The transcriptomic changes associated with the development of social parasitism in the honeybee *Apis mellifera capensis*

Denise Aumer[1*], **Fiona Mumoki**[2], **Christian W.W. Pirk**[2], **Robin F. A. Moritz**[1,2,3]

[1] Department of Molecular Ecology, Martin-Luther University Halle-Wittenberg, Hoher Weg 4, 06099 Halle (Saale), Germany

[2] Department of Zoology and Entomology, University of Pretoria, Private Bag X20 Hatfield 0028, Pretoria, South Africa

3 German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany

* Corresponding author: denise.aumer@zoologie.uni-halle.de, Phone: (0049)-345-5526235

Keywords: worker reproduction / reproductive dominance / social parasitism / *Apis mellifera capensis* / thelytoky / pseudoqueen

Acknowledgement: We are grateful to the members of the Research Unit FOR2281 'Sociality and the reversal of the fecundity-longevity trade-off', especially to Daniel Elsner and Mark Harrison, for their valuable input regarding the data analysis. We also thank Nikita Venter, Cathy Bester and Alex Nepomuceno for their assistance in handling the experimental bees. Financial support was granted by the Deutsche Forschungsgemeinschaft (RFAM).

Abstract

Social insects are characterised by the division of labour. Queens usually dominate reproduction, whereas workers fulfil non-reproductive age-dependent tasks to maintain the colony. Although workers are typically sterile, they can activate their ovaries to produce their own offspring. In the extreme, worker reproduction can turn into social parasitism as in *Apis mellifera capensis*. These intraspecific parasites occupy a host colony, kill the resident queen and take over the reproductive monopoly. Because they exhibit a queenlike behaviour and are also treated like queens by the fellow workers, they are so-called "pseudoqueens". Here we compare the development of parasitic pseudoqueens and social workers at different time points using fat body transcriptome data. Two complementary analysis methods – a principal component analysis and a time course analysis – led to the identification of a core set of genes involved in the transition from a social worker into a highly fecund parasitic pseudoqueen. Comparing our results on pseudoqueens with gene expression data of honeybee queens revealed many similarities. In addition, there was a set of specific transcriptomic changes in the parasitic pseudoqueens that differed from both, queens and social workers, which may be typical for the development of the social parasitism in *A. m. capensis*.

Introduction

Eusocial insects are characterised by the reproductive division of labour between the female castes: The workers remain sterile and fulfil tasks to maintain the colony including brood care and foraging, whereas the highly fecund queens monopolise reproduction (Winston 1987). In honeybees, *Apis mellifera* L., queens enjoy the reproductive monopoly because they suppress ovary activation in workers with their queen's mandibular gland pheromones (QMP; Butler 1959; Slessor et al. 1988; Plettner et al. 1993; Winston and Slessor 1998; Hoover et al. 2003). The genetic basis associated with enhanced fecundity and suppression of worker ovary development is well studied. A key role is played by the multifunctional yolk protein Vitellogenin (Vg) that is significantly higher expressed in fecund queens compared to non-reproductive workers (Engels 1974; Engels and Fahrenhorst 1974; Amdam et al. 2003; Corona et al. 2007) and likely involved in ovary activation (Fleig 1995; Lin et al. 1999; Seehuus et al. 2007). In addition, comparisons between newly emerged and sexually mature queens (Wu et al. 2016), as well as between the two female castes identified many DEGs. Importantly, metabolic genes, including oxidoreductases, and immune genes were significantly higher expressed in queens compared to workers (Grozinger et al. 2007), suggesting profound differences between the two castes.

However, even though worker sterility is the rule, workers can occasionally activate their ovaries. This usually happens whenever the queen is lost and cannot be replaced. Laying workers normally produce male offspring by arrhenotokous parthenogenesis (Dzierzon 1845; Winston 1987). As an exception, workers of the South African subspecies *A. m. capensis* Escholtz, the Cape honeybee, can produce female offspring by thelytokous parthenogenesis (Onions 1912). In addition, generally *A. m. capensis* workers are distinctively different to workers of all other honeybee subspecies in traits related to fertility: They are usually characterized by a high number of ovarioles (Hepburn and Crewe 1991; Phiancharoen et al. 2010), the presence of a well-developed spermatheca (Onions 1914; Ruttner 1988), the ability for rapid ovary activation (Anderson 1963; Ruttner and Hesse 1981; Moritz and Hillesheim 1985), the secretion of queenlike proportions of QMP components (Hemmling et al. 1979; Crewe and Velthuis 1980; Moritz et al. 2002) and the evading of policing by other workers (Moritz et al. 1999; Simon et al. 2001). Genetically the entire phenotype seems to be controlled by a single gene

(Ruttner 1988; Lattorff et al. 2005, 2007; Aumer et al. 2017), although this gene has not yet been identified (Aumer et al. 2017). Because of all their reproductive capacities, *A. m. capensis* workers can develop into so-called "false queens" or "pseudoqueens" (Sakagami 1958; Velthuis et al. 1990). These pseudoqueens do not only regain fertility, but they also act like queens and are treated like such by the surrounding sterile workers (Velthuis et al. 1990; Moritz et al. 2000). However, whether an *A. m. capensis* worker develops into a pseudoqueen strongly depends on its environment. While in queenright *A. m. capensis* colonies workers usually refrain from reproduction (Hoover et al. 2003; Härtel et al. 2006), under queenless conditions they strongly compete among each other for the reproductive control. Only the most dominant workers that produce the highest amounts of queenlike mandibular gland secretions can establish themselves as pseudoqueens in the queenless hive. The other workers remain sterile due to the strong queenlike mandibular gland signal of the pseudoqueens (Moritz et al. 1996, 2000, 2004; Härtel et al. 2011). In turn, whenever *A. m. capensis* workers are in a colony of another *A. mellifera* subspecies, they can escape the suppressive effect of the host queen's mandibular gland pheromones and can activate their ovaries in the presence of the queen (Pirk et al. 2002; Martin et al. 2002b). Therefore, in colonies of other honeybee subspecies *A. m. capensis* workers regularly develop into pseudoqueens (Johannsmeier 1983, 1992; Koeniger and Würkner 1992; Martin et al. 2002b; Neumann and Hepburn 2002). As a consequence *A. m. capensis* pseudoqueen workers are potent social parasites that can gain entry to foreign colonies, replace the host colony's queen, take over the nest and start producing clonal offspring by thelytokous parthenogenesis (Neumann and Moritz 2002). The parasite's offspring manipulates the host workers in order to get additional feeding, which has been suggested to be a causal mechanism for the enhanced reproductive capacities of *A. m. capensis* social parasites (Beekman et al. 2000). At the end of the 20th century the social parasitism of *A.m. capensis* workers had major impact on the commercial beekeeping in South Africa. A single clonal lineage of *A. m. capensis* workers (Baudry et al. 2004) invaded colonies of the adjacent honeybee subspecies *A. m. scutellata* and killed many thousand colonies of South African beekeepers, which is known as the so called "Capensis Calamity" (Allsopp 1992; Allsopp and Crewe 1993). As these parasitic bees are still infesting *A. m. scutellata* colonies

and causing extensive colony losses in the North of South Africa (Pirk et al. 2014), they provide an excellent test system to study the transitions from a social to a parasitic worker.

Recently, two theoretical models for the evolution of social parasitism in eusocial insects were proposed (Cini et al. 2015). Both models are based on the assumption that eusociality evolved from a solitary lifestyle through specialisation of the female castes. Queens and workers arise from the same genome, but display great phenotypic plasticity due to differential gene expression (Barchuk et al. 2007) and caste-specific gene regulatory mechanisms (f. e. via epigenetics, Lyko et al. 2010; Weiner and Toth 2012). Social parasites of eusocial species are often closely related to their hosts (= Emery's rule, Emery 1909) and usually engage in reproduction (Wilson 1971). Therefore, they might have evolved from their eusocial ancestor by either 'deleting' the worker phenotype and only expressing the queen phenotype (= 'Phenotype Deletion Model'), or they evolved by developing a new/modified phenotype (= 'Phenotype Shift Model') (Cini et al. 2015). At the transcriptomic level, the 'Phenotype Deletion Model' predicts no overlap between the gene expression of the social parasites and their host workers, but similar gene expression patterns to their host queens. In contrast, according to the 'Phenotype Shift Model' the gene expression of social parasites would reveal some novel patterns that are different to both – the host queen and worker gene expression. A combination of the two models is also potentially possible (Cini et al. 2015). To test these models a comparison of social parasites with their host's female castes seems to be obvious (Cini et al. 2015). A few studies using such systems exist, supporting the 'Phenotype Shift Model', like in the wasp genus *Polistes* (Cini et al. 2015). Further, in the social parasites of the ant genera *Pogonomyrmex* and *Vollenhovia* no evidence for a 'deletion' of genes that belong to the worker phenotype was found (Smith et al. 2015).

The intraspecific social parasitism of *A. m. capensis* workers is a highly appealing experimental model system to study the molecular mechanisms of reproductive social parasitism. It offers a system in which we can compare socially sterile and highly fecund parasitic individuals of the same species and from the same caste with the same genetic and physiological constraints. *A. m. capensis* parasitic pseudoqueens are characterized by many queenlike physiological and behavioural traits. In addition, they also display some parasite-specific features. For example after invading a colony, they can engage in lethal fights with the host queen (Moritz et al. 2003) to take over the reproductive

monopoly. As they also express a queenlike bouquet of the mandibular gland pheromones, one would expect a queenlike gene expression in combination with some novel parasite-specific patterns at the transcriptomic level. The fat body is known to be the central organ for the fundamental physiological and behavioural changes that control reproduction in honeybees (Amdam et al. 2012). Therefore, the identification of gene cascades in the fat body that contribute to the transition from a sterile social bee into a highly fecund parasite, might provide insights to whether the phenotype shift or deletion pathway was followed in the evolution of worker parasitism in the honeybee. Here we compare the development of parasitic pseudoqueens with that of social sterile workers at different time points, using full transcriptome data of fat bodies.

Methods

Experimental set-up and Sample collection

Two neighbouring South African honeybee subspecies - *A. m. capensis* and *A. m scutellata* - were used for this experiment. Brood frames of both subspecies with mostly sealed brood were collected from managed colonies that were kept at the University of Pretoria, South Africa. The *A. m. scutellata* brood frames were taken from 20 queenright *A. m. scutellata* colonies. The *A. m capensis* brood was obtained from a queenless *A. m. scutellata* host colony that was heavily infested by parasitic *A. m. capensis* laying workers. The infestation level was not only recognized by the absence of the *A. m. scutellata* queen and the presence of many laying *A. m. capensis* pseudoqueens that elicited retinue behaviour by the remaining *A. m. scutellata* workers, but also by the existence of scattered arched capped brood that is characteristic for *A. m. capensis* brood in *A. m. scutellata* colonies. All brood frames were kept in an incubator (37 °C, high relative humidity). Freshly emerged workers were collected daily and small groups were housed in standard hoarding cages (115 x 100 x 140 mm, Köhler et al. 2013). To avoid colony related differences, the freshly emerged *A. m. scutellata* workers were randomly mixed in a bucket before they were put into the hoarding cages. Each cage was provided with a piece of comb, water, food (1:1 honey/castor sugar mixture) and pollen *ad libitum* and kept in incubators at 35 °C and 60% relative humidity. To sustain the group size consistently, the cages were checked daily and dead bees were replaced by freshly emerged ones. This highly

controlled simplistic experimental set-up offers equal conditions and the possibility to closely observe the experimental groups as well as single individuals. In total three experimental groups were prepared as follows:

1) Parasitic pseudoqueen: A single *A. m. capensis* clone introduced into a group of 100 *A. m. scutellata* workers. Because *A. m. capensis* workers are reproductively dominant over *A. m. scutellata* workers, the Cape bee is the only individual that becomes reproductively active by suppressing the ovary activation of the others (Neumann and Hepburn 2002). This reproductive dominance was easily recognizable by the queenlike behaviour of the pseudoqueen and the release of retinue behaviour in the *A. m. scutellata* workers.
2) Social *A. m. capensis* worker: A group composed of 50 *A. m. capensis* workers. Because the *A. m. capensis* workers compete for reproductive dominance (Moritz et al. 1996, 2000, 2004; Härtel et al. 2011), the cages were observed until the reproductive dominants were identified (usually 5-10 minutes). One of the workers that did not show any queenlike behaviour and that did not elicit any retinue behaviour by the other workers was sampled.
3) Social *A. m. scutellata* worker: A group of 100 *A. m. scutellata* workers. One randomly chosen worker was sampled.

Since queenless colonies are more susceptible for an infestation by *A. m. capensis* workers than queen-right colonies (Tribe 1983; Woyke 1995), we used the queenless state in the experimental groups to facilitate the development of laying workers. The small group sizes (100 workers) allowed the reproductive dominant individuals to quickly establish themselves as pseudoqueens in the cages and the risk of being killed by the host workers (Velthuis 1976; Crewe 1984; Calis et al. 2005) was lower than it would have been in bigger groups. Also in the *A. m. capensis* worker control groups with only 50 bees, reproductive dominant *A. m. capensis* workers that elicited retinue behaviour in their fellow workers were observed. However, for this control group only social workers that had not elicited retinue behaviour were sampled. As in either group size (50 or 100 bees) reproductive dominant pseudoqueens and social workers were obtained, the group size did not seem to have a major effect on the experimental approach.

To follow the different temporal transitions in the transcriptome from a freshly emerged worker to a parasitic pseudoqueen, sampling was done at four consecutive time points: three, four, seven and eight days after emergence. At about days 3 and 4 *A. m. capensis* workers start to produce queenlike mandibular gland components (Simon et al. 2001; Okosun et al. 2015). By days 7 and 8 the ovaries of *A. m. capensis* workers are fully activated (Ruttner and Hesse 1981; Hepburn et al. 1991; Martin et al. 2002a) and the pseudoqueens are established in the group (Moritz et al. 2004). For every type of experimental group and time point, three replicate cages were set-up. Immediately after sampling the workers, their ovaries and fat bodies were dissected in ice cold Ringer's solution on dry ice. Dissections were carried out as described by Carreck et al. (2013). Whole ovaries and all visible fat body tissue were sampled. In addition to the behaviour observations, as proof for the phenotype the developmental stage of the ovaries of all sampled workers was determined, using the five categories of Hess (1942). For the transcriptome analyses total RNA of each fat body was solvent extracted separately, following a modified protocol from Chomczynski and Sacchi (1987). The quality and quantity of all samples was measured using a NanoDrop™ 2000 (Watham, Massachusetts) and three samples of good quality (260/280 ratio: ~2.0, 260/230 ratio: 2.0-2.2) and equal quantity per experimental group and time point were combined for sequencing.

Sequencing

All samples were sent to the Beijing Genomics Institute (BGI) in Hong Kong, China for QC analyses, library preparation and transcriptome sequencing. RNA integrities and quantities were re-checked using the Agilent 2100 BioAnalyser (Agilent, Santa Clara, CA) and a QubitTM Fluorometer (Invitrogen, Carlsbad CA). Library construction was done using the illumina® TruSeq RNA Library Prep Kit (Illumina, San Diego CA) as preparation for whole transcriptome paired-end sequencing (read length of 100 bp) on an illumina® HiSeq4000. The resulting sequences were submitted to the Sequence Read Archive at NCBI (Number*XXXXX*).

Processing of raw reads

Statistical analysis and quality control of all libraries was done with FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqch). Afterwards the adapters were trimmed from the sequences using Trimmomatic (version 0.36, Bolger et al. 2014). The resultant trimmed paired reads were mapped against the latest version of the honey bee genome Amel_4.5 (Honeybee Genome Consortium 2014, Elsik et al. 2014) using Bowtie2 (version 2-2.2.9, Langmead et al. 2009; Langmead and Salzberg 2012) and TopHat2 (version 2.0.10, Kim et al. 2013) with default settings. The accepted hits per library were sorted using SAMtools (Li et al. 2009) to generate the input file for HTSeq (version 0.7.1, Anders et al. 2015) that was used to count the number of reads per gene per library.

Identification of genes involved in parasitic pseudoqueen development

To identify genes that are involved in the transition from a sterile social worker into a highly fecund pseudoqueen, we performed a Principal Component Analysis (PCA) and a time course analysis.

The PCA was done to analyze the overall effect of the variables 'experimental group' and 'time' and to reveal the genes that contribute most to the variance in the data set. The PCA included all samples and all genes with a non-zero read count across samples. The data were variance-stabilized using the Bioconductor package *DESeq2* (Love et al. 2014) and analysed regarding the two principal components 'experimental group' and 'time' using the function 'plotPCA'. PC1 refers to the three experimental groups: *A. m. capensis* parasitic pseudoqueen, *A. m. capensis* social worker and *A. m. scutellata* social worker. PC2 refers to the consecutive sampling time points: days 3, 4, 7 and 8. For both principal components all genes with positive and negative scores were extracted. The gene expression of the genes with either positive or negative scores for PC2 'time' were tested for significant differences among the experimental groups using the Kruskal-Wallis rank sum test for each time point separately, followed by the post-hoc Dunn-Test (Dunn 1964). Significance levels were set to $p \leq 0.01$. For both principal components the top 20 genes with the highest and lowest scores that contribute most to the differentiation of the samples were analysed further.

To examine the temporal dynamics in the transcriptomic changes among the three experimental groups we used the Bioconductor package *Next maSigPro* (Nueda et al. 2014). The algorithm identifies genes with a significant different expression over time and among experimental groups, and it clusters genes with similar expression patterns together. The analysis was run with a false discovery rate of 5%. In order to avoid false positive signals (Nueda et al. 2014), the input data were reduced to the strongest significantly differentially expressed genes (adjusted p-value of ≤ 0.01 and an absolute value of ≥ 2 log2-fold-change). To identify these genes the samples of day 3 and 4 per experimental group were considered as 'early' replicates, while the samples of day 7 and 8 per experimental group were considered as 'late' replicates. Afterwards all possible pairwise comparisons within and among the experimental groups were carried out using the Bioconductor package *DESeq2* (Love et al. 2014). Due to the fact that they are potentially important candidate genes, the top 80 genes that contribute most to the principal components 'experimental group' and 'time' in the PCA were added to the input data for the time course analysis. Normalization of the input data was done with the Bioconductor package *DESeq2* (Love et al. 2014). Before running the time course analysis, the number of groups/clusters that are most likely for the input data had to be specified (Nueda et al. 2014). To identify this number the Bioconductor package *mclust* (Fraley et al. 2012) that implements Bayesian inference and the elbow method were used.

Functional analysis

For the functional analysis of the identified gene sets the program InterProScan (version-5.25-64.0, Jones et al. 2014) was used to retrieve the GO terms. The subsequent GO enrichment analyses were carried out using the Bioconductor package *TopGO* using the 'fisher' test (Alexa et al. 2006). All genes with a non-zero read count across all samples were included in the universe. The GO enrichment analyses were done for the top 20 genes with the highest and lowest scores for the principal components 'time' and 'experimental group' of the PCA, as well as for all genes included in interesting clusters revealed by the time course analysis. The genes were assigned to two categories: biological processes and molecular functions.

Results

Phenotype of experimental workers

To identify the phenotype of the experimental workers, we used behavioural observations and determined their ovary developmental stage using the Hess scale. Following Hess (1942) the ovaries were categorized into five groups, where stage 1 typifies completely inactivated ovaries and stage 5 represents fully activated ovaries, ready to lay eggs.

In general, *A. m. capensis* workers are characterized by rapid ovary activation (Anderson 1963; Ruttner and Hesse 1981; Moritz and Hillesheim 1985). Accordingly, at days 3 and 4 the ovaries of the experimental *A. m. capensis* workers were already moderately developed (n=12, on average stage 3.00 ± 0.58 SD). At days 7 and 8 all *A. m. capensis* workers were characterized by a high degree of ovary activation (n=12, on average stage 4.08 ± 0.49 SD). In contrast, the *A. m. scutellata* workers sampled at days 3 and 4 showed low levels of ovary activation (n=6, on average 1.33 ± 0.47 SD), with almost no change until days 7 and 8 (N=6, on average 1.50 ± 0.50 SD). Only workers that develop into reproductive dominant pseudoqueens show queenlike behaviour and elicit retinue behaviour in their host workers. Therefore, the *A. m. scutellata* and the *A. m. capensis* workers that did not elicit retinue behaviour served as social worker controls. Hence both ovary development analyses and behavioural observations provided decisive evidence for the phenotype of the experimental workers.

Statistical analysis of raw reads

The sequenced libraries comprised on average 65,900,726 (± 6,190,602 SD) reads per library including 98.5% (± 0.96%, SD) forward/reverse read pairs of which 70% (± 7.47% SD) were mapped to the reference genome. On average 10,613 (± 52 SD) genes were expressed per library. A total of 11,136 genes were expressed over all samples, representing 79 % of the known honeybee genes (Elsik et al 2014). Therefore our sequencing results are in line with other recent RNAseq studies in honeybees (Niu et al. 2014; Doublet et al. 2016), or they even exceed their numbers 2-10 times (Aufauvre et al. 2014; Wu et al. 2016; Wu et al. 2017).

Principal component analysis

The PCA including all samples and all 11,136 genes with a non-zero read count across samples showed a clear distinction of both the experimental groups (PC1) and the consecutive time points of sampling (PC2). The first principal component (PC1) contributes 32 % to the variance in the data set and clearly splits the experimental groups. The second principal component (PC2) contributes 19 % to the total variance and nicely reflects the consecutive time points. The two late samples (days 7 and 8) of the parasitic pseudoqueen group are clearly separated from all other samples (Fig. 1).

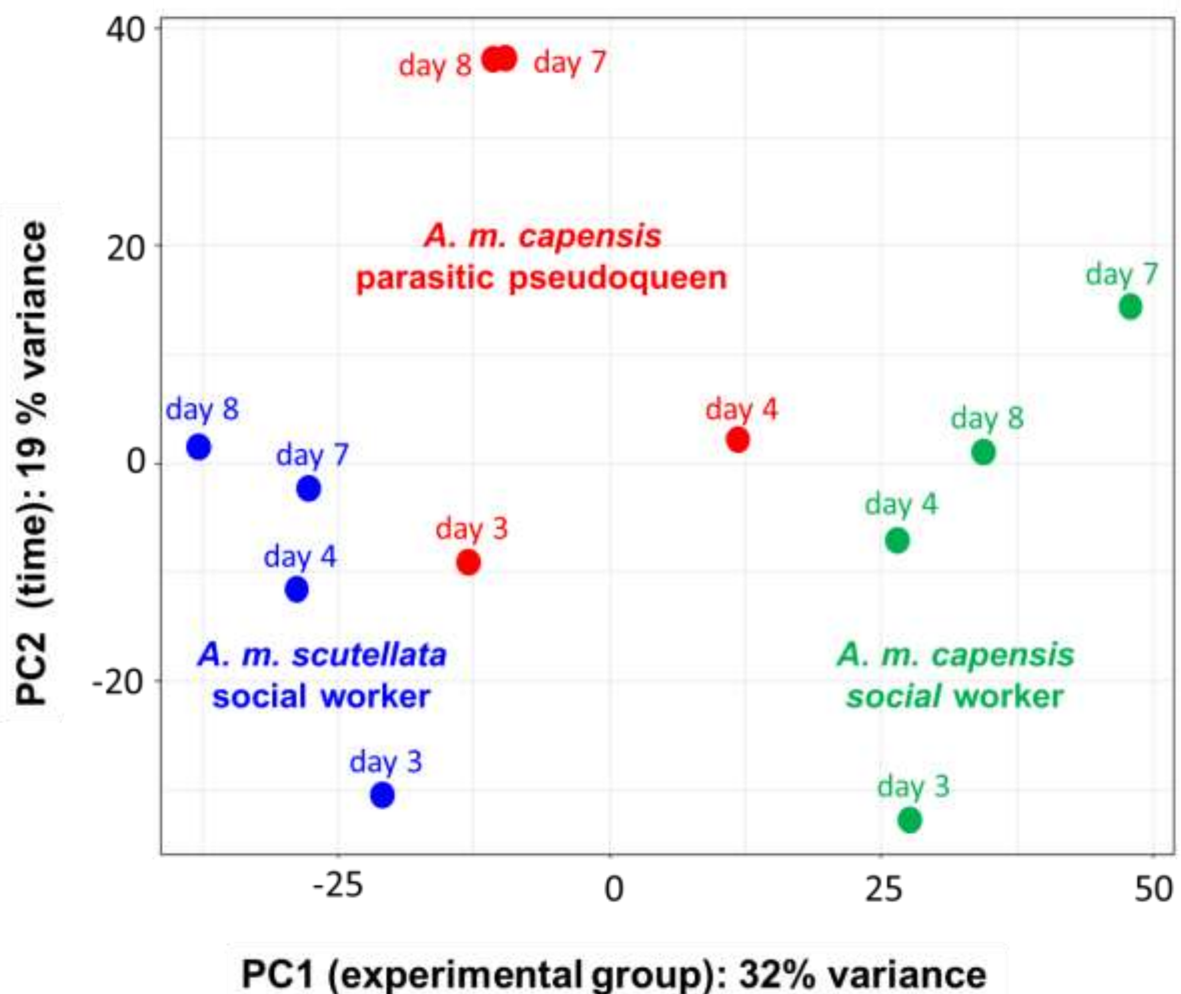

Fig. 1: Principal component analysis of variance stabilized data of all samples (day 3, 4, 7, 8) and all 11,136 genes with a non-zero read count across samples. PC1 refers to the 'experimental group' and PC2 to the consecutive time points of sampling (day 3, 4, 7 & 8). The three experimental groups are displayed in different colours: Red – parasitic *A. m. capensis* pseudoqueens, green – social *A. m. capensis* worker, blue – social *A. m. scutellata* worker

The 5,439 genes with positive scores for PC2 'time' are more strongly up-regulated over time in the parasitic pseudoqueens than in the two control groups. While there was no significant difference among the three experimental groups at days 3 and 4, the average gene expression is significantly higher in the pseudoqueens (1.33x higher) than in the two control groups at days 7 and 8 (Fig. 2a; n=5439, Kruskal-Wallis test: day7: p=1.841e-07, day8: p=7.15e-05; post-hoc Dunn test for all pairwise comparisons between pseudoqueens and controls: d7: $p \leq 8.2973e-03$, d8: $p \leq 1.7313e-03$). In contrast, for the 5,697 genes with negative scores for PC2 the opposite can be observed: they are more strongly down-regulated over time in the pseudoqueens than in the two control groups. Again there is no significant difference among the three experimental groups at days 3 and 4. Whereas at days 7 and 8 the pseudoqueens have on average a significantly lower gene expression (0.78x) compared to the control groups (Fig. 2b; n=5697, Kruskal-Wallis test: day7: p=0.0006, day8: p=8.476e-05; post-hoc Dunn test for all pairwise comparisons between pseudoqueens and controls: d7: $p \leq 0.0020$, d8: $p \leq 0.0015$).

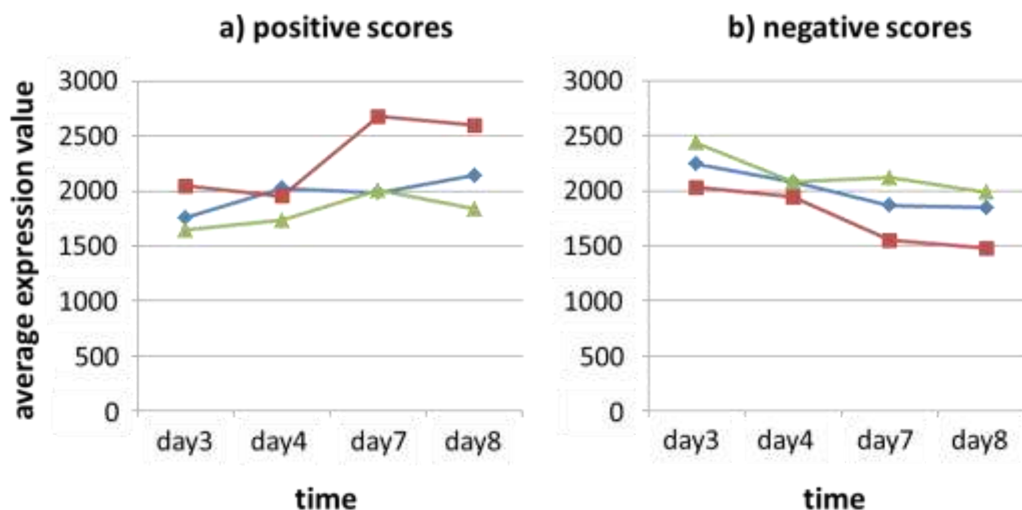

Fig. 2: The average expression value of all genes with positive (a) or negative scores (b) for PC2 'time' at the four consecutive time points: day 3, 4, 7 and 8. The three experimental groups are displayed in different colours: Red – parasitic *A. m. capensis* pseudoqueens, green – social *A. m. capensis* worker, blue – social *A. m. scutellata* worker. a) Positive scores: n=5439, Kruskal-Wallis test: day3: n.s., day4: n.s., day7: p=1.841e-07, day8: p=7.15e-05; post-hoc Dunn test for all pairwise comparisons between parasitic pseudoqueens and social controls: day3: n.s., day4: n.s., d7: $p \leq 8.2973e-03$, d8: $p \leq 1.7313e-03$. b) Negative scores: n=5697, Kruskal-Wallis test: day3: n.s., day4: n.s., day7: p=0.0006, day8: p=8.476e-05; post-hoc Dunn test for all pairwise comparisons between parasitic pseudoqueens and social controls: day3: n.s., day4: n.s., d7: $p \leq 0.0020$, d8: $p \leq 0.0015$

The total of 80 genes with the highest and lowest scores (20 genes each) for PC1 'experimental group' and PC2 'time' are listed in Online Resource 1.

Time course analysis

The pairwise comparisons among and within the experimental groups revealed a total of 451 significantly differentially expressed genes (details in Online Resource 2). Even though 20 of the 80 genes with the highest or lowest scores in the PCA were not listed in the group of genes, we nevertheless included them as potentially important candidate genes in the input for the time course analysis.

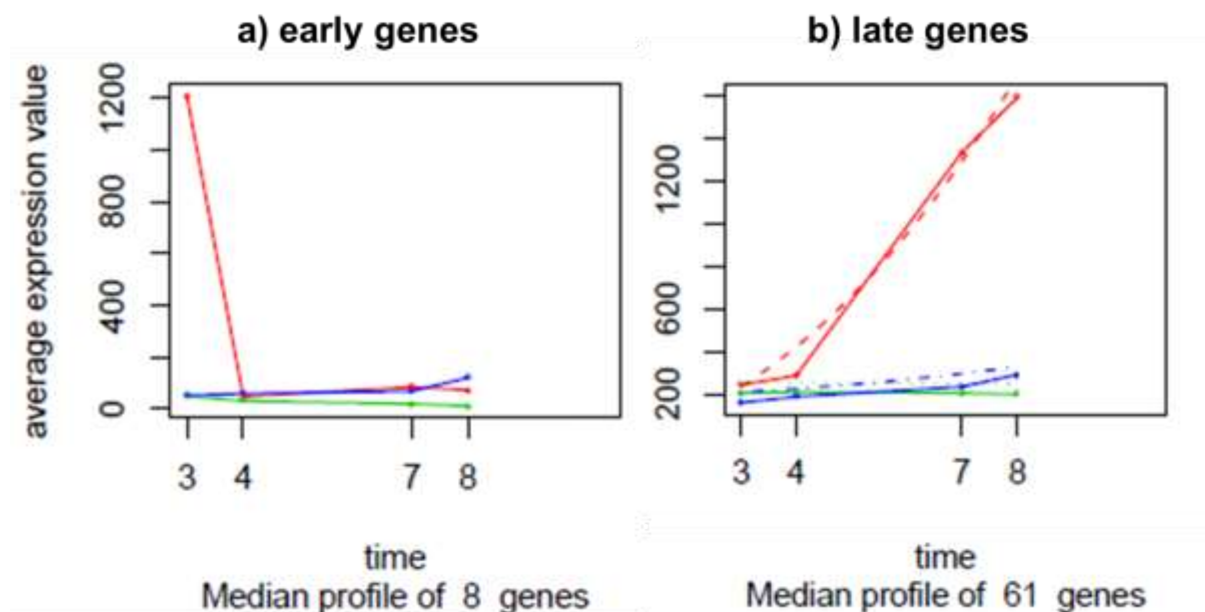

Fig. 3: The two clusters including genes that were significantly differentially expressed over time in the parasitic pseudoqueens but not in the two social control groups. Given are the average expression values (solid lines) including the regression curves (dashed lines) over time (from day 3 to day 8). a) early genes, b) late genes. The three experimental groups are displayed in different colours: Red – parasitic *A. m. capensis* pseudoqueens, green – social *A. m. capensis* worker, blue – social *A. m. scutellata* worker

In total 245 of the 471 genes were significantly differentially expressed in the pseudoqueens over time. Based on correlating expression patterns, these genes were grouped into seven clusters (shown in Online Resource 3). Two clusters included genes that were significantly differentially expressed

over time in the parasitic pseudoqueens, but not in the two control groups (*A. m. capensis* and *A. m. scutellata* social workers). These are therefore most likely involved in the transition from a social sterile worker into a parasitic pseudoqueen. One cluster includes eight early genes that are highly expressed in the first days, but down-regulated later (Fig. 3a 'early genes'). The other cluster includes 61 late genes that are up-regulated over time in the parasitic pseudoqueens, but not in the two control groups (Fig. 3b 'late genes'). Details about the genes in these two clusters are given in Online Resource 4.

Complementary analysis methods reveal similar results

Both the PCA and the time course analysis resulted in similar findings. 65% of the 20 genes with the highest scores for PC2 'time' that contribute most to the differentiation of the two late samples (days 7 and 8) of the pseudoqueen to the rest of the samples in the PCA, are also among the 'late genes' (Fig. 3b) that are significantly up-regulated over time in the pseudoqueens, but not in the two control groups, identified by the time course analysis.

Gene functions

The functional GO enrichment analysis of the 'late genes' identified by the time course analysis revealed that among others the biological processes defense response, innate immune response, sensory perception of smell and the molecular functions oxidoreductase activity, odorant binding and olfactory receptor activity were significantly enriched. This is consistent with the results of the GO enrichment analysis of the genes with the highest scores for the principal component 'time' of the PCA: Among others the same biological processes and molecular functions were significantly enriched. Details for all GO enrichment analyses are given in Online Resource 5.

In addition, the 'late genes' revealed by the time course analysis (Fig. 3b) comprise not only protein coding genes, but also a considerable number (23%) of long non-coding RNAs (lncRNAs) including *Ks-1* (Kenyon cell/small-type preferential gene-1, Sawata et al. 2002).

Discussion

We here compared the development of parasitic pseudoqueens with that of social workers at different time points. A core set of candidate genes involved in the parasitic pseudoqueen development was extracted using two different complementary analytical approaches, with both yielding similar results. As hypothesized, pseudoqueens share many features with real queens, but the analyses also revealed particular transcriptomic changes that may elucidate the specific developmental trajectories in becoming social parasites.

Genes related to ovary activation

A. m. capensis workers are characterized by rapid ovary activation (Anderson 1963; Ruttner and Hesse 1981; Moritz and Hillesheim 1985). Accordingly, genes that are putatively involved in ovary activation were up-regulated in the *A. m. capensis* pseudoqueens. One interesting candidate gene in this context is the multi-functional protein Transferrin (Trf, Kucharski and Maleszka 2003). Trf is significantly up-regulated in queens compared to workers (Grozinger et al. 2007) and it might be positively correlated to the ovary activation in honeybees (Koywiwattrakul and Sittipraneed 2009). Further, even though in comparison with the control groups not significantly differentially expressed, the parasitic pseudoqueens of this study were characterised by high expression levels of *Vitellogenin* (*Vg*) – especially at days 3 and 4. In general, the pleiotropic yolk protein Vg has been considered to be particularly important for ovary maturation in honeybees (Engels 1974; Engels and Fahrenhorst 1974; Fleig 1995; Lin et al. 1999; Seehuus et al. 2007). Queens are characterised by high Vg levels throughout their lives (Engels and Fahrenhorst 1974). Also young workers are characterised by a high Vg titer (Amdam et al. 2003), but this amount decreases with time and completely drops with the transition from in-hive tasks to foraging (Engels and Fahrenhorst 1974; Pinto et al. 2000). This temporal change is related to the age polyethism in workers (Amdam et al. 2004). Only during the nursing phase is Vg needed by the hypopharyngeal glands (HPGs) to synthesize brood food. However, under hopelessly queenless and therefore brood-decreasing conditions most Vg is no longer transported to the HPGs but to the growing ovaries (Amdam et al. 2003; Seehuus et al. 2007; Wegener et al. 2009). Consequently, based on their higher Vg levels, young workers can activate their

ovaries more effectively than older ones (Lin et al. 1999). Thus, this is in line with the 'Reproductive Ground Plan Hypothesis' (RGPH, West-Eberhard 1996; Amdam et al. 2004) that predicts that queens and young workers are characterised by common physiological traits and similar gene expression patterns, resembling the reproductive phase of the common solitary ancestor that continuously switched tasks between foraging to provision the nest and reproducing. Whereas foragers are characterised by gene expression patterns similar to those in the foraging phase of the ancestor. Accordingly, Grozinger et al. (2007) found that reproductive honeybee females - both queens and workers - primarily up-regulate genes that are typically associated with the nursing state (Whitfield et al. 2003). Dominant *A. m. capensis* workers do not engage in nursing tasks (Hillesheim et al. 1989), but rather invest in the development of reproductive traits. Therefore the synthesised Vg is most likely transported to the ovaries right from the beginning, resulting in the *capensis*-typical rapid ovary activation (Anderson 1963; Ruttner and Hesse 1981; Moritz and Hillesheim 1985).

Other similarities between pseudoqueens and queens

In addition, other similarities between pseudoqueens and regular queens were found that cannot be directly related to ovary activation (Grozinger et al. 2007). These include the up-regulated oxidoreductases and immune genes in the 7-8 day old parasitic pseudoqueens. Both oxidoreductases and immune genes improve the honeybees' resistance to endo- and exogenous harmful substances, which might also contribute to the queen's extended lifespan compared to workers (reviewed in Remolina and Hughes 2008). Oxidoreductases are involved in oxidative processes and contribute therefore to the antioxidative stress response and detoxification (Li et al. 2007). Immune genes including antimicrobial peptides (f. e. *defensin* and *apidaecin*) are part of the innate immune system and protect honeybees against infections (Danihlik et al. 2015).

Genes related to establishing reproductive dominance

Parasitic *A. m. capensis* workers face the task of developing reproductive dominance, particularly in the presence of the queen. They need to overcome two large obstacles to successfully establish themselves as the fertile dominant individual:

1) the presence of the queen and the suppressive effect of her QMP on worker ovary action (Butler 1959; Slessor et al. 1988; Plettner et al. 1993; Winston and Slessor 1998; Hoover et al. 2003)
2) the aggressions of other colony members due to their ovary activation (Anderson 1968; Crewe 1984; Calis et al. 2005)

In general, *A. m. capensis* workers can activate their ovaries in the presence of a queen (Pirk et al. 2002). One way to overcome the suppressive effect of the queen's QMP is a simple behavioural trait: moving away and staying in regions of the colony distant from the queen (Moritz et al. 2002; Pirk et al. 2002). The set of 'early genes' that is highly expressed in the first days of the parasitic pseudoqueen life, but not in the two control groups, may be particularly important to characterise this trait. Although we did not analyse brain or antennal tissue, the neuronal gene *nAChRa5* (neuronal acetylcholine receptor subunit alpha-10) which was among the 'early genes' might nevertheless be relevant in this context, as gene products of the fat body can be transported to the target tissue (e. g. Vg, Haunerland and Shirk 1995; Amdam et al. 2003). *nAchRa5* might therefore be involved in the specific behavioural transitions enabling the pseudoqueens to avoid the queen rather than being attracted by her QMP in the colony (Moritz et al. 2002; Pirk et al. 2002). In addition, the odorant binding protein encoding gene *OBP14*, revealed by the PCA, might be relevant in this context. *OBP14* is strongly down-regulated in the parasitic pseudoqueens, but up-regulated in sterile workers (Wanner et al. 2007) and queens (Wu et al. 2016). Functionally, *OBP14* might be involved in the QMP perception of workers (Wanner et al. 2007). Therefore, the down-regulation of this gene might be an important mechanism used by parasitic workers to resist the suppressive effects of QMP. In queens this might not be necessary. Because, even though they are continuously exposed to QMP, it does not seem to have any inhibiting effect on their ovary development. In queens *OBP14* might rather act as lipophilic compound carrier that might be involved in caste-specific physiological functions (Pelosi et al. 2005; Forêt and Maleszka 2006; Forêt et al. 2007; Niu et al. 2014).

In the later developmental phase, pseudoqueens have been shown to engage in lethal fights with the host queen (Moritz et al. 2003) that are elicited by the tergal gland pheromones of the queen (Pflugfelder and Koeniger 2003). The set of 'late genes' that is up-regulated over time in the

pseudoqueens, but not in the two control groups, comprises two odorant receptor genes (*odorant receptor 30a-like* and *odorant receptor 13a-like*) that might be involved in the perception of the queen's tergal gland pheromones. In addition, the 'late genes' comprise many defense genes. Among others, the gene '*venom dipeptidylpeptidase IV*' (also named '*Api m 5'*), which is a toxic component of the honeybee venom (Li et al. 2013). Cardoen et al. (2011) suggested that the up-regulation of aggression-related genes including *venom dipeptidylpeptidase IV* in older workers is a preparation for defense tasks or foraging. In contrast, in queens '*venom dipeptidylpeptidase IV*' is down-regulated in the sexually mature state (Wu et al. 2016). Laying queens normally do not need to engage in aggressive fights, whereas developing pseudoqueens are often involved in fights with the host queen (Moritz et al. 2003) or need to defend themselves against attacking host workers (Velthuis 1976; Anderson 1968; Crewe 1984; Calis et al. 2005). Therefore, a highly toxic venom and the up-regulation of defense genes in general, may be an important trait during parasitic take-overs of host colonies (Neumann and Moritz 2002). Further, the observed up-regulation of immune genes might be related to this trait. Since pseudoqueens that engage in fights might be injured, a highly active innate immune system might enhance the healing process and thus survival.

Finally, many lncRNAs were up-regulated in the parasitic pseudoqueens over time, but not in the two control groups. LncRNAs do not encode proteins, but they display important biological functions in many organisms (reviewed in Bonasio and Shiekhattar 2014) and are also highly abundant in the honeybee genome (Jayakodi et al. 2015). However, they have not yet been extensively studied. Nevertheless, the functions of a few lncRNAs have been addressed already and they seem to be involved in important physiological and behavioural traits, including immune response (Jayakodi et al. 2015; Satyavathi et al. 2017), larval ovary development (Humann and Hartfelder 2011; Humann et al. 2013) and potentially the age polyethism in workers (Tadano et al. 2009). In general, lncRNAs regulate gene expression at the transcriptional and translational level by either activating or supressing transcription or translation (Huarte 2013; Yoon et al. 2013; Bonasio and Shiekhattar 2014; Long et al. 2017). Therefore, the lncRNAs that were up-regulated in the pseudoqueens of this study might

harbour the potential to regulate the transition from a social sterile worker into a highly fecund parasitic pseudoqueen on the transcriptional and/or translational level.

Conclusion

In conclusion, our results support a combination of the 'Phenotype Deletion Model' and 'Phenotype Shift Model' for the evolution of social parasitism (Cini et al. 2015). Even though we did not observe a complete loss of the worker characteristic gene expression in the parasitic pseudoqueens, we found many similarities to the gene expression patterns of regular honeybee queens (which partly supports the 'Phenotype Deletion Model'). The parasite-specific transcriptomic changes that discriminate the parasitic pseudoqueens to both - queens and social workers – prove the 'Phenotype Shift Model'. These specific changes are likely involved in the development of the social parasitism in *A. m. capensis*. However, we want to emphasize that this study is dealing with correlative gene expression data. Only direct evidence, obtained via functional analyses (e. g. via gene knockdown experiments using RNAi), can finally prove the functional role of the identified genes in contributing to the shift from the social to the parasitic phenotype. Nevertheless, this study is a first step towards the understanding of the genetic background of the social parasitism in *A. m. capensis,* which remains an important problem in South African beekeeping (Pirk et al. 2014). Even decades after the 'Capensis Calamity', the parasite continues to destroy thousands of colonies annually.

Acknowledgements

We are grateful to the members of the Research Unit FOR2281 "Sociality and the reversal of the fecundity-longevity trade-off," especially to Daniel Elsner and Mark Harrison, for their valuable input regarding the data analysis. We also thank Nikita Venter, Cathy Bester, and Alex Nepomuceno for their assistance in handling the experimental bees.

Funding information

Financial support was granted by the Deutsche Forschungsgemeinschaft (RFAM).

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Alexa A, Rahnenfuhrer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 2: 1600–1607

Allsopp MH (1992) The Capensis Calamity. South African Bee J 64: 52-54

Allsopp MH, Crewe R (1993) The Cape honeybee as a Trojan horse rather than the hordes of Jenghiz Khan. Am Bee J 133: 121-123

Amdam GV, Fennern E, Havukainen H (2012) Vitellogenin in Honey Bee Behavior and Lifespan. In: Galizia C, Eisenhardt D, Giurfa M (eds) Honeybee Neurobiology and Behavior. Springer, Dordrecht

Amdam GV, Norberg K, Fondrik MK, Page RE (2004) Reproductive ground plan may mediate colony-level selection effects on individual foraging behaviour in honey bees. PNAS 101: 11350-11355

Amdam GV, Norberg KA, Hagen A, Omholt SW (2003) Social exploitation of vitellogenin. Proc Natl Acad Sci USA 100: 1799–1802

Anders S, Pyl PT, Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166-169

Anderson RH (1963) The laying worker in the Cape honeybee, *Apis mellifera capensis*. J Apic Res 2: 85-92

Anderson RH (1968) The effect of queen loss on colonies of *Apis mellifera capensis*. S Afri J Agri Sci 11: 383-388

Aufauvre J, Misme-Acouturier B, Viguès B, Texier C, Delbac F, Blot N (2014) Transcriptome Analyses of the Honeybee Response to *Nosema ceranae* and Insecticides. PLoS ONE 9: e91686

Aumer D, Allsopp MH, Lattorf MHG, Moritz RFA, Jarosch-Perlow A (2017) Thelytoky in Cape honeybees (*Apis mellifera capensis*) is controlled by a single recessive locus. Apidologie 48: 401-410

Barchuk AR, Cristino AS, Kucharski R, Costa LF, Simoes ZLP, Maleszka R (2007): Molecular determinants of caste differentiation in the highly eusocial honeybee *Apis mellifera.* BMC Dev Biol 7: 70

Baudry E, Kryger P, Allsopp M, Koeniger N, Vautrin D. Mougel F, Cornuet J-M, Solignac M (2004) Whole-genome scan in thelytokous-laying workers of the Cape honey bee (Apis mellifera capensis): central fusion, reduced recombination rates and centromere mapping using half-tetrad analysis. Genetics 167: 243-252

Beekman M, Calis JNM, Boot WJ (2000) Parasitic honeybees get royal treatment. Nature 404, 723

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics: btu170

Bonasio R, Shiekhattar R (2014) Regulation of Transcription by Long Noncoding RNAs. Annu Rev Genet 48: 433-455

Butler CG (1959) The source of substance produced by a queen honeybee which inhibits development of the ovaries of the workers of her colony. Proc Roy Ent Soc A 34: 137-138

Calis JNM, Allsopp MH, Boot WJ (2005) Establishment of Cape honey bees as social parasites in African honeybee colonies. Proc Neth Entomol Soc Meet 16: 81-90

Cardoen D, Wenseleers T, Ernst UR, Danneels EL, Laget D, de Graaf DC, Schoofs L, Verleyen P (2011) Genome-wide analysis of alternative reproductive phenotypes in honeybee workers. Mol Ecol 20: 4070-4084

Carreck NL, Andree M, Brent CS, Cox-Foster D, Dade HA, Ellis JD, Hatjina F, van Englesdorp D (2013) Standard methods for Apis mellifera anatomy and dissection. J Api Res 52: 1-40

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidnium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162: 156-159

Cini A, Patalano S, Segonds-Pichon A, Busby GB, Cervo R, Sumner S (2015) Social parasitism and the molecular basis of phenotypic evolution. Front Genet 6: 32

Corona M, Velarde RA, Remolina S, Moran-Lauter A, Wang Y, Hughes KA, Robinson GE (2007) Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity. Proc Natl Acad Sci USA 104: 7128–7133

Crewe RM (1984) Differences in behaviour and morphology between capensis and adansonii. S Afr Bee J 5: 16-21

Crewe RM, Velthuis HHW (1980) False queens: A consequence of mandibular gland signals in worker honeybees. Naturwissenschaften 67: 467-469

Danihlik J, Aronstein K, Petrivalsky M (2015) Antimicrobial peptides: A key component of honey bee innate immunity. J Apic Res 54: 123-136

Doublet V, Paxton RJ, McDonnel CM, Dubois E, Nidelet S, Moritz RFA, Alaux C, Le Conte Y (2016) Brain transcriptomes of honey bees (*Apis mellifera*) experimentally infected by two pathogens: Black queen cell virus and *Nosema ceranae.* Genomics Data 10: 79-82

Dunn OJ (1964) Multiple comparison using RANK sums. Technometrics 6: 241-252

Dzierzon J (1845) Chodowanie pszczół - Sztuka zrobienica złota, nawet z zielska, in: Tygodnik Polski Poświęcony Włościanom (pl), Issue 20, Pszczyna (Pless)

Elsik CG, Worley KC, Bennett AK, Beye M, Camara F, Childers CP, et al. (2014) Finding the missing honey bee genes: lessons learned from a genome upgrade. BMC genomics 15: 86

Emery C (1909) Über den Ursprung der dulotischen, parasitischen und myrmekophilen Ameisen. Biol Zentralbl 29: 352–362

Engels W (1974) Occurrence and significance of vitellogenins in female castes of social Hymenoptera. Am Zool 14: 1229-1237

Engels W, Fahrenhorst H (1974) Alters- und kastenspezifische Veränderungen der Haemolymph-Protein-Spektren bei *Apis mellificia.* Roux's Arch 174: 285–296

Fleig R (1995) Role of the follicle cells for yolk uptake in ovarian follicles of the honey-bee *Apis-mellifera* L (Hymenoptera, Apidae). Int J insect morphol embryol 24: 427–433

Fraley C, Raftery AE, Murphy TB, Scrucca L (2012) mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. Technical Report No. 597, Department of Statistics, University of Washington

Grozinger CM, Fan Y, Hoover SER, Winston ML (2007) Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). Mol Ecol 16: 4837-4848

Haunerland NH, Shirk PD (1995) Regional and functional differentiation in the insect fat body. Annu Rev Entomol 40: 121–145

Härtel S, Wossler T, Moltzer G-J, Crewe R, Moritz RFA, Neumann P (2011) Pheromone-mediated reproductive dominance hierarchies among pseudo-clonal honeybee workers (*Apis mellifera capensis*). Apidologie 42: 659-668

Hemmling C, Koeniger N, Ruttner F (1979) Quantitative Bestimmung der 9-Oxodecensäure im Lebenszyklus der Kapbiene (*Apis mellifera capensis* Escholtz). Apidologie 10: 227-240

Hepburn HR, Crewe RM (1991) Portrait of the Cape honeybee, *Apis mellifera capensis*, Apidologie 22: 567-580

Hepburn HR, Magnuson P, Herbert L, Whiffler LA (1991) The development of laying workers in field colonies of the Cape honey bee. J Api Res 30: 107-112

Hess G (1942) Ueber den Einfluss der Weisellosigkeit und des Fruchtbarkeitsvitamins E auf die Ovarien der Bienenarbeiterin (Ein Beitrag zur Frage der Regulationen im Bienenstaat). Beih Schweiz Bienen Ztg 2: 33–111

Hillesheim E, Koeniger N, Moritz RFA (1989) Colony performance in honeybees (*Apis mellifera capensis* Esch.) depends on the proportion of subordinate and dominant workers. Behav Ecol Socioboil 24: 291-296

Hoover SER, Keeling CI, Winston ML, Slessor KN (2003) The effect of queen pheromones on worker honey bee ovary development. Naturwissenschaften 90: 477-480

Huarte M (2013) LncRNAs have a say in protein translation. Cell Res 23: 449-451

Humann FC, Hartfelder K (2011) Representational Difference Analysis (RDA) reveals differential expression of conserved as well as novel genes during castespecific development of the honey bee (Apis mellifera L.) ovary. Insect Biochem Mol Biol 41: 602–612

Humann FC, Tiberio GJ, Hartfelder K (2013) Sequence and Expression Characteristics of Long Noncoding RNAs in Honey bee Caste Development – Potential Novel Regulators for Transgressive Ovary Size. PLoS ONE 8: e78915

Jayakodi M, Jung JW, Park D, Ahn Y-J, Lee S-C, Shin S-Y, Shin C, Yang T-J, Kwon H-W (2015) Genome-wide characterization long intergenic non-coding RNAs (lincRNAs) provides new insight into viral diseases in honey bees *Apis cerana* and *Apis mellifera.* BMC Genomics 16: 680-692

Johannsmeier MF (1983) Experiences with the Cape bee in the Transvaal. S Afr Bee J 55: 130-138

Johannsmeier MF (1992) The Cape bee: a problem in the Transvaal. Plant Protection Research Institute. Pretoria, Leaflet

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, and Hunter S (2014) InterProScan 5: genome-scale protein function classification. Bioinformatics 30: 1236-1240

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14: R36

Koeniger G, Würkner W (1992) Die Kap Honigbiene (*Apis mellifera capensis)* Natürliche Verbreitung und die Schwierigkeit der Haltung unter unseren klimatischen Bedingungen. Die Biene 128: 583-588

Köhler A, Nicolson SW, Pirk CWW (2013) A new design for honey bee hoarding cages for laboratory experiments. J Api Res 52: 12-14

Koywiwattrakul P, Sittipraneed S (2009) Expression of Vitellogenin and Transferrin in Activated Ovaries of Worker Honey Bees, *Apis mellifera*. Biochem Genet 4: 19-26

Kucharski R, Malezska R (2003) Transcriptional profiling reveals multifunctional roles for transferrin in the honeybee *Apis mellifera.* J Insect Sci 3: 27-34

Langmead B, Salberg SL (2012) Fast gapped read-alignment with Bowtie 2. Nat Meth 9: 357-359

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25

Lattorff HMG, Moritz RFA, Crewe RM, Solignac M (2007) Control of reproductive dominance by the thelytoky gene in honeybees. Biol Lett 3: 292-295

Lattorff HMG, Moritz RFA, Fuchs S (2005) A single locus determines thelytokous parthenogenesis of laying honeybee workers (*Apis mellifera capensis*). Heredity 94: 533-537

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R and 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics 25: 2078-2079

Li X, Schuler MA, Berenbaum MR (2007) Molecular Mechanisms of Metabolic Resistance to Synthetic and Natural Xenobiotics. Annu. Rev. Entomol 52: 231-253

Lin H, Winston ML, Haunerland NH, Slessor KN (1999) Influence of age and population size on ovarian development and of trophallaxis on ovarian development and vitellogenin titres of queenless worker honeybee (Hymenoptera: Apidae). Can Entomol 131: 695–706

Li R, Zhang L, Fang Y, Han B, Lu X, Zhou T, Feng M, Li J (2013) Proteome and phosphoproteome analysis of honeybee (Apis mellifera) venom collected from electrical stimulation and manual extraction of the venom gland. BMC Genomics 14: 766-778

Long Y, Wang X, Youmans DT, Chech TR (2017) How do lncRNAs regulate transcription. Sci Adv 3: eaao2110

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15: 550

Lyko F, Foret S, Kucharski R, Wolf S, Falckenhayn C, Maleszka R (2010) The honey bee epigenomes: Differential methylation of brain DNA in queens and workers. PLOS Biol 8: e1000506

Martin SJ, Beekman M, Wossler TC, Ratnieks FLW (2002a) Parasitic Cape honeybee workers, *Apis mellifera capensis*, evade policing. Nature 415: 163-165

Martin S, Wossler T, Kryper P (2002b): Usurpation of African *Apis mellifera scutellata* colonies by parasitic *Apis mellifera capensis* workers. Apidologie 33: 215-232

Moritz RFA, Crewe RM, Hepburn HR (2002) Queen avoidance and mandibular gland secretion of honeybee workers (*Apis mellifera* L.). Insect Soc 49: 86-91

Moritz, RFA, Hillesheim E (1985) Inheritance of dominance in honeybees (*Apis mellifera capensis*). Behav Ecol Sociobiol 17: 87-89

Moritz RFA, Kryger P, Allsopp MH (1996) Competition for royalty in bees. Nature 384: 31

Moritz RFA, Kryger P, Allsopp MH (1999) Lack of worker policing in the Cape honeybee (*Apis mellifera capensis*). Behaviour 136: 1079-1092

Moritz RFA, Lattorff HMG, Crewe RM (2004) Honeybee workers (*Apis mellifera capensis*) compete for producing queen-like pheromone signals. Proc Roy Soc Lond B 271: 98-100

Moritz RF, Pflugfelder J, Crewe RM (2003) Lethal fighting between honeybee queens and parasitic workers (*Apis mellifera*). Naturwissenschaften 9: 378-381

Moritz RFA, Simon UE, Crewe RM (2000) Pheromonal contest between honeybee workers (*Apis mellifera capensis*). Naturwissenschaften 87: 395-397

Neumann P, Hepburn R (2002) Behavioural basis for social parasitism of Cape honeybees (*Apis mellifera capensis*). Apidologie 33: 165-192

Neumann P, Moritz RFA (2002) The Cape honeybee phenomenon: the sympatric evolution of a social parasite in real time? Behav Ecol Sociobiol 52: 271-281

Niu D, Zheng H, Corona M, Lu Y, Chen X, Cao L, Sohr A, Hu F (2014) Transcriptome comparison between inactivated and activated ovaries of the honey bee *Apis mellifera* L. Insect Mol Biol 23: 668-681

Nueda MJ, Tarazona S, Conesa A (2014) Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. Bioinformatics 30: 2598-2602

Okosun OO, Yusuf AA, Crewe RM, Pirk CW (2015) Effects of age and Reproductive Status on Tergal Gland Secretions in Queenless Honey bee Workers, Apis mellifera scutellata and A. m. capensis. J Chem Ecol 41: 896-903

Onions GW (1912) South African "fertile-worker bees". S Afr Agric J 1: 720-728

Onions GW (1914) South African "fertile worker bees". Union S Afr Agric J 7:44-46

Pflugfelder J, Koeniger N (2003) Fight between virgin queens (Apis mellifera) is initiated by contact to the dorsal abdominal surface. Apidologie 34: 249–256

Phiancharoen M, Pirk CWW, Radloff SE, Hepburn R (2010) Clinal nature of the frequencies of ovarioles and spermathecae in Cape worker honeybees, *Apis mellifera capensis*. Apidologie 41: 129-134

Pinto LZ, Bitondi MMG, Simões ZLP (2000) Inhibition of vitellogenin synthesis in *Apis mellifera* workers by juvenile hormone analogue, pyriproxyfen. J Insect Physiol 46: 153-160

Pirk CWW, Human H, Crewe RM, vanEngelsdorp D (2014) A survey of managed honey bee colony losses in the Repulic of South Africa – 2009 to 2011. J Api Res 53: 35-42

Pirk CWW, Neumann P & Hepburn HR (2002) Egg laying and egg removal by workers are positively correlated in queenright Cape honeybee colonies (Apis mellifera capensis) Apidologie 33: 203-211

Plettner E, Slessor K, Winston M, Robinson G, Page R (1993) Mandibular gland components and ovarian development as measures of caste differentiation in the honey bee (*Apis mellifera l.*). J Insect Physiol 39: 235-240

Remolina SC, Hughes KA (2008) Evolution and mechanisms of long life and high fertility in queen honey bees. AGE 30: 177-185

Ruttner F. (1988) Biogeography and taxonomy of honeybees, Springer-Verlag, Berlin

Ruttner F, Hesse B (1981) Rassenspezifische Unterschiede in Ovarentwicklung und Eiablage von weisellosen Arbeiterinnen der Honigbiene *Apis mellifera l*. Apidologie 12: 159-183

Sakagami SF (1958) The false queen: fourth adjustive response in dequeened colonies, Behaviour 8: 280–296

Sawata M, Yoshino D, Takeuchi H, Kamikouchi A, Ohashi K, Kubo T (2002) Identification and punctate nuclear localization of a novel noncoding RNA, Ks-1, from the honeybee brain. RNA 8: 772-785

Seehuus S-C, Norberg K, Krekling T, Fondrk K, Amdam GV (2007) Immunogold Localization of Vittellogenin in the Ovaries, Hypopharyngeal Glands and head Fat Bodies of Honeybee Workers, *Apis Mellifera.* J Insect Sci 7: 52

Simon U, Moritz RFA, Crewe RM (2001) The ontogenetic pattern of mandibular gland components in queenless worker bees (*Apis mellifera capensis* Esch.). J Insect Physio 47: 735-738

Slessor KN, Kaminski L-A, King GGS, Borden JH, Winston ML (1988) Semiochemical basis of the retinue response to queen honey bees. Nature 332: 354-356

Smith CR, Helms Cahan S, Kemena C, Brady SG, Yang W, Bornberg-Bauer E,Eriksson T, Gadau J, Helmkampf M, Gotzek D, Okamoto Miyakawa M., Suarez AV, Mikheyev A (2015) How do genomes create novel phenotypes? Insights from the loss of the worker caste in ant social parasites. Mol Biol Evol 32: 2919-2931

Tadano H, Yamazaki Y, Takeuchi H, Kubo T (2009) Age- and division-of-labour-dependent differential expression of a novel non-coding RNA, Nb-1, in the brain of worker honeybees, Apis mellifera L. Insect Mol Biol 18: 715-726

Tribe GD (1983) What is the Cape bee? S Afr Bee J 55:77–87

Velthuis HHW (1976) Egg laying, aggression and dominance in bees. Proc 15th Int Congr Entomol, Washington: 436-449

Velthuis HHW, Ruttner F, Crewe RM (1990) Differentiation in reproductive physiology and behaviour during the development of laying worker honeybees. In: Engels W (Ed.) Social Insects. Springer Verlag, Berlin Heidelberg, New York: 231-243

Wanner KW, Nicols AS, Walden KKO, Brockmann A, Luetje CW, Robertson HM (2007) A honey bee odorant receptor for the queen substance 9-oxo-2-decenoic acid. PNAS 104: 14383-14388

Wegener J, Huang ZY, Lorenz MW, Bienefeld K (2009) Regulation of hypopharyngeal gland activity and oogenesis in honey bee (Apis mellifera) worker. J Insec Physiol 55: 716-725

Weiner SA, Toth, AL (2012) Epigenetics in social insects: a new direction for understanding the evolution of castes. Genet Res Int: Article ID 609810

West-Eberhard MJ (1996) Wasp societies as microcosms for the study of development and evolution. In: Turillazzi S, West-Eberhard MJ, editors. Natural history and evolution of paper-wasps. Oxford (UK): Oxford University Press: 90–317

Wilson EO (1971) The Insect Societies. Harvard Bellknap, Cambridge

Winston ML (1987) The Biology of the Honey Bee. Cambridge, Massachusetts: Harvard University Press, London

Winston ML, Slessor KN (1998) Honey bee primer pheromones and colony organization: gaps in our knowledge. Apidologie 29: 81-95

Whitfield CW, Cziko AM, Robinson GE (2003) Gene expression profiles in the brain predict behaviour in individual honey bees. Science 302: 296–299

Woyke J (1995) Invasion of Capensis bee. In: Magnuson P (Ed.) Proceedings of the First International Electronic Conference on the Cape Bee Problem in South Africa, 5-30 June 1995, PPRI, Pretoria: 35

Wu X-B, Wang Z-L, Li S-Y, Zeng Z-J (2016) Transcriptome comparison between newly emerged and sexualy matured bees of Apis mellifera. J Asia Pac Entomol 19: 893-897

Wu Y, Zheng H, Corona M, Pirk C, Meng F, Zheng Y, Hu F (2017) Comparative transcriptome analysis on the synthesis pathway of honey bee (*Apis mellifera*) mandibular gland secretions. Sci Rep 7: 4530

Yoon J-H, Abdelmohsen K, Gorospe M (2013) Post-transcriptional gene regulation by long noncoding RNA. J Mol Biol 425: 3723-3730