the splice pattern of the queen (queen line) that expresses both transcript isoforms for both exons. Minus (−) or plus (+) signs indicate the absence or presence of exon 5. Long or short indicates the unspliced or spliced isoform of exon 7.
The missing protein domains of the spliced transcripts may essentially modify the access of *gemini* to different DNA-binding domains of other genes, which eventually result in the observed phenotypic differences. Exon 7 is part of the CP2 DNA-binding domain, and the deletion of 8 aa is expected to have an impact on the DNA-binding capacity of the resulting protein and therefore on the worker phenotype. Similar effects have been shown in human *LBP-1d* (*UBP-1*, another CP2 member) transcript isoforms that have a deletion in the putative DNA-binding domain, generated by alternative splicing. Those transcripts are no longer able to bind DNA and repress DNA binding of other isoforms by forming heteromers (28). Because exon 5 is not part of the CP2 DNA-binding domain, a mechanism other than DNA binding must control the molecular switch controlling the onset of egg laying. A common mechanism for controlling the action of transcription factors is the alteration of dimerization domains that cause isoform variability (22). For example, in *grainyhead*, another member of the CP2 family, a mutant lacking the N-terminal activation domain binds to the full-length *grainyhead* isoform and hence inhibits its ability to homodimerize (29).

To further understand the mechanisms controlling the observed splice patterns, we sequenced all noncoding regions of *gemini*, including the promoter region, which may also comprise sequence stretches necessary for accurate splicing (30–32). We did not observe any consistent sequence differences in the nonflanking intronic regions or in introns flanking exon 7. However, all parasitic Cape honeybee workers share a common deletion of 9 bp in the flanking intron of exon 5—the exon that controls ovary activation in arrhenotokous honeybees and that is differentially spliced in thelytokous honeybees. Both European *A. m. mellifera* and African *A. m. scutellata* show the identical sequence without the 9-bp deletion in this flanking intron, confirming that it is not only a marker of African honeybees but perhaps the functional site controlling the mode of worker parthenogenesis. Interestingly, and similar to other intronic splice enhancer (ISE) motifs, this *thelytoky associated element 1* (*tae1*) sequence is purine rich and short in length. Differential splicing has repeatedly been shown to be controlled by such ISE motifs located downstream of alternate exons (33, 34). The 9-bp deletion in *A. m. capensis* may therefore have far-reaching effects on the splice-site recognition of the upstream laying cassette exon 5. As this was the only consistent sequence difference we found in the four flanking introns of the spliced exons, it might be possible that *tae1* not only affects the exon 5 switch but also the splicing of exon 7, as it has been shown that more distant splice regulatory motifs can control splicing of non-neighbor exons (32).
5. Alternative splicing of *gemini*

Predictions of the secondary structure of intron 5_6 of *gemini* suggest that this sequence is part of a stem/loop conformation (Mfold) (35). With the deletion of this motif, the stem integrating the relevant bases is predicted to disappear. This conformational change might influence the pre-mRNA splicing process via different mechanisms (36). Either structural elements might hinder the accessibility of *cis* regulatory sequences by splicing factors, or the altered secondary structure might interfere with the regulatory mechanisms required for successful and accurate splicing. Furthermore, it remains possible that the deletion itself is directly involved in the assembly of putative splice factors. Serine/arginine-rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs) have been shown to interact with such purine-rich motifs in mammals (33) and hence directly control the altered splice pattern of *gemini* in honeybees.

Irrespective of the actual molecular switch mechanisms, we have found a gene that is alternatively spliced in a caste and type of parthenogenesis-specific manner and central to the regulation of worker fertility. If the 9-bp deletion with the ISE motif (*tae I*) flanking the ovary activation exon in nonparasitic honeybee workers proves to be the allelic form of the thelytoky gene in future studies, it would take as little as nine nucleotides to turn an altruistic worker into one that can become a social parasite.

**Material and Methods**

**Honeybee Samples for *gemini* Expression.** *A. m. capensis* workers were reared from brood produced by social parasitic workers in an *A. m. scutellata* host colony at the University of Pretoria and are therefore homozygous for the thelytoky allele. The brood frames were incubated at 35 °C and ~60% relative humidity. Emerging bees were kept in hoarding cages containing 120 individuals with food and water ad libitum. After 16 d of incubation, the bees were killed and the ovaries dissected and stored in RNAlater until RNA preparation. *A. m. carnica* workers and queens were held at the apiary of the Martin-Luther-Universität Halle-Wittenberg. Sealed brood frames with young nurse bees and emerging brood were removed from the colony and kept queenless. After 14 d, ~200 workers (aged ~7–21 d) were randomly sampled from the frames and immediately flash-frozen in liquid nitrogen. *A. m. carnica* queens were reared using standard apicultural techniques and kept in small colonies of ~1,000 worker bees until day 7 after the onset of egg laying. All sampled bees were stored at −80 °C. Worker ovaries were classified using the categories of Hess (37) by pooling classes 1 and 2 as undeveloped and 3–5 as developed.
RNA Preparation. Ovaries from each individual were manually homogenized, and RNA extraction followed the standard TRIzol (Invitrogen) protocol with subsequent DNase (Promega) digestion. RNA quality and quantity was photometrically assessed. Equal amounts of RNA were immediately reverse-transcribed with M-MLV H− Point Mutant Reverse Transcriptase (Promega) using oligo(dT) primer (0.5 µg/µL; Promega) according to the manufacturer’s instructions.

Splice Variant Detection. Because exon 1 was never transcribed in any of the samples, cDNA was amplified using primers 4I and 18II (Table S1) spanning all exons except the first one (resulting in a PCR product of 2,299 bp; Amel_4.0) (38) with long-range PCR Enzymes (Finnzymes and Fermentas) following the manufacturers’ instructions. The PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) and thereafter cloned into the pGEM-T Vector using the TA cloning kit (Invitrogen). Transformation was done in JM109 competent cells according to the manufacturer’s instructions (Promega). Plasmids were purified and inserts sequenced and aligned using CLC Free Workbench 3 (CLC Bio). The annotated exons 11 and 12 are not transcribed, and because the annotated exons 14 and 15 are not separated by an intron, they are fused to exon 14. Moreover, the transcripts contain 179 bp at the 3’end, which had been originally annotated to be untranscribed genomic DNA in the Amel_4.0 genome sequence.

Rapid Amplification of cDNA Ends (RACE)-PCR. To complete the transcript sequences, RACE-PCRs were conducted. A 3’end cDNA amplification was done according to Scotto-Lavino et al. (39), and the nested PCRs were adapted to the BiothermTM DNA Polymerase (Genecraft) using 0.2 mM dNTPs, 0.3 µM of each primer (gemini-specific forward primers: sp7.1I; 7.2I and 7.3I in PCR1, 15I in PCR2; Table S1) and 0.5 units polymerase in a total of 10 µL. The PCR programs consisted of 5 min at 95 °C, followed by 40 cycles of 30 s at 95 °C, 30 s at 54 °C, and 60 s at 72 °C, and a final extension of 20 min at 72 °C. The PCR products were either directly sequenced after gel elution or cloned (see above) if the quality of the PCR products was insufficient for direct sequencing. Plasmids having the right insert were purified and sequenced as mentioned above. For the amplification of 5’cDNA ends the protocol of Scotto-Lavino et al. (40) was modified. After reverse transcription and poly(A)-tailing
poly(T) *gemini*-specific primers (QgemI; Qgem4I; Table S1) were used for amplification. Tailing the primers with 15 thymine residues ensured that these primers were only bound to the 5´end of the tailed cDNAs. The transcript specific primers sp5II (which spans the boundaries of exon 5 and 6) and 5II (which is situated in exon 5) were used as reverse primers. Purified PCR products were sequenced.

**RT-PCR.** *ATF2, gemini, glycerol-phosphate-dehydrogenase (GPDH), ribosomal protein 49 (rp49), and elongation factor 1α (EF) quantification was conducted with real-time PCR using SYBR Green (Bio-Rad) assays following the manufacturer’s instructions. The four different splice products were analyzed separately, because the PCRs of the combined altered transcript isoforms (+5.L7; +5.S7; −5.L7; −5S7) did not yield adequate products. Each isoform was quantified with primers specific for the two exons and their spliced counterpart (Table S1). Each sample was run in duplicate. The realtime PCR cycling profile consisted of a 3-min incubation at 95 °C, followed by 39 cycles of 15 s at 95 °C and 30 s at 54 °C for annealing, and 30 s at 72 °C for extension and data collection. Melting curve analyses were performed between 50 °C and 90 °C, reading the fluorescence at 1 °C increments. Additionally, the purity of the PCR products was visually verified on 1.8% agarose gels. After baseline subtraction using the global minimum trend option, C(t) values were calculated by the Opticon Monitor 3 software (Bio-Rad) using a single SD over cycle range.

**Data Analysis.** The relative gene expression of *ATF2* was calculated as the efficiency-corrected gene expression of *ATF2* compared with the efficiency- corrected gene expression of the housekeeping genes *GPDH* (NM001014994), *rp49* (AF441189.1), and *EF* (NM001014993). The expression ratios of the different transcript forms of *gemini* were calculated from the efficiency-corrected gene expressions for each *gemini* amplicon. The efficiency-corrected relative expression of the individual transcript isoforms was calculated by normalizing the *gemini* gene expression to the two housekeeping genes *GPDH* and *EF*. PCR efficiencies of every PCR product were determined by pooling cDNA from every sample used in the calculations. PCR efficiencies were determined using serial dilutions covering a 10^4-fold template dilution range (41). Nonparametric Mann–Whitney U tests were used for statistical comparisons adjusting the P values for replicate tests.
5. Alternative splicing of *gemini*

**gemini** (Exon 5) **Knockdown by RNAi.** Our aim was to activate ovaries in arrhenotokous workers by reducing the abundance of specific splice variants. Because transcripts missing exon 5 were not detected by qPCR in nonreproductive arrhenotokous workers, we did not conduct a silencing experiment to even further reduce this already rare transcript. Instead we targeted transcripts containing exon 5, which were highly abundant in all arrhenotokous workers, to substantially alter the ratio of unspliced vs. spliced transcripts and eventually activate the ovaries of *A. m. carnica* workers. Hence, two *gemini*-specific siRNA template sequences within the cassette exon 5 (78 bp) were selected using the siRNA target designer version 1.6 (Promega), which also scrambled the target-specific siRNA sequences resulting in siRNAs to be used as negative controls. The selected siRNA sequences were blasted against the honeybee genome to avoid potential off-target effects in genes other than *gemini*. None of the siRNAs shared sequence similarities longer than 14 nt (16- to 18-nt–long stretches of homology are suggested as the maximum acceptable length in RNAi studies per Ambion siRNA design guidelines). Furthermore, none of the genes with the highest sequence similarities was found twice when blasting both siRNAs individually. The two siRNAs obtained (for sequences, see Table S2) cover the 3´ and the 5´ regions of the template. Using this siRNA mixture decreased the chance of affecting unintended targets. siRNA was synthesized using the T7 Ribomax Express RNAi System (Promega) according to the manufacturer’s instructions but with an extended incubation time of up to 2 h. The siRNA quality and quantity was assessed by capillary gel electrophoresis and photometry.

**Honeybees and siRNA Feeding.** Newly emerged *A. m. carnica* workers (arrhenotokous, 1–2 d old) were sampled from a single brood frame and divided into three groups: untreated bees, bees treated with scrambled siRNA (which has no sequence similarity to any bee-specific gene), and a *gemini*-specific siRNA treatment group. Bees in the latter group were fed with a mixture of equal amounts of the two *gemini*-specific siRNAs. In each case, 35–40 newly emerged bees were put in wooden cages (10.5 × 13 cm) provided with a small piece of comb and pollen ad libitum at 34 °C. The bees were fed daily with 1.5 mL 50% sugar water containing 1 µg siRNA per individual. After 14 d, a time span after which ovary activation can easily be detected, bees were killed and screened for ovary activation in a doubleblind test. The samples were stored in coded tubes by a person with no knowledge of the treatments, to allow
for an unbiased examination of the ovaries by a second person without any prior knowledge of the experimental groups. Ovaries were classified as developed when ovarioles contained mature eggs. This procedure was repeated in two consecutive trials, using two brood frames from two different colonies. The quantification of the transcript changes were conducted as discussed previously.

**Sequencing of Noncoding Regions of gemini and Characterization of tae1.** The putative promoter region and genomic sequences between the exons were analyzed using phenol/chloroform-extracted genomic (gDNA) from parasitic A. m. capensis clone workers and A. m. scutellata workers from the University of Pretoria. PCRs were run with gDNA from one individual, each with intron-spanning primers (Table S1). Sequencing of fragments <1 kb were done by standard PCRs using the BiothermTM DNA Polymerase (Genecraft), following the manufacturer’s instructions. Products were directly sequenced. Fragments >1 kb were processed by long-range PCRs using the Kapa2G Robust PCR Kit (Peqlab). The PCR programs consisted of an initial heating at 95 °C for 3 min, followed by 40 cycles of 30 s at 95 °C, 30 s at 54 °C, and 2 min at 72 °C and ended with a final elongation at 72 °C for 20 min. PCR products were purified (QIAquick PCR Purification Kit; Qiagen) and cloned (see above). Clones were sequenced by primer walking, and sequences were compared with the annotated honeybee genome (Amel_4.0). Sequence differences in intron 5_6 (tae1) in A. m. capensis workers were thereafter confirmed by PCRs spanning the region of interest using primers I2500I and I21000II and subsequent sequencing in four workers of each subspecies. Length differences caused by the deletion of the tae1 were thereafter confirmed in 130 A. m. capensis workers and 10 A. m. scutellata workers by PCRs using primers gem_taeI and gem_taeII.

**Acknowledgments**

We thank Berit Langer, Christoph Eller, Denise Kleber, and Petra Leibe for their technical assistance. We are also grateful for the constructive comments of two unknown reviewers, which substantially improved the manuscript. Financial support was granted by the Deutsche Forschungsgemeinschaft (R.F.A.M.) and the National Research Foundation of South Africa.
References

5. Alternative splicing of gemini


Supporting Information

Fig. S1. GPDH-, rp49-, and EF-normalized gene expression of the transcription factor homologs ATF2 and gemini in ovaries of different castes. Both genes are situated within the th locus and were prime candidate genes causing thelytoky in the African subspecies A. m. capensis. Significant differences between different castes are indicated by **P = 0.0025 (Mann–Whitney U test).
Fig. S2. GPDH- and EF-normalized gene expression of the four different transcripts of *gemini* after splicing of exon 5 (Upper) and exon 7 (Lower) (schematic on Right) plotted on a logarithmic scale. Different letters (P ≤ 0.05) and asterisks (*P ≤ 0.05; **P ≤ 0.01) indicate significant differences (Mann–Whitney U tests with P values corrected for multiple testing where necessary). n.d., not detected.
**Fig. S3.** Alignment of the intron between the spliced exon 5 and the following exon. The first line represents the annotated sequence (Amel 4.0), whereas the other two stand for the
intrinsic sequences of the African subspecies A. m. scutellata and A. m. capensis. The 9-bp deletion (tae1) spans from base 201 to 210 within the parasitic cape-clone sequence.
Table S1. Primer sequences

<table>
<thead>
<tr>
<th>Method</th>
<th>Primer</th>
<th>Sequence (5'→3')</th>
<th>Product size (bp)</th>
<th>qRT Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Splice site detection</strong></td>
<td>4I</td>
<td>TTACGGCACGGATTAATACAACCTCA</td>
<td>2299</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18II</td>
<td>CGCAATAAGTTGCTTTTGTCTTAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5'-RACE</strong></td>
<td>QgemI</td>
<td>TTTTTTTTTTTTTTTTGTATACTGAAATGC</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Qgem4I</td>
<td>TTTTTTTTTTTTTTTTTGCGTTAGATCAG</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>5II</td>
<td>ACCATTTCAGCGTAAACTGCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sp5II</td>
<td>ACATGTTTTAAACATATTGGTA</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td><strong>3'-RACE</strong></td>
<td>16I</td>
<td>TGTTTAGTTTACGGATATGGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5I</td>
<td>CATCCATGGCAGTTATAAAGCA</td>
<td>210</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td>6II</td>
<td>ATGATGTTAGGAGGATATCTGT</td>
<td>165</td>
<td>2.04</td>
</tr>
<tr>
<td><strong>qRT of gemini</strong></td>
<td>Sp7.2I</td>
<td>AATTTTGAAGTGGACAATGC</td>
<td>196</td>
<td>1.97</td>
</tr>
<tr>
<td></td>
<td>8II</td>
<td>TTTGGTAGGATCCCACATAATCTTCT</td>
<td>220</td>
<td>2.21</td>
</tr>
<tr>
<td></td>
<td>Sp7.3I</td>
<td>CATTCAAACGTTCCTTTTTCTC</td>
<td>168</td>
<td>2.16</td>
</tr>
<tr>
<td></td>
<td>6II</td>
<td>ATGATGTTAGGAGGATATCTGT</td>
<td>243</td>
<td>1.79</td>
</tr>
<tr>
<td></td>
<td>Sp5II</td>
<td>CATTCAAACGTTCCTTTTTCTC</td>
<td>582</td>
<td>2.15</td>
</tr>
<tr>
<td><strong>qRT of ATF-2</strong></td>
<td>ATF2_I</td>
<td>GATTTGAGAAATGCAAGGGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATF2_II</td>
<td>TGGATACCGTTTCTGTCTGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>House-keeping genes</strong></td>
<td>Rp49I</td>
<td>TGGTGACATTTGACAGAGCA</td>
<td>1550</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rp49II</td>
<td>CCATTGCAATTTACGACCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sequencing of promotor region</strong></td>
<td>0I</td>
<td>TTTCAACTGACTAACCAGATTG</td>
<td>1070</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1I</td>
<td>GAAATACCGCAGAAATGCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I0_1 R4</td>
<td>ACCTTTGGATGAAAGCCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>I1_2 F1</td>
<td>TCTCATTTTCTTCGCGTT</td>
<td>2102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I2_3 R2</td>
<td>AACCGAAGCAAATGCGATG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I2_3 F1</td>
<td>TTTCCAGGTTCGCGCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I2_3 R2</td>
<td>ATGATGTTAGGAGGATATCTG</td>
<td>592</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I2_3 F2</td>
<td>TTTCTCATTCAAACGACG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I2_3 R3</td>
<td>ATGATGTTAGGAGGATATCTG</td>
<td>741</td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>I2_3 F4</td>
<td>TTTGCGAGCTTCTGGTCTG</td>
<td>1760</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I2_3 R8</td>
<td>CCCGTCTTCTTTCGACG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I2_3 F7</td>
<td>TTTGCGAGCTTCTGCTG</td>
<td>2279</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3II</td>
<td>CTTGCTCGAACGATTTGGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>3I</td>
<td>TTTACGGCAGGATAACTACA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4II</td>
<td>ATGATGTTAGGAGGATATCGT</td>
<td>1564</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I1_64I</td>
<td>TCTATATTTGAGGCAAGCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>5II</td>
<td>CTTTTGGGTTAAATACCGTCA</td>
<td>1056</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I1_2459II</td>
<td>TTTGCTCAGTAGGTTCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>I1_2459II</td>
<td>TTTGCTCAGTAGGTTCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5I</td>
<td>TTTGCTCAGTAGGTTCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I2500I</td>
<td>GATAAAGATTTACATTATTTACT</td>
<td>2477</td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>I21000II</td>
<td>GATAAAGATTTACATTATTTACT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I2500I</td>
<td>TACACAAGCTGATAAAATCTC</td>
<td>2577</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I21000II</td>
<td>GATAAAGATTTACATTATTTACT</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>6I</td>
<td>AAGAACAAATGGAACAACTCGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7II</td>
<td>ATGATGTTACCGTCAATAGCAA</td>
<td>467</td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>7I</td>
<td>CAATCCGAGCTGCAAATAGCAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8II</td>
<td>TTTGGTAGGATCCCACATAATCTT</td>
<td>552</td>
<td></td>
</tr>
<tr>
<td><strong>Introns 7-13</strong></td>
<td>7I</td>
<td>CAATCCGAGCTGCAAATAGCAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13II</td>
<td>ATACGTTGAGACAACTCGGAAGCA</td>
<td>5542</td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>12I</td>
<td>AATTACCCCTGAGGATCAATAGCAA</td>
<td>354</td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>14II</td>
<td>AACATATACACCGAGGATTCGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intron sequencing</strong></td>
<td>14I</td>
<td>GCATGTATTTCGTTGGAACCAATAGGA</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td><strong>Intron 14_15</strong></td>
<td>15II</td>
<td>CGCAATAAGTTCCTTTTCTTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>tael detection</strong></td>
<td>gem_taeI</td>
<td>ACGAGTCCGAAAGCTGGAAGCCTG</td>
<td>122/113</td>
<td></td>
</tr>
<tr>
<td></td>
<td>gem_taeII</td>
<td>ACGAGTCCGAAAGCTGGAAGCCTG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table S2. siRNA sequences used for siRNA synthesis

<table>
<thead>
<tr>
<th>siRNA Name</th>
<th>Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1.FW</td>
<td>GGATCCTAATACGACTCACTATAGTAATGCAAGCAAGGCTTC</td>
</tr>
<tr>
<td>1.1.RV</td>
<td>AAGTAAATGCAAGCAAGGCTTGCTTGTAGTGTAGGATCC</td>
</tr>
<tr>
<td>1.2.FW</td>
<td>GGATCCTAATACGACTCACTATAGTAATGCAAGCAAGGCTTC</td>
</tr>
<tr>
<td>1.2.RV</td>
<td>AAGAAGGCCCTGGCTTGTAGTATAGTGTAGGATCC</td>
</tr>
<tr>
<td>2.1.FW</td>
<td>GGATCCTAATACGACTCACTATAGTAATGCAAGCAAGGCTTC</td>
</tr>
<tr>
<td>2.1.RV</td>
<td>AAGGCCCTGGCTTGTAGTATAGTGTAGGATCC</td>
</tr>
<tr>
<td>2.2.FW</td>
<td>GGATCCTAATACGACTCACTATAGTAATGCAAGCAAGGCTTC</td>
</tr>
<tr>
<td>2.2.RV</td>
<td>AAGGCCCTGGCTTGTAGTATAGTGTAGGATCC</td>
</tr>
<tr>
<td>scr1.FW</td>
<td>GGATCCTAATACGACTCACTATAGTAATGCAAGCAAGGCTTC</td>
</tr>
<tr>
<td>scr1.RV</td>
<td>AAGATATCAAGCAAGGCTTGCTTGTAGTGTAGGATCC</td>
</tr>
<tr>
<td>scr2.FW</td>
<td>GGATCCTAATACGACTCACTATAGTAATGCAAGCAAGGCTTC</td>
</tr>
<tr>
<td>scr2.RV</td>
<td>AAGTCTAGGACCTCTCAGAGGCTTGCTTGTAGTGTAGGATCC</td>
</tr>
</tbody>
</table>

T7 promoter regions are underlined.

Table S3. Sample overview: Laying worker characteristics of *A. mellifera capensis* and other *A. mellifera subspecies* and their control by the thelytoky gene

<table>
<thead>
<tr>
<th>All other <em>A. m. subspecies</em></th>
<th><em>A. m. capensis</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Worker offspring</td>
<td>male, haploid</td>
</tr>
<tr>
<td>Mode of parthenogenesis</td>
<td>arrhenotoky</td>
</tr>
<tr>
<td><em>Thelytoky</em>-allele (<em>tae1</em>-element)</td>
<td>/-</td>
</tr>
<tr>
<td>Splice pattern</td>
<td>-/+</td>
</tr>
<tr>
<td>Non-laying</td>
<td>5+5/L7 S7</td>
</tr>
<tr>
<td>Laying</td>
<td>Laying</td>
</tr>
<tr>
<td>-5/L7 S7</td>
<td>5+5/L7 S7</td>
</tr>
</tbody>
</table>

The predominant splice forms are underlined.
6. Summary

The understanding of proximate mechanisms of altruistic behaviour and reproductive abstinence in eusocial insects is based on a broad empirical and theoretical fundament. Nevertheless, the evolutionary origins of genetic mechanisms responsible for altruism still remain elusive. Cape honeybees possess several reproductive traits including the unique ability to parthenogenetically produce diploid female offspring even in the presence of the queen (Pettey 1922; Pirk et al. 2002; Beekman et al. 2009). Workers of this subspecies can produce queen like amounts of the queen pheromones (Plettner et al. 1993; Crewe and Velthuis 1980; Zheng et al. 2010), undermine the queen’s reproductive monopoly and establish genetically almost identical clones of themselves as social parasitic pseudoqueens in subsequent generations (Holmes et al. 2010; Moritz et al. 2011). The parasitic Cape bees may thus be in a “potential reverse social evolution” overcoming the hegemony of the queen by challenging her status as the only reproductive individual (Gadagkar 1997). Thus this subspecies of the honeybee is a most promising model organism for studies on the evolution of reproductive altruism and its genetic background.

6.1. Functional gene studies of the $th$ locus

The core of the parasitic capensis worker syndrome lies in the thelytokous parthenogenesis. Genetic analyses revealed that this trait (Ruttner 1988), as well as two other reproductive worker traits are under control of a single recessive allele (Lattorff et al. 2005; Lattorff et al. 2007). The mapped region contains 15 annotated genes which putatively code for proteins covering very different physiological functions. The work presented here shows about half of them to be expressed in a caste specific manner in honeybee ovaries (see chapter 4), similar to the differential expression of the transcriptome in brains of workers and queens (Grozinger et al. 2007). It is not surprising to see this large number of differentially expressed genes in the light of the fundamental phenotypic differences between the two female honeybee castes. Yet only three genes in the putative $th$ region consistently correlated with the reproductive capacity of the individual (worker non-reproductive < worker reproductive < queen) as one would expect for a gene involved in the regulation of reproductive hierarchies.

One of the genes expressed in a caste specific manner –gemini– has been regarded as a putative candidate causing thelytoky in cape bees. Gemini is a member of the CP-2
transcription factor family and its ortholog in *Drosophila* is known to interact with the spindle-F protein which is involved in oocyte axis determination and oocyte microtubule cytoskeleton organization (Giot et al. 2003). This is of particular interest, as the diploidy of worker laid eggs in Cape bees is restored by an abnormal spindle rotation during meiosis (Verma and Ruttner 1983). Since *gemini*’s splice pattern correlates with the two female castes and modes of parthenogenesis and whose exon 5 was shown to directly control ovary activation in arhenotokous workers, it is now regarded as the most promising candidate within the *th* locus to maintain or in case of Cape bees undermine the reproductive monopoly of the queen caste (see chapter 5).

### 6.2. *gemini* – functional and evolutionary aspects

Epigenetic DNA methylation, which is dependent on different nutritional states, was suggested to be linked to gene regulation via the control of alternative splicing (Lyko et al. 2010). Indeed *gemini* shows differential methylation in queens and workers (in both larvae and in adult brains), and might partly regulated by epigenomic modifications (Lyko et al. 2010; Lyko and Maleszka 2011; personal communication) in addition to the allelic differences suggested in this thesis. Clusters of methylated CpGs close to the differentially spliced exons may adjust the DNA accessibility after methylation regulating differential spliced transcripts. Thus methylation of *gemini* might be the link between differential nutrition of larvae of different castes (reviewed in Rembold 1964) and the control of alternative splicing, which in turn controls ovary activation. Moreover, Cape honeybee workers are preferentially fed in foreign host colonies (Beekman et al. 2000; Allsopp et al. 2003). This preferential feeding in combination with other behavioural and physiological pre-dispositions increases their reproductive potential (Neumann and Hepburn 2002; Zheng et al. 2010). Thus, methylation of *gemini* might provide the link between the environmental benefit of preferential feeding and alternative splicing which correlates with the parasitic phenotype of Cape honeybees.

Such a combination of genetic and nutritional determination of caste has been shown in the stingless bee *Melipona beecheii* (Jarau et al. 2010). The proportion of females (25 %) developing into queens after the provision of food where geraniol has been added, supports the two-locus two allele model previously stated by Kerr (1948, 1950a, b). Double heterozygous individuals can develop into queens if fed with
sufficient geraniol whereas all other genotypes develop into the worker caste irrespective of the diet.

Another example of a single allele control of eusocial structures has previously been shown in ants. Polymorphism of a colony’s social organization in the red imported fire ant, *Solinopsis invicta*, is associated with allelic variation of a single gene, General protein-9 (*Gp-9*) (Ross 1997; Gotzek and Ross 2007). Allelic states of *Gp-9* in workers decide on the development of monogynous or polygynous colonies (Ross 1997; Ross and Keller 1998; Krieger and Ross 2002; Gotzek et al. 2007) which changes the colonial phenotype including aggression, colony foundation and reproductive output of colonies (Bourke and Franks 1995). The product of *Gp-9* is a member of the insect odorant binding protein (OBP) family (reviewed in Gotzek and Ross 2009). *Gp-9* was suggested to function as a molecular chemoreception transducer where the queens’ pheromones serve as ligands. Workers which carry different alleles differ in their binding properties of these ligands and therefore exhibit different queen recognition capabilities (Krieger and Ross 2002; Krieger 2005).

*Gp-9* is accepted generally as a “green-beard” gene (Keller and Ross 1998) in coding for a perceptible trait which is recognized by others and which in the end leads to preferential treatment of its carriers (Dawkins 1976). The latter aspect also holds true for selfish genes which by definition encode for phenotypes that increase their transmission to future generations (Dawkins 1976). Following this definition, the *tae* (see chapter 5) which is very likely involved in the regulation of thelytokous parthenogenesis in Cape bees can be regarded as a selfish genetic element. With the promotion of asexual reproduction and thus the transmission of identical genotypes into the next generation *tae* promotes its own spread through a population. The wild type allele on the other hand which is associated with the arrhenotokous mode of parthenogenesis and a non-queenlike phenotype can be regarded as an altruistic counterpart. According to the inclusive fitness theory (Hamilton 1964), maintenance and spread of altruistic alleles is governed by factors that may also counteract selfish genetic elements. Empirical data on less fit honeybee colonies consisting of dominant workers of Cape bees that do not engage in other tasks than egg-laying (Hillesheim et al. 1989) suggest that colony level selection may indeed decrease the number of selfish genetic elements in *A. m. capensis*.

Loss of fitness of honeybee colonies due to reproductive workers may also be the reason for the maintenance of two distinct *Apis* subspecies in South-Africa differing
substantially in their mode of parthenogenesis. A hybrid zone exists between the endemic ranges of *A. m. capensis* and *A. m. scutellata* in which both subspecies interbred without causing a breakdown of the zone or loss of the distinct characteristics of each subspecies (Hepburn and Crewe 1991). Here workers of queenless colonies produce male offspring by arrhenotokous parthenogenesis as well as female offspring by thelytokous parthenogenesis (Hepburn and Crewe 1991). Furthermore, F1 matings of *A. m. capensis* and *A. m. scutellata* hybrid colonies produce viable colonies without signs of reproductive parasitism by *A. m. capensis* (Crewe and Allsopp 1994). Beside the theory of pure ecologically driven stabilization of the hybrid zone (Hepburn and Crewe 1991), genetically mixed colonies may show a reduced fitness from the inadequate regulation of worker reproduction (Beekman et al. 2008). Moreover, molecular aspects like chromosomal incompatibilities between parental populations (Fel-Clair et al. 1996) as can be seen in mice and regulatory incompatibilities between transcription factors and the genes that they control (Noor and Feder 2006) may also reduce the fitness of hybrids compared to their parental types. The latter aspect might be of particular interest as the work presented here shows that the transcription factor *gemini*, probably due to an alteration of the intronic sequence flanking the alternatively spliced exon 5 (*tae*; see chapter 5), is differentially spliced between workers of arrhenotokous and thelytokous subspecies. Knock-down techniques developed and presented in this thesis (see chapter 2 and 3) allowed to determine the phenotypic outcome of selective repression of transcripts containing exon 5. Indeed, exon 5 deficient transcripts significantly increased the frequency of reproductive arrhenotokous workers. Moreover, the amount of exon 5 deficient transcripts is generally higher in thelytokous Cape bees and in case of reproductive Cape workers resembles the splice pattern of arrhenotokous queens. Thus, altered splice patterns between the different subspecies might affect different down-stream gene cascades which in turn cause the queen-like phenotype of *Capensis* bees. This might predispose Cape honeybee workers to evade reproductive altruism and might also be responsible for the speciation process between *A. m. capensis* and *A. m. scutellata*.

In conclusion, this thesis provides an important clue on the genetic control which maintains the well balanced system of reproductive hierarchies in honeybee colonies. For the first time this work shows how alternative splicing generates multiple transcript isoforms which regulate the reproductive physiology of female honeybees. These results offer new possibilities to study the genetic mechanisms behind the alternate
splice pattern of gemini which might involve methylation processes as well as subsequent gene cascades. Once the actual allelic switch has been molecularly and functionally characterised, it will be possible to study selection processes that maintain the balance between altruistic A. m. scutellata workers and selfish in A. m. capensis in the Southern part of Africa. Even now it appears as an exceptional example, how a single selfish genetic element may reverse evolutionary processes leading to a complete break down of eusocial structures and bringing back selfish individuals which exclusively act for their own benefit.

6.3. References


7. Zusammenfassung


7.1. Funktionelle Genstudien am th-Locus

7. Zusammenfassung

der beiden weiblichen Kasten. Nur drei Gene in dieser Region zeigten ein
Expressionsmuster entsprechend dem individuellen reproduktiven Potential (sterile
Arbeiter ‹ legende Arbeiter ‹ Königin).

Eines der Gene (\textit{gemini}), dass in Kastenabhängiger Art und Weise exprimiert
wird, war vom Beginn der Untersuchungen an als eines der interessantesten
Kandidatengene betrachtet worden. \textit{Gemini} codiert einen Transkriptionsfaktoren der
zur CP-2 Familie gehört. Dessen Ortholog interagiert in \textit{Drosophila} mit dem Spindel-F
Protein, dass für die Achsenfestlegung in der Oocyte und für die Organisation des
Mikrotubuli-Cytoskelletts in der Oocyte verantwortlich ist (Giot et al. 2003). Letzteres
ist von besonderem Interesse, da die Diploidie in den von Kapbienen gelegten Eiern
durch eine unnormale Spindelrotation während der Meiose verursacht wird (Verma und
Ruttner 1983). \textit{Geminis} Transkriptmuster nach alternativem Spleißen korreliert sowohl
mit der Art der untersuchten Kasten als auch mit den unterschiedlichen Arten der
Parthenogenese. Zudem sind Transkripte, die das Exon 5 enthalten, direkt für die
Steuerung der Ovaraktivierung in arrhenotoken Bienen verantwortlich (siehe Kapitel
5). So kann \textit{gemini} nach dieser Studie als der augenscheinlichste und wahrscheinlichste
Kandidat unter den 15 Genen des \textit{th} locus betrachtet werden, der die thelytoke
Parthenogenese in Kapbienen verursacht und damit für die Untergrabung des
Fortpflanzungsmonopols der Königin verantwortlich ist.

7.2. \textit{gemini}-funktionelle und evolutionäre Aspekte

Epigenetische Steuerung mittels DNA-Methylierung wurde als möglicher
Mechanismus zur Steuerung der Genexpression mittels alternativem Spleißen
vorgeschlagen (Lyko et al. 2010). Dieser regulative Mechanismus wird unter anderem
durch unterschiedliche Ernährungsmuster gesteuert. Auch bei \textit{gemini} zeigten sich
variable Methylierungsmuster in larvalen und adulten Arbeitern und Königinnen (Lyko
et al. 2010; Lyko und Maleszka 2011; mündlicher Austausch), welches neben oder
verursacht durch die allelischen Unterschiede für das unterschiedliche Spleißmuster
verantwortlich sein könnten. Ansammlungen methylierter CpG-Inseln in der Nähe von
alternativ gespleißten Exons könnten nach methylierungsbedingten
Konformationsänderungen den Zugang zur DNA verändern, und damit das Entfernen
verschiedener Exonbereiche verhindern. So könnten unterschiedliche
Methylierungsmuster die Verbindung zwischen kastenspezifischer unterschiedlicher
Ernährung der Larven (zusammengefasst von Rembold 1964) und alternativem Spleißen
7. Zusammenfassung

Solch eine Kombination aus genetischer und ernährungsbedingter Steuerung reproductiver Hierarchien wurde auch in stachellosen Bienen gefunden (Jarau et al. 2010). 25 % aller Bienen die mit Geraniol angereichertem Futter aufgezogen wurden, entwickelten sich zu Königinnen. Dieser Prozentsatz unterstützt die These, dass die Königinnenwerdung in Melipona von zwei Regionen, die heterozygot ausgeprägt sein müssen, gesteuert wird (Kerr 1948, 1950a, b). Demnach entwickeln sich doppelt heterozygote Individuen nur dann zur Königin, wenn dem Futtersaft ausreichend Geraniol zugesetzt worden ist. Alle anderen Genotypen entwickeln sich zwangsläufig zu Arbeiterinnen, unabhängig davon, wie sie ernährt werden.

7. Zusammenfassung

Fähigkeiten die Königin wahrzunehmen und akzeptieren je nach Genotyp eine unterschiedliche Anzahl an Königininnen (Krieger und Ross 2002; Krieger 2005).


Ein, durch sich selbst reproduzierende Arbeiter verursachter Selektionsnachteil könnte eine schlechtere Anpassungsfähigkeit von Kapbienen zur Folge haben. Dies wiederum könnte die Ursache sein, dass in Süd-Afrika in räumlicher Nähe zwei verschiedene Unterarten zu finden sind, die sich vor allem in ihrer Art der parthenogenetischen Fortpflanzung unterscheiden. Zwischen den endemischen Gebieten dieser beiden Unterarten findet sich eine Vermischungszone. Hier kommt es zum genetischen Austausch, ohne das die Zone oder die Charakteristika der beiden Unterarten verschwinden (Hepburn und Crewe 1991). In dieser Zone werden in weisellosen Völkern sowohl männlicher Arbeiternachwuchs mittels Arrhenotokie, als auch weiblicher Arbeiterinnennachwuchs mittels thelytoker Parthenogenese generiert.
7. Zusammenfassung


Zusammenfassend lässt sich sagen, dass diese Arbeit essentielle Einblicke in die genetische Kontrolle gewährt, die das Gleichgewicht zwischen Reproduktion und Altruismus in eusozialen Insektenkolonien bewahrt. Zudem zeigt diese Arbeit zum ersten Mal, dass auch in Honigbienen die Vielfalt an Transkripten durch alternatives Spleiß hergestellt wird. Diese unterschiedlichen Transkriptvarianten haben direkten
8. Appendix

8.1 Declaration on the contributions to the manuscripts/papers on which this thesis is based


Experimental set-up; injections and microscopy, sample preparations and qRT-measurements, data analyses, statistical analyses, manuscript writing.


Experimental set-up; injections, sample preparations and qRT-measurements, data analyses, statistical analyses, manuscript writing.


Experimental set-up; sample preparations and qRT-measurements, RNAi, data analyses, statistical analyses, manuscript writing.

4. Jarosch, A. und Moritz, R.F.A Molecular characterization of a genomic region which controls reproductive hierarchies in honeybees (Apis mellifera sp.). Submitted to Insect Biochemistry and Molecular Biology.

Experimental set-up; sample preparations and qRT-measurements, data analyses, statistical analyses, manuscript writing.
8.2. Acknowledgements

This work was conducted in the Institute of Biology within the Molecular Ecology research group of Prof. Robin F. A. Moritz with financial support from the Deutsche Forschungsgemeinschaft (DFG) and the National Research Foundation of South Africa.

First of all I would like to thank Prof. Robin F. A. Moritz for introducing me into the field of evolutionary genetics of honeybees. I am grateful for the support during all stages of this thesis and especially for his encouragement after disappointing days in the lab and after every rejection we received during the years. Moreover, without our long and in the end fruitful discussions this thesis would not exist in its current form.

Furthermore I would like to thank my co-authors Eckart Stolle and Prof. Crewe for the support and improvements on the manuscripts. Those also got substantially improved by comments of other persons who were Prof. Rob Paxton, Dr. Tom Murray, Dr. Bernd Kraus, Jonathan Kidner, Silvio Erler and Dr. Mathias Becher.

My thanks to the Developmental genetics group of Prof. Reuter and especially to Manueler Gabler for giving me the opportunity to work in their lab and for their support on their LSM. I also thank all of the students which were part of this work during their practical and diploma thesis. Among them I would like to give a special mention to Berit Langer and Christoph Eller, who set up protocols in the lab and conducted experiments which were essential for this thesis.

My very special thanks to all current and some former members of the MolEcol Research group in Kröllwitz. The warm, creative and very often amusing atmosphere made working much more enjoyable. My special regards to Dr. Michael Lattorff, Dr. Stephan Härtel, Dr. Mathias Becher, Dr. Rodolfo Jaffé, Dr. Bernd Kraus, Silvio Erler, Dr. André Walter, Dr. Stephan Wolf, Mario Popp, Dieter Behrens (I appreciate your humour!), Ben Barth, Holger Scharpenberg and Bertrand Fouks. Especially the female members of this group which are Petra Leibe + Weber, Denise Kleber, Anett Huth-Schwarz and Theresa Toev need to be mentioned as they did not only assist in the lab but moreover became good friends. It was a pleasure to work with all of you.

Last but not least I need to thank all of my friends, my family and Gabriel for their support and patience. This work is dedicated to you.
8.3 Curriculum vitae

Antje Jarosch

Date of birth 06.01.1981
Place of birth Hennigsdorf, Germany
Marital status unmarried
Nationality German

Education

1987 – 1993 Elementary school – Velten Stadt
Degree: Abitur (Grade: 2.0, ‘good’)

Higher Education

10/2000 – 03/2006 Studies of Biology
Humboldt-Universität zu Berlin
Major: biochemistry
Minors: microbiology, ecology

Diploma thesis: Leibniz Institute of Freshwater Ecology and Inland Fisheries
„Oxidative stress caused by cyanobacteria in Daphnia magna: Comparisons between enzymatic reactions and genetic regulations“
Degree: Diploma (Grade: 1.1, ‘very good’)

04/2006 – 04/2009 Research assistant
Martin-Luther-University Halle-Wittenberg, Institute for Biology-Zoology, Molecular Ecology
DFG-Project: “Genetic control of reproductive hierarchies in honeybees (Apis mellifera)”
(coordinator: Prof R.F.A. Moritz)
8. Appendix

04/2010 – 06/2011 **Research assistant**

Martin-Luther-University Halle-Wittenberg, Institute for Biology-Zoology, Molecular Ecology

BMBF Project Fugapis: “Functional genome analysis of disease resistance in honeybees (*Apis mellifera*)”

(Coordinator: Prof. Dr. Robin F.A. Moritz)

Since 04/2006 **Ph.D. student**

Martin-Luther-University Halle-Wittenberg, Institute for Biology-Zoology, Molecular Ecology

Supervisor: Prof. Dr. Robin F.A. Moritz

Title: “Genetic control of reproductive hierarchies in honeybees”
8.4. Publications

Peer review


Jarosch, A. und Moritz, R.F.A Molecular characterization of a genomic region which controls reproductive hierarchies in honeybees (Apis mellifera sp.). Submitted to Insect Biochemistry and Molecular Biology.

Jarosch A., Gätschenberger H., Gimple O., Frey E., Rosenkranz P., Tautz J. and Moritz R.F.A. A single QTL in charge of Noduli production is suspected to be a major antimicrobial defence mechanism in honeybees (Apis mellifera), in preparation.
8.5. Erklärung

Hiermit erkläre ich an Eides statt, dass diese Arbeit von mir bisher weder an der Naturwissenschaftlichen Fakultät I - Biowissenschaften der Martin-Luther-Universität Halle-Wittenberg noch an einer anderen wissenschaftlichen Einrichtung zum Zweck der Promotion eingereicht wurde.


Halle (Saale), den 17. November 2011

........................................
Antje Jarosch