Heredity (2011) 106, 894–903

www.nature.com/hdy

ORIGINAL ARTICLE

Genetic architecture of ovary size and asymmetry in European honeybee workers

O Rueppell[1], JD Metheny[1,4], T Linksvayer[2,5], MK Fondrk[2,6], RE Page Jr[2] and GV Amdam[2,3]
[1]*Department of Biology, University of North Carolina at Greensboro, Greensboro, NC, USA;* [2]*School of Life Sciences, Arizona State University, Tempe, AZ, USA and* [3]*Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Aas, Norway*

The molecular basis of complex traits is increasingly understood but a remaining challenge is to identify their co-regulation and inter-dependence. Pollen hoarding *(pln)* in honeybees is a complex trait associated with a well-characterized suite of linked behavioral and physiological traits. In European honeybee stocks bidirectionally selected for *pln*, worker (sterile helper) ovary size is pleiotropically affected by quantitative trait loci that were initially identified for their effect on foraging behavior. To gain a better understanding of the genetic architecture of worker ovary size in this model system, we analyzed a series of crosses between the selected strains. The crossing results were heterogeneous and suggested non-additive effects. Three significant and three suggestive quantitative trait loci of relatively large effect sizes were found in two reciprocal backcrosses. These loci are not located in genome regions of known effects on foraging behavior but contain several interesting candidate genes that may specifically affect worker-ovary size. Thus, the genetic architecture of this life history syndrome may be comprised of pleiotropic, central regulators that influence several linked traits and other genetic factors that may be downstream and trait specific.
Heredity (2011) **106,** 894–903; doi:10.1038/hdy.2010.138; published online 3 November 2010

Keywords: worker caste; social evolution; Apis mellifera; QTL; trait network; ovarioles

Introduction

Organisms embody numerous biological traits that are influenced by interactive gene networks. Therefore, most traits are mutually dependent through either direct, specific co-regulation or developmental, physiological or life-history constraints. These traits cannot be satisfactorily understood if studied in isolation because they may constrain or enable each other's evolution (Steppan *et al*., 2002; Sih *et al*., 2004). Interconnectivity is particularly important for complex traits, with potentially many genes, environmental factors and their interactions involved (Phillips and Belknap, 2002).

Life-history and behavioral traits are exemplary complex traits. Despite considerable research, their genetic analysis is in most cases still in its initial stages. Recently however, some molecular factors of significant importance have been identified for a few complex traits (Robinson *et al*., 2008). For example, the ectodysplasin signaling pathway influences the armor of sticklebacks (Colosimo *et al*., 2005), natural variation in *Caenorhabditis elegans* foraging is mediated by allelic variation in neuropeptide receptor *npr-1* (de Bono and Bargmann, 1998), and natural variation in the cyclic GMP-dependent protein kinase (PKG) affects learning in *Drosophila* (Mery *et al*., 2007). Additionally, the *PKG* gene has been associated with feeding (Kaun *et al*., 2007) and natural variation in foraging behavior in several taxa (Debelle *et al*., 1989; Ben-Shahar *et al*., 2002; Lucas and Sokolowski, 2009). The pleiotropy of PKG exemplifies that comparative, in-depth studies may reveal suites of traits that are affected by single genes. These genes might belong to central molecular networks that are responsible for the evolution of syndromes (Sih *et al*., 2004) across species.

One of the best-studied natural syndromes is the pollen hoarding (*pln)* syndrome in honeybees (*Apis mellifera* L) that consists of a broad suite of correlated behavioral and physiological traits, including foraging behavior, behavioral ontogeny, gustatory and visual responsiveness, associative learning, locomotor activity, vitellogenin and juvenile hormone dynamics, and ovary size in essentially sterile helper females called workers (Page and Amdam, 2007; Page *et al*., 2007). Most insights into the *pln* syndrome have come from two honeybee strains that have been bidirectionally selected for *pln* (high versus low amounts of pollen stored in the colony) and showed correlated changes in worker foraging behavior (Page and Fondrk, 1995), behavioral ontogeny (Pankiw and Page, 2001), sucrose responsiveness and learning (Scheiner *et al*., 2001), locomotion (Humphries

Correspondence: Dr O Rueppell, Department of Biology, University of North Carolina at Greensboro, 312 Eberhart Building, 1000 Spring Garden Street, Greensboro, NC 27403, USA.
E-mail: olav_rueppell@uncg.edu
[4]*Current address: West Virginia University, Biochemistry Department, 1 Medical Center Drive, Morgantown, WV 26506, USA.*
[5]*Current address: Centre for Social Evolution, Department of Biology, University of Copenhagen, Universitetsparken 15, Copenhagen DK-2100, Denmark.*
[6]*Current address: Department of Entomology, University of California at Davis, Davis, CA 95616, USA.*
Received 9 June 2010; revised 14 September 2010; accepted 27 September 2010; published online 3 November 2010

et al., 2005) and light responsiveness (Tsuruda and Page, 2009). Many of these trait correlations have been reconfirmed at the individual level in wild type worker bees or by comparing Africanized and European honeybee populations in North America (Pankiw and Page, 2000; Pankiw, 2003; Amdam *et al*., 2006a). These phenotypic associations between worker reproductive traits (levels of the yolk precursor vitellogenin and ovary size) and the *pln* syndrome have lead to the reproductive ground plan hypothesis, a framework developed from the ovarian ground plan hypothesis (West-Eberhard, 1987, 1996). It proposes that reproductive control modules have been co-opted by social evolution to govern behavioral specialization among honeybee workers (Amdam *et al*., 2004, 2006a, b; Page *et al*., 2007).

Genetic support for the reproductive ground plan hypothesis has been provided by quantitative trait loci (QTL) studies of the *pln* syndrome. QTL mapping is a direct, genome-wide approach to search for causal segregating genetic variation of complex genetic traits (Mackay, 2001). It is particularly powerful in the highly recombining honeybee genome because the resulting QTL regions correspond to short physical stretches of DNA (Hunt *et al*., 2007). Initial studies of two central aspects of the *pln* syndrome, colony-level *pln* and individual foraging preference in workers, described three QTL that were labeled *pln* QTL (Hunt *et al*., 1995; Page *et al*., 2000). A later study confirmed direct and interaction effects of these *pln* QTL and revealed a fourth *pln* QTL (Rüppell *et al*., 2004). Additionally, crosses between the high- and low-*pln* strains identified one additional QTL for sucrose responsiveness (Rueppell *et al*., 2006) and three new QTL (*aff)* for the rate of behavioral ontogeny, measured as the age of first foraging (Rueppell *et al*., 2004; Rueppell, 2009). Sucrose responsiveness and the age of first foraging were also significantly influenced by the previously mapped *pln* QTL (Rueppell *et al*., 2004, 2006), suggesting pleiotropy between the different behavioral aspects of the *pln* syndrome. As predicted by the reproductive ground plan hypothesis, allelic variation at two behavioral *pln* QTL (*pln2* and *pln3*) also affected the reproductive trait ovary size, measured as the number of ovariole filaments, in worker honeybees (Wang *et al*., 2009).

The analysis of the *pln* QTL has suggested that the insulin/insulin-like signaling (IIS) pathway has a central role in the *pln* syndrome because genes involved in IIS were significantly overrepresented in the QTL regions (Hunt *et al*., 2007). Consistent gene-expression differences in *HR46* and *PDK1*, two IIS associated candidate genes located in *pln2* and *pln3*, respectively, suggest that these genes are involved in pollen strain divergence and ovary-size differences (Wang *et al*., 2009). Furthermore, RNAi-mediated gene knockdown of the insulin receptor substrate (*IRS*) gene, a candidate for *pln4*, leads to shifts in foraging preference (Wang *et al*., 2010). IIS mediates nutritional signals, and thus regulates growth, reproduction and lifespan in many organisms, including insects (Wu and Brown, 2006). Therefore, it is a prime candidate for a central molecular mechanism that coordinates multiple behavioral changes. In many insects, IIS influences the synthesis of vitellogenin, a yolk precursor protein that has been co-opted during honeybee social evolution (Amdam *et al*., 2003) with far-reaching hormonal consequences (Guidugli *et al*., 2005) that could explain the connection between worker social behavior and ovary size (Amdam *et al*., 2007).

Ovary size is not only an important aspect of the *pln* syndrome, but it is generally important for understanding insect life history evolution (Orgogozo *et al*., 2006; Bergland *et al*., 2008) and the evolution of reproductive division of labor in social insects (Lattorff *et al*., 2007; Oxley *et al*., 2008; Linksvayer *et al*., 2009b). In honeybees, the queen caste shows much larger ovaries with 150–180 ovarioles per ovary than the worker caste with usually 2–12 ovarioles per ovary (Winston, 1987). The number of ovarioles per ovary varies extensively between honeybee workers (Ruttner and Hesse, 1981) and substantial genetic variation has been reported from natural populations (Thuller *et al*., 1996; Linksvayer *et al*., 2009b). Some of the molecular and cellular mechanisms that determine worker ovary size have been identified (Schmidt Capella and Hartfelder, 1998, 2002). However, little is known about the genetics of the natural variation of ovary size in populations, such as the difference between the high- and low-*pln* strains. High strain bees have on average almost 2.5–3 ovarioles per ovary more than low-strain bees (Amdam *et al*., 2006a; Wang *et al*., 2009). However, the effect size of the *pln* QTL on ovary size is modest (Wang *et al*., 2009) and only 5.4% of the total phenotypic variance is explained by the two significant QTL. Thus, other direct genetic factors likely contribute to these ovary size differences.

Here, we report on a comprehensive QTL mapping experiment on four interrelated traits of ovarian anatomy of honeybee workers using reciprocal backcrosses that were derived from a series of crosses between the high- and low-*pln* strains and were initially analyzed for specific effects of the *pln* QTL on ovary size (Wang *et al*., 2009). After a preliminary analysis of selective DNA pools (see Supplementary Information), we conducted QTL mapping analyses in two reciprocal backcrosses based on individual genotypes. The results indicate three significant and three suggestive new QTL for worker ovary size. Our combined use of microsatellite and single-nucleotide polymorphism (SNP) markers allowed for an immediate localization of the QTL (c.f. Rueppell, 2009) to compare the genetic architecture between traits and backcrosses and to identify candidate genes for future studies.

Materials and methods

Experimental crosses

We used the well-established high- and low-*pln* strains (Page and Fondrk, 1995) to set up a series of hybrid and reciprocal high- and low-backcrosses (HBC and LBC, respectively) between them (Wang *et al*., 2009). Both parental colonies and all eight hybrid, eight HBC, and seven LBC colonies were screened for worker ovary size by counting the ovariole number in one, randomly chosen ovary in 20 workers per colony (dissection method described below).

Queens of one selected HBC and one selected LBC colony were simultaneously confined on empty comb to maximize their egg laying. Pairs of the egg-filled combs were transferred into common, unrelated rearing hives. The resulting HBC and LBC workers were anaesthetized on ice and dissected in random order within 24 h of their

emergence. The head and thorax of each bee were frozen until DNA was isolated (Wang *et al*., 2009). DNA quantity and quality were examined on a Nanodrop spectrophotometer and the DNA samples were diluted to 100 ng μl^{-1} or re-purified if a contamination was indicated. The abdomen of each bee was dissected immediately to count ovarioles. It was pinned into a dissection tray with one needle through the opening to the petiolus and the other needle through the sting chamber. The cuticle was cut on both sides and across the second anterior segment. Both ovaries were exposed and transferred onto a microscope slide to count the number of ovarioles.

The smaller ovariole number was recorded as minimum ovariole number, the number of the larger ovary as maximum ovariole number. These two variables were averaged to compute the mean ovariole number. Ovary asymmetry was computed as the difference between maximum and minimum ovariole number divided by their sum (Palmer and Strobeck, 1986), assigning individuals with zero ovarioles on both sides an asymmetry score of zero. Exclusion of the individuals with zero ovarioles did not significantly alter the results. Other asymmetry metrics, such as the difference or the ratio between maximum and ovariole number were also evaluated, but results did not differ significantly from the reported results. Individuals with missing or incomplete ovary information were omitted from the study. Individual data departed significantly from normality and thus non-parametric tests and descriptive statistics were chosen where appropriate.

QTL analyses

Based on preliminary data from 1136 SNPs (Whitfield *et al*., 2006) that were genotyped in a selective, pooled DNA QTL mapping analysis (see Supplementary Information), 280 SNPs across the genome were chosen for genotyping a random subsample of 160 individuals from each backcross. Genotyping was performed by MALDI-TOF (Matrix-assisted laser desorption/ionization-time-of-flight) mass spectrometry (Sequenom, San Diego, CA, USA) (Ragoussis *et al*., 2006) with automated genotype calling, according to Sequenom standards. Non-polymorphic loci were omitted from the analyses. Based on the Amel4.0 genomic location of the remaining SNPs and the most complete linkage map for the honeybee (Solignac *et al*., 2007) we determined significant gaps (>30 cM) in the SNP marker coverage of the genome. These gaps were filled by genotyping 96 individuals at polymorphic microsatellite loci in these genomic regions. Microsatellite loci were selected from existing ones (Solignac *et al*., 2007) or directly from the genome sequence by searching for dinucleotide repeat motifs and designing primers with Primer3 (Rozen and Skaletsky, 2000).

We used a tailed-primer approach (Schuelke, 2000), with IRD-labeling for detection on LiCor's 4300 DNA Analyzer (Lincoln, NE, USA). Alleles were amplified with a touchdown PCR protocol, decreasing the annealing temperature from 68 to 48 °C (Schug *et al*., 2004). PCR reactions were carried out in 10 µl and contained 10 ng of template DNA, 200 µM dNTPs, 120 nM forward primer, 360 nM reverse primer, 50 nM of IRD-labeled M13 primer, 2 mM $MgCl_2$, standard PCR buffer and 0.2 U of Taq polymerase. PCR products of different size and label were combined and analyzed on 25 cm gels with 1000 V for 2–3 h. Genotypes were scored in duplicate.

SNP and microsatellite data were combined and used to construct a genomic map with Mapmaker 3.0b (Lander and Botstein, 1989; Lincoln *et al*., 1993) as basis for QTL mapping. This map was compared with existing estimates (Solignac *et al*., 2007) for differences in marker ordering and recombination distance between adjacent markers. Markers that caused significant departure (more that 5% and 5 cM change in map distance) from previous estimates were double-checked for genotyping errors and where significant differences persisted, the existing, high density map (Solignac *et al*., 2007) and our best map model were both used for QTL mapping.

Using the computer program MapQTL 4.0 (van Ooijen *et al*., 2002) markers were evaluated by single marker analysis (Kruskal–Wallis tests) before interval mapping (interval size 5 cM) was performed. Markers near putative QTL with a logarithm (base 10) of odds score (LOD) > 2.0 were selected as co-factors for multiple QTL mapping (MQM). Genome-wide LOD significance thresholds for each trait were empirically determined by permutation tests (Churchill and Doerge, 1994). Pair-wise epistasis between all identified, significant and suggestive QTL in both backcrosses was tested by analysis of variance using the nearest genetic marker as factors. Significance thresholds were Bonferroni-corrected to account for the multiple testing. Higher-order interactions could not be evaluated in a meaningful way because of our limited sample size.

Results

Ovary phenotypes

The workers of the parental high-*pln* (high) strain colony had significantly more ovarioles (median:4.5 (quartiles: 3.0–6.75) per ovary) than workers of the low strain parental colony (2.0 (1.0–3.0); Mann–Whitney $Z_{(20,20)} = -4.47$, $P < 0.001$). Overall, the hybrid colonies (2.0 (1.0–3.0)) were closer to the low strain (Figure 1) but significant

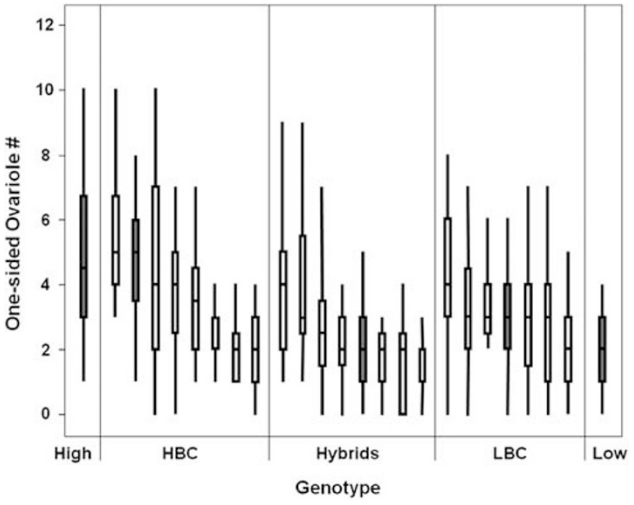

Figure 1 One-sided ovary size distributions in the parental 'high' and 'low' strain colonies, the resulting hybrids and backcross populations. Medians, interquartile ranges and total ranges are shown, based on sampling one ovary from 20 workers per colony. Darkened boxes highlight the colonies used in generating the mapping populations.

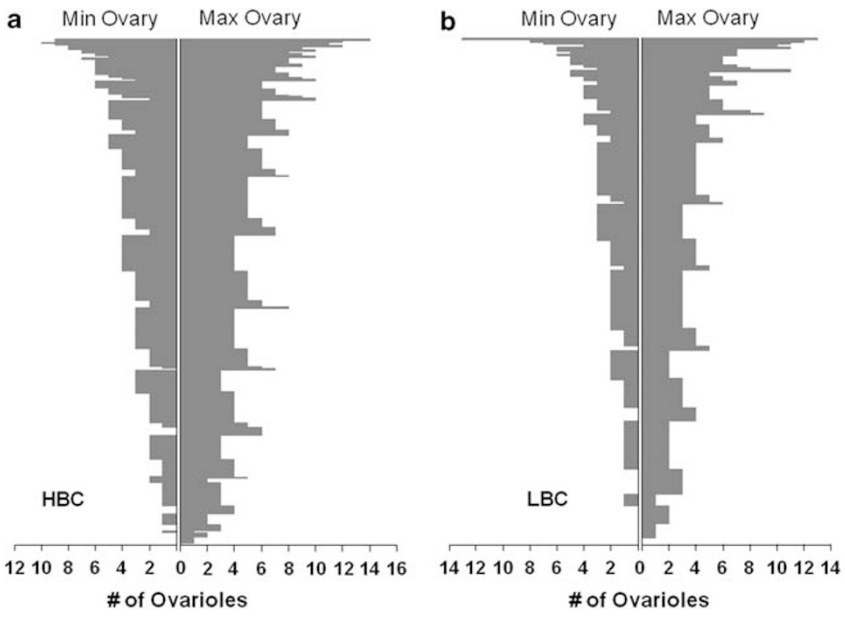

Figure 2 Phenotypic distributions of worker ovary size in the two mapping populations HBC (**a**) and LBC (**b**). Individual workers are represented as horizontal lines, sorted in descending order of mean ovary size. The left 'Min Ovary' displays the ovariole number of the smaller ovary side, and the corresponding line on the right ('Max Ovary') displays the number of ovarioles in the larger-side ovary.

heterogeneity among the individual hybrid populations existed (Kruskal–Wallis $\chi^2 = 36.4$, df = 7, $P < 0.001$). The hybrid colony (2.0 (1.0–3.0)) that was chosen as queen source for subsequent backcrosses was not significantly different from the low strain parent ($Z_{(20,20)} = -0.18$, $P = 0.862$) but had significantly less ovarioles than the high strain parent colony ($Z_{(20,20)} = -4.18$, $P < 0.001$). The HBC colonies (3.0 (2.0–5.0)) were overall more similar to the high strain than to their hybrid parent (Figure 1). The HBC colony that was chosen for mapping (5.0 (3.50–6.0)) was not significantly different from its high strain parent ($Z_{(20,20)} = -0.33$, $P = 0.758$) but had significantly more ovarioles than the hybrid ($Z_{(20,20)} = -3.7$, $P < 0.001$). Workers in the LBC colonies had more ovarioles (3.0 (2.0–4.0)) than in either of their parental colonies (hybrid and low strain). Accordingly, the LBC colony used for QTL mapping had workers with more ovarioles (3.0 (2.0–4.0)) than the hybrid ($Z_{(20,20)} = -2.6$, $P = 0.010$) and low strain ($Z_{(20,20)} = -2.9$, $P = 0.003$) parent colony (Figure 1).

The mapping populations, gathered from one particular HBC and LBC colony, showed a wide phenotypic variation (Figure 2) and were consistent with the screening results described above. In the HBC, the medians of minimum and maximum ovary size measurements were 3.0 (2.0–4.0) and 5.0 (4.0–6.0), respectively. The resulting mean ovary size and ovary asymmetry scores showed a median of 4.0 (3.0–5.0) and 0.20 (0.11–0.40), respectively. In the LBC, the medians of minimum and maximum ovary size measurements were 2.0 (1.0–3.0) and 3.0 (3.0–4.0), respectively. The resulting mean ovary size and ovary asymmetry scores showed a median of 2.5 (2.0–3.5) and 0.25 (0.14–0.50), respectively. Compared with the LBC, the HBC had a significantly larger minimum ($Z_{(392,393)} = 8.4$, $P < 0.001$), maximum ($Z_{(392,393)} = 10.5$, $P < 0.001$) and mean ovary size ($Z_{(392,393)} = 10.5$, $P < 0.001$), as well as asymmetry scores ($Z_{(392,393)} = 2.3$–2.9, all $P < 0.019$).

The size of both ovaries in a worker were significantly correlated in both backcrosses (HBC: Spearman's $R = 0.58$, $n = 392$, $P < 0.001$; LBC: $R = 0.61$, $n = 393$, $P < 0.001$). Ovary asymmetry was unrelated to maximum ovary size (HBC: $R = 0.01$, $n = 392$, $P = 0.780$; LBC: $R = 0.01$, $n = 393$, $P = 0.899$) but negatively correlated to minimum ovary size (HBC: $R = -0.74$, $n = 392$, $P < 0.001$; LBC: $R = -0.70$, $n = 393$, $P < 0.001$).

QTL analyses

The HBC map contained 231 SNP and microsatellite markers and covered all chromosomes (linkage groups) with an average marker spacing of 17.9 cM. The marker placement resulted in 98.8% of the total genome map (Solignac *et al*., 2007) within 20 cM of at least one genetic marker. The largest coverage gap between two markers was 48 cM on chromosome 2. A lack of genetic variability in this mapping population was indicated in three genome regions by an unusual number of monomorphic microsatellite loci. In the first region (contigs 9.21–9.23), we found 18 out of 20 microsatellites tested to be monomorphic. On contigs 12.25–12.27 seven out of eight and on contig 13.09 five out of six tested microsatellites were monomorphic.

In the HBC, one significant QTL was identified by interval mapping on contig 3.38 (near SNP marker 'est2829', peak at 13.3 Mb), affecting minimum ovariole number (LOD = 4.1, effect size of 1.6 ovarioles, explaining 14.0% of the total phenotypic variance), maximum ovariole number (LOD = 4.7, effect size 1.9, 15.4% variance explained) and mean ovariole number (LOD 5.1, effect size 3.4 16.7% variance explained). Multiple QTL mapping increased the LOD scores of 4.7 (effect size 1.6, 14.0% variance explained), 5.4 (effect size 1.9, 16.6% variance explained) and 5.9 (effect size 3.5, 17.8% variance explained), respectively (Figure 3a), and

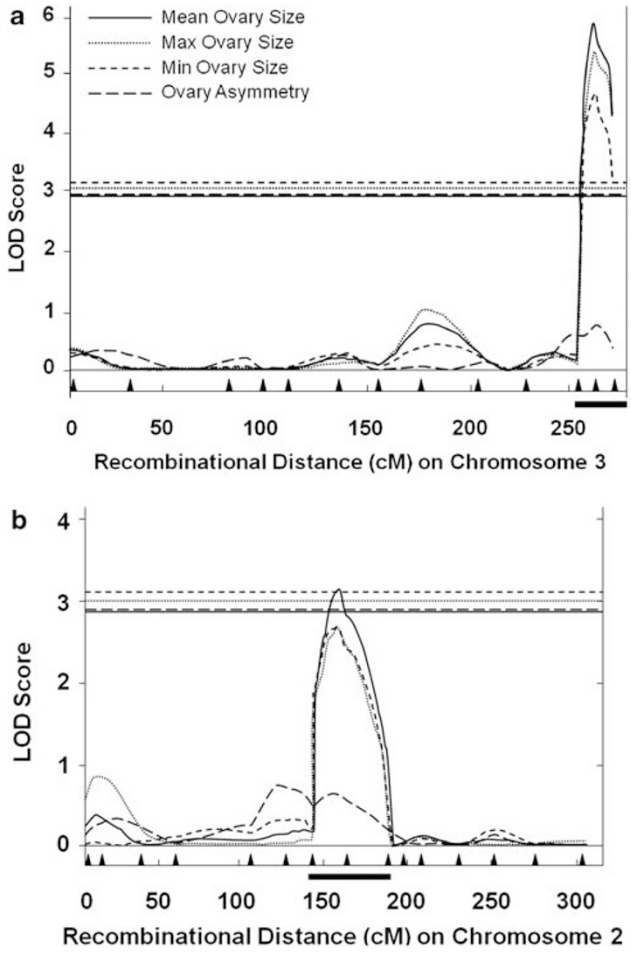

Figure 3 Two significant QTL on chromosome 3 (**a**) and 2 (**b**) for worker ovary size detected with conventional, individual QTL mapping showed the high genetic correlation between the three ovary size measurements (minimum, maximum and total ovary size) but provided little evidence for QTL overlap with ovary asymmetry, in spite of the phenotypic, negative correlation between ovary asymmetry and minimum ovary size.

indicated another significant QTL on contig 2.34 (near microsatellite marker 'At082', peak at 10.7 Mb) with LOD scores for minimum ovariole number of 2.7 (effect size 1.2, 8.8% variance explained), maximum ovariole number of 2.7 (effect size 1.4, 9.4% variance explained), and mean ovariole number of 3.1 (effect size 2.7, 10.5% variance explained) (Figure 3b).

One additional suggestive QTL (LOD > 2.0) between contigs 8.16 and 8.19 (near SNP marker 'est6637', peak at 4.1 Mb) was identified by interval mapping with the strongest effect on maximum ovary size (LOD = 2.3, effect size 1.5, 9.3% variance explained), followed by mean ovary size (LOD = 2.2, effect size 2.5, 8.6% variance explained) and minimum ovary size (LOD = 1.6, effect size 1.0, 6.1% variance explained). MQM models did not increase effect size or significance of this locus. No QTL was detected for ovary asymmetry. The empirically determined genome-wide LOD significance thresholds for interval mapping were 2.9 for total ovary size, 3.0 for maximum ovary size, 3.1 for minimum ovary size and 2.9 for ovary asymmetry.

In the LBC, 221 SNP and microsatellite markers covered all linkage groups, with an average marker

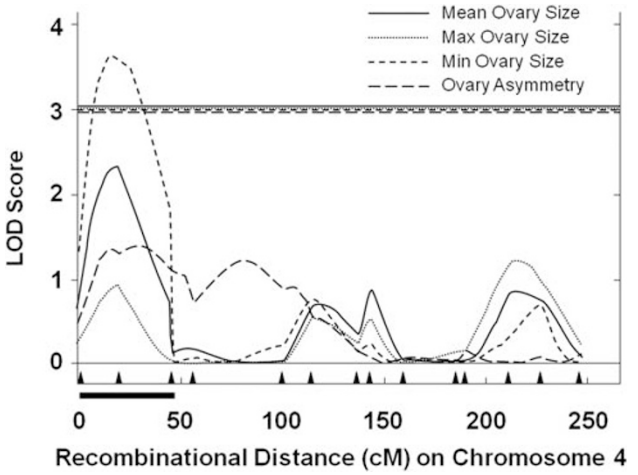

Figure 4 One significant QTL for minimum ovary size was detected on chromosome 4, showing minor effects on total ovary size, ovary asymmetry and maximum ovary size.

spacing of 17.8 cM. Markers were spaced so that 98.4% of the total mapable genome was within 20 cM of at least one marker. The largest coverage gap between two adjacent makers was 46 cM on chromosome 3. A lack of genetic variability in the mapping population was suggested in regions 12.25–12.27 and 15.35–15.37 because we, respectively, tested nine and six monomorphic microsatellite loci before identifying one polymorphic marker.

One significant QTL for minimum ovariole number was identified on contig 4.7 (near microsatellite marker K0433B, peak at 1.8 Mb) with LOD = 3.4 (effect size 1.2 ovarioles, explaining 15.2% of the total phenotypic variance). A suggestive QTL was located on contig 15.35 (near microsatellite marker 6425, peak at 9.3 Mb) with LOD = 2.3 (effect size 1.1, 11.8% variance explained). MQM mapping increased the LOD for the significant QTL to 3.7 (effect size 1.2, 15.0% variance explained) (Figure 4) and for the suggestive QTL on contig 15.35 to LOD = 2.7 (effect size 1.2, 12.7% variance explained). These two regions were also included in the MQM model for mean ovariole number despite smaller effects (contig 4.7: LOD = 2.3, effect size 2.0, 8.5% variance explained; contig 15.35: LOD = 2.0, effect size 2.1, 9.5% variance explained) but they did not show significant or suggestive effects on maximum ovariole number or ovary asymmetry.

The MQM analysis of mean ovariole number suggested an additional QTL (LOD = 2.1, effect size 2.0, 8.4% variance explained) on contig 13.05 near SNP marker ahb4072 (peak at 1.8 Mb). No QTL was detected for ovary asymmetry. The empirically determined genome-wide LOD significance thresholds were 3.0 for all four traits.

Pair-wise epistasis tests among all six significant and suggestive QTL did not reveal any significant epistatic effects in the HBC or the LBC.

Candidate genes

For the significant QTL on chromosomes 2, 3 and 4 from the individual QTL mapping experiment, we determined the 1.5 LOD support regions (indicated by black bars under *x*-axis in Figures 3 and 4) from the interval

mapping LOD traces to evaluate the positional candidates for putative functional evidence of involvement in worker ovary size determination. On chromosome two (Figure 3b), 107 gene models (see Supplementary Information) were found in the QTL region. Of these, 43 were hypothetical loci or genes without close homologs of known functions. Among the remaining, five genes were particularly promising: 1) LOC408691, the homolog of *Roc1a*, a ubiquitin-protein ligase that has been associated with cell proliferation; 2) LOC411671, the homolog of *Ric*, a member of the *Ras*-kinase family that is in general involved in intra-cellular signaling, including the IIS pathway and anti-apoptotic functions; 3) LOC726899, the homolog of *thread*, a ubiquitin–protein ligase that is involved in apoptosis and has been associated with several ovary defects in *Drosophila*; 4) LOC412132, the homolog of *genghis khan*, a protein kinase that is involved in intracellular signaling and actin regulation; and 5) LOC408694, the homolog of fimbrin, another regulator of actin.

The region of the QTL on chromosome three (Figure 3a) contained 84 genes (see Supplementary Information), with 13 hypothetical loci or genes without ascribable function. The top candidate in this region was LOC551949, the homolog of *loki* (alternatively known as *chk2*), which is involved in signal transduction controlling apoptosis and germ cell development. Numerous second-tier candidates exist, including LOC413805 (homolog of *Nop60B*), LOC413817 (homolog of *IP3K1*), several other intracellular signaling components, including LOC410904, which presumably is interacting with *actin* and Rho-GTPases, and the *NDP kinase 6* homolog LOC724131. The QTL region on chromosome four (Figure 4) included 65 genes including 10 hypothetical loci (see Supplementary Information). The best candidates among the remaining 55 genes were identified as the *orb* homolog (LOC411907), a putative mRNA binding factor, and the gene LOC726966, a peptidase with homologies to *C42D8.5a* and a putative role in hormone processing.

Discussion

Previously, we have shown that variation in ovary size correlates with differences in foraging behavior (Amdam *et al*., 2006a; Rueppell *et al*., 2008) and with allelic variation in two genome regions that had been identified as *pln* QTL (Wang *et al*., 2009). These associations between worker ovary size and foraging division of labor provide support for the reproductive ground plan hypothesis of social evolution in honeybees and confirm that ovary size is a trait of the *pln* syndrome (Page and Amdam, 2007). Similar to previously studied behavioral traits of the *pln* syndrome (Rueppell *et al*., 2004, 2006), worker ovary size is affected pleiotropically by some of the *pln* QTL (Wang *et al*., 2009) but further QTL exist that exceed the phenotypic effect of the *pln* QTL. This suggests that most traits of the *pln* syndrome may be affected by a mixture of central loci that affect the whole suite of traits and peripheral, trait-specific loci. The newly described QTL could be trait-specific, downstream elements, but we cannot rule out more general functions because targeted studies on their pleiotropic effects on other aspects of the *pln* syndrome have not yet been performed.

The ovary size of honeybee workers (*Apis* spp.) is unusually variable compared with other non-*Apis* bee species (Michener, 2000) and our crosses show that significant genetic variation for this trait exists among European honeybees. This variation in worker ovary size may be adaptive through effects on colony level division of labor (Page and Amdam, 2007). Thus, ovary size may be under diversifying selection rather than be directionally selected for by individual worker reproductive opportunities, as might be the case in other *A. mellifera* populations (Phiancharoen *et al*., 2010). The repeated findings of significant genetic variation for worker ovary size in *A. mellifera* may also be explained mechanistically: The pronounced phenotypic plasticity for ovary size between the queen and worker castes is based on differential expression of multiple genes (Barchuk *et al*., 2007) and any allelic variation in these genes can also lead to genetic variation for worker ovary size (Linksvayer *et al*., 2009b).

The phenotypic differences among the total of 25 colonies that were screened in this study, including the high- and low-*pln* strain sources, represented a combination of direct and indirect genetic effects because each larval cohort was reared in its own hive (Linksvayer *et al*., 2009a). Indirect effects through larval feeding have a significant part in worker ovary size (Hoover *et al*., 2006; Wegener *et al*., 2009). Yet, only modest differences in ovary size were detected between the HBC and LBC workers that were raised in their own hives (used in screening) and the corresponding mapping populations that were subsequently raised in unrelated, common rearing hives. Regardless of the size of indirect genetic effects, the differences between our HBC and LBC mapping populations and the QTL identified in this study cannot be explained by different social environments but represent differences of the individual larvae (that is, direct genetic effects).

Three significant QTL for worker ovary size but no QTL for ovary asymmetry were detected. Two of the three significant QTL were detected in the HBC and one in the LBC with no overlap among them or between them and the three suggestive QTL regions that we also reported. This lack of correspondence between the HBC and LBC was also found in previous QTL studies of reciprocal backcrosses of the selected *pln* strains (Rueppell *et al*., 2004, 2006; Rüppell *et al*., 2004; Rueppell, 2009). It could be explained by allelic differences for all QTL between the two hybrid queens that produced the two backcrosses. However, it seems more likely to be due to dominance and/or genetic background effects, which may be common for complex traits in general (Sinha *et al*., 2006). We suggest that the natural genetic architecture of worker ovary size within European honeybees is comprised of additive and non-additive effects. However, our QTL did not display any detectable epistatic effects among each other, which is consistent with previous results of additive *pln* effects on ovary size (Wang *et al*., 2009) but differs from the genetic architecture of foraging specialization, the central behavioral component of the *pln* syndrome (Rüppell *et al*., 2004).

The most significant of the individual QTL on chromosome 3 (Figure 3a) displayed a LOD score of almost six, explaining 17% of the phenotypic variation, comparable with the most significant QTL identified for inter-specific ovary size differences in *Drosophila*

(Orgogozo *et al*., 2006). Although this estimate is likely an overestimate (Beavis, 1998), it represents a QTL of major effect. The remaining QTL explain decreasing amounts of variation and our results are compatible with a model of multiple genetic factors of varying effect size. In addition to the three suggestive QTL explaining between 8.4 and 12.7% of the total phenotypic variation, the small but significant effects of the *pln* QTL on ovary size (Wang *et al*., 2009) suggest that several more genes may be involved that could not be detected in our genome-wide QTL analysis because of its inherent lack of statistical power (Otto and Jones, 2000; Benjamini and Yekutieli, 2005). However, we have detected several major factors that contribute to the direct genetic differences in ovary size between the high- and low-*pln* strains of honeybees and are consequently a part of the genetic architecture of the more general *pln* syndrome. We employed stringent, genome-wide significance criteria instead of methods based on false discovery rates (Benjamini and Yekutieli, 2005) and clearly distinguish between significant and suggestive QTL (Lander and Kruglyak, 1995).

In contrast to previous mapping studies of ovary size (Orgogozo *et al*., 2006; Bergland *et al*., 2008; Linksvayer *et al*., 2009b), our independent measurement of both ovaries in each individual allowed for an assessment of how different the ovary size can be within worker bees. The intra-individual phenotypic differences were quantified as an asymmetry score without distinguishing between the left and right side, although some evidence for directional asymmetry for ovariole number in worker honeybees exists (Chaud-Netto and Bueno, 1979). The genotypic differentiation of asymmetry between HBC and LBC was weak and no significant or suggestive QTL could be detected. This outcome was surprising because minimum ovary size was negatively correlated to ovary asymmetry, explaining half of its phenotypic variation. Minimum ovary size differed strongly between the backcrosses and it was significantly affected by the identified QTL, explaining 31% in the HBC and 27% in the LBC of the phenotypic variation. These results suggest that the portion of the phenotypic variation that is responsible for the negative correlation (individuals with smaller minimum ovary size being more asymmetrical) is largely independent of the identified QTL and thus must be attributed to either environmental, indirect genetic or interaction effects or unidentified genetic factors of presumably small effect.

The phenotypic correlation between the size of the two ovaries of workers was weaker than anticipated and both mapping populations contained workers with highly asymmetric values, such as ovaries with 11 and 0, 10 and 2 or 8 and 1 ovarioles. However, symmetry was common and the left and right ovaries of over half of the workers differed by no more than one ovariole from each other. Additionally, maximum and minimum ovariole scores displayed very similar LOD profiles across the genome. This suggests that the identified QTL do not separately affect the minimum or maximum ovary size. Instead, they influence the mean ovary size with correlated changes of minimum and maximum.

In addition to our main QTL analyses based on individual genotypes, we analyzed genotype data from selective, fractionated DNA pools as a preliminary QTL mapping strategy. In contrast to theoretical expectations and previous studies (Darvasi and Soller, 1994; Docherty *et al*., 2007; Korol *et al*., 2007; Wang *et al*., 2007) the two analyses showed little agreement (see Supplementary Information). This could be because of the fact that different individuals from the same mapping populations were analyzed. However, an empirical follow-up test (see Supplementary Information) suggested that this case of incongruence between the two approaches is most likely due to technical errors during the pooled DNA SNP genotyping (see Supplementary Information; Mariasegaram *et al*., 2007) and consequently we discounted the results from the pooled analyses. These results suggest that pooled mapping may not always be accurate and should be verified with individual mapping when possible (Rueppell, 2009).

The three significant QTL with confidence intervals of a combined length of 6.2 Mb contained 245 gene models. This represents approximately 2.6% of the clonable genome size and 2.4% of the official gene set of *A. mellifera*. Although these numbers are too high for a comprehensive study of all positional candidate genes, our results reconfirm that QTL confidence intervals in the honeybee are comparatively small and thus informative (Hunt *et al*., 2007). Strong variation exists for how much information is available for individual candidate genes, and therefore our selection of plausible functional candidates is necessarily biased toward well-studied genes.

Based on the available information about ovary development in honeybee workers (Dedej *et al*., 1998; Schmidt Capella and Hartfelder, 1998, 2002; Hepperle and Hartfelder, 2001), our search for functional candidates focused on gene predictions with presumed functions in cellular apoptosis or proliferation, hormone and signal transduction, and actin regulation. Several genes were found in each category and the most promising candidates were identified by at least two independent lines of evidence. However, our candidates did not contain well-known central regulators, such as *PKG*, *vitellogenin* or *IIS* genes.

For the QTL on chromosome 2, the homolog of *thread* seems most promising because *thread* (or *Diap1*) is a central inhibitor of cellular apoptosis by caspase inhibition (Steller, 2008) and ecdysteroid-mediated activity differences have been associated with different levels of apoptosis (Yin *et al*., 2007). These molecular functions seem to be evolutionarily conserved (Steller, 2008). In addition, *thread* mutations have ovary degeneration phenotypes in *Drosophila* (Rodriguez *et al*., 2002). For the strongest QTL, found on chromosome 3, the homolog of *loki* is our top candidate. *Loki* (or *Dmchk2*) is a serine/threonine kinase that is also involved in the regulation of apoptosis (Xu *et al*., 2001) and has a role in germ cell development (Oishi *et al*., 1998). The top candidate gene of the fourth chromosome QTL, that mostly affected minimum ovary size, was the homolog of *orb* because *orb* acts in *Drosophila* as a transcriptional regulator that is responsible for the control of polyadenylation in the oocycte (Castagnetti and Ephrussi, 2003) and has been associated with small, undeveloped ovary phenotypes (Lantz *et al*., 1994).

In both backcrosses, we identified several regions with a pronounced lack of heterozygosity in the investigated microsatellite loci. Only by screening further loci away from these regions did we find polymorphic markers in our mapping populations. One of these regions was

identical between the HBC and LBC. Given that, overall 47% of all successfully-amplified loci were polymorphic, even five adjacent monomorphic microsatellites have only a 2% probability to occur by chance alone. These genomic regions could have originated from convergent genetic drift in both selected strains that were used to establish the mapping populations. However, their independent establishment and periodic outbreeding to unrelated commercial honeybees makes this explanation unlikely. These regions seem also too large (8–67 cM) to represent selected haplotype blocks although the largest region (scaffolds 9.21–9.23) covers multiple genes that are assumed to be under selection (Zayed and Whitfield, 2008). Two remaining potential explanation for these areas are gene conversion (Ziegler *et al*., 2009) and segregation distorters (Hurst and Werren, 2001), such as genes that have a dominant lethal effect in workers.

Our study primarily contributes to the understanding of the genetic architecture of the *pln* syndrome, which is important for understanding social evolution. In particular, ovary size relates to caste development (Linksvayer *et al*., 2009b), worker reproductive potential (Makert *et al*., 2006) and division of labor among workers (Amdam *et al*., 2006a; Wang *et al*., 2009). However, our results also further the understanding of genetic influences on ovary size of insects in more general terms. Honeybees provide an ideal model system to study ovary size determination because (I) extensive phenotypic variation exists, (II) much is known about the cellular processes of ovary development (Hartfelder and Emlen, 2005), and (III) predictable changes during caste differentiation can be linked to developmental processes, such as gene expression patterns (Evans and Wheeler, 1999; Barchuk *et al*., 2007). We have identified a tractable number of QTL and candidate genes that provide complementary data for evolutionary genetic studies of this important, yet understudied trait.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This study would not have been possible without the practical help of Kristen Ward, Kathe Isle, Michael Munday, Megan Wallrichs, Abby Stanley and Anny Pena. Furthermore, we would like to thank Charlie Whitfield (UIUC) for essential help with the pooled DNA SNP genotyping. Financial support for this project was provided be the National Science Foundation (postdoctoral fellowship to TAL and Grants 0615502 and 0634182 to OR and GVA), the Agriculture and Food Research Initiative of the USDA National Institute of Food and Agriculture (2010–65104–20533 to OR), and the National Institute on Aging (NIA P01 AG22500 to REP). GVA was additionally supported by the Norwegian Research Council (180504 and 185306) and the PEW Foundation.

References

Amdam GV, Csondes A, Fondrk MK, Page RE (2006a). Complex social behaviour derived from maternal reproductive traits. *Nature* **439**: 76–78.

Amdam GV, Nilsen KA, Norberg K, Fondrk MK, Hartfelder K (2007). Variation in endocrine signaling underlies variation in social life history. *American Naturalist* **170**: 37–46.

Amdam GV, Norberg K, Fondrk MK, Page Jr RE (2004). Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. *Proc Natl Acad Sci USA* **101**: 11350–11355.

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

Amdam GV, Norberg K, Page RE, Erber J, Scheiner R (2006b). Downregulation of vitellogenin gene activity increases the gustatory responsiveness of honey bee workers (*Apis mellifera*). *Behav Brain Res* **169**: 201–205.

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 Developm Biol* **7**: 70.

Beavis WD (1998). QTL analysis: power, precision, and accuracy. In: Paterson AH (ed) *Molecular Dissection of Complex Traits*. CRC Press: Boca Raton, FL, pp 145–162.

Ben-Shahar Y, Robichon A, Sokolowski MB, Robinson GE (2002). Influence of gene action across different time scales on behavior. *Science* **296**: 741–744.

Benjamini Y, Yekutieli D (2005). Quantitative trait loci analysis using the false discovery rate. *Genetics* **171**: 783–790.

Bergland AO, Genissel A, Nuzhdin SV, Tatar M (2008). Quantitative trait loci affecting phenotypic plasticity and the allometric relationship of ovariole number and thorax length in *Drosophila melanogaster*. *Genetics* **180**: 567–582.

Castagnetti S, Ephrussi A (2003). Orb and a long poly(A) tail are required for efficient oskar translation at the posterior pole of the *Drosophila* oocyte. *Development* **130**: 835–843.

Chaud-Netto J, Bueno OC (1979). Number of ovarioles in workers of *Apis mellifera adansonii* and *Apis mellifera ligustica*: a comparative study. *J Apicult Res* **18**: 260–263.

Churchill GA, Doerge RW (1994). Empirical threshold values for quantitative trait mapping. *Genetics* **138**: 963–971.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, Grimwood J *et al.* (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**: 1928–1933.

Darvasi A, Soller M (1994). Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. *Genetics* **138**: 1365–1373.

de Bono M, Bargmann CI (1998). Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* **94**: 679–689.

Debelle JS, Hilliker AJ, Sokolowski MB (1989). Genetic localization of foraging (for)-a major gene for larval behavior in *Drosophila melanogaster*. *Genetics* **123**: 157–163.

Dedej S, Hartfelder K, Aumeier P, Rosenkranz P, Engels W (1998). Caste determination is a sequential process: effect of larval age at grafting on ovariole number, hind leg size and cephalic volatiles in the honey bee (*Apis mellifera carnica*). *J Apicult Res* **37**: 183–190.

Docherty SJ, Butcher LM, Schalkwyk LC, Plomin R (2007). Applicability of DNA pools on 500 KSNP microarrays for cost-effective initial screens in genomewide association studies. *Bmc Genomics* **8**: 214.

Evans JD, Wheeler DE (1999). Differential gene expression between developing queens and workers in the honey bee, *Apis mellifera*. *Proc Natl Acad Sci USA* **96**: 5575–5580.

Guidugli KR, Nascimento AM, Amdam GV, Barchuk AR, Omholt S, Simoes ZLP *et al.* (2005). Vitellogenin regulates hormonal dynamics in the worker caste of a eusocial insect. *Febs Lett* **579**: 4961–4965.

Hartfelder K, Emlen DJ (2005). Endocrine control of insect polyphenism. In: Iatrou K, Gilbert L and Gill S (eds). *Comprehensive Molecular Insect Science*, Vol. 3 Elsevier: Amsterdam, pp 651–703.

Hepperle C, Hartfelder K (2001). Differentially expressed regulatory genes in honey bee caste development. *Naturwissenschaften* **88**: 113–116.

Hoover SER, Higo HA, Winston ML (2006). Worker honey bee ovary development: seasonal variation and the influence of larval and adult nutrition. *J Comp Physiol B* **176**: 55–63.

Humphries MA, Fondrk MK, Page RE (2005). Locomotion and the pollen hoarding behavioral syndrome of the honeybee (*Apis mellifera* L). *J Comp Physiol A* **191**: 669–674.

Hunt GJ, Amdam GV, Schlipalius D, Emore C, Sardesai N, Williams CE *et al.* (2007). Behavioral genomics of honeybee foraging and nest defense. *Naturwissenschaften* **94**: 247–267.

Hunt GJ, Page Jr RE, Fondrk MK, Dullum CJ (1995). Major quantitative trait loci affecting honey bee foraging behavior. *Genetics* **141**: 1537–1545.

Hurst GDD, Werren JH (2001). The role of selfish genetic elements in eukaryotic evolution. *Nature Reviews Genetics* **2**: 597–606.

Kaun KR, Riedl CAL, Chakaborty-Chatterjee M, Belay AT, Douglas SJ, Gibbs AG *et al.* (2007). Natural variation in food acquisition mediated via a *Drosophila* cGMP-dependent protein kinase. *J Experim Biol* **210**: 3547–3558.

Korol A, Frenkel Z, Cohen L, Lipkin E, Soller M (2007). Fractioned DNA pooling: a new cost-effective strategy for fine mapping of quantitative trait loci. *Genetics* **176**: 2611–2623.

Lander ES, Botstein D (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185–199.

Lander ES, Kruglyak L (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* **11**: 241–247.

Lantz V, Chang JS, Horabin JI, Bopp D, Schedl P (1994). The Drosophila orb RNA-binding protein is required for the formation of the egg chamber and establishment of polarity. *Genes & Development* **8**: 598–613.

Lattorff HMG, Moritz RFA, Crewe RM, Solignac M (2007). Control of reproductive dominance by the thelytoky gene in honeybees. *Biology Letters* **3**: 292–295.

Lincoln SE, Daly MJ, Lander ES (1993). *Technical Report*. Whitehead Institute for Biomedical Research: Cambridge, MA.

Linksvayer TA, Fondrk MK, Page RE (2009a). Honeybee social regulatory networks are shaped by colony-level selection. *American Naturalist* **173**: E99–E107.

Linksvayer TA, Rueppell O, Siegel A, Kaftanoglu O, Page Jr RE, Amdam GV (2009b). The genetic basis of transgressive ovary size in honey bee workers. *Genetics* **183**: 693–707.

Lucas C, Sokolowski MB (2009). Molecular basis for changes in behavioral state in ant social behaviors. *Proc Natl Acad Sci USA* **106**: 6351–6356.

Mackay TFC (2001). The genetic architecture of quantitative traits. *Ann Rev Gen* **35**: 303–339.

Makert GR, Paxton RJ, Hartfelder K (2006). Ovariole number-a predictor of differential reproductive success among worker subfamilies in queenless honeybee (*Apis mellifera* L) colonies. *Behav Ecol Sociobiol* **60**: 815–825.

Mariasegaram M, Robinson NA, Goddard ME (2007). Empirical evaluation of selective DNA pooling to map QTL in dairy cattle using a half-sib design by comparison to individual genotyping and interval mapping. *Genetics Selection Evolution* **39**: 267–283.

Mery F, Belay AT, So AKC, Sokolowski MB, Kawecki TJ (2007). Natural polymorphism affecting learning and memory in *Drosophila*. *Proc Natl Acad Sci USA* **104**: 13051–13055.

Michener CD (2000). *The Bees of the World*. Johns Hopkins University Press: Baltimore, MD.

Oishi I, Sugiyama S, Otani H, Yamamura H, Nishida Y, Minami Y (1998). A novel *Drosophila* nuclear protein serine threonine kinase expressed in the germline during its establishment. *Mech Develop* **71**: 49–63.

Orgogozo V, Broman KW, Stern DL (2006). High-resolution quantitative trait locus mapping reveals sign epistasis controlling ovariole number between two *Drosophila* species. *Genetics* **173**: 197–205.

Otto SP, Jones CD (2000). Detecting the undetected: estimating the total number of loci underlying a quantitative trait. *Genetics* **156**: 2093–2107.

Oxley PR, Thompson GJ, Oldroyd BP (2008). Four quantitative trait loci that influence worker sterility in the honeybee (*Apis mellifera*). *Genetics* **179**: 1337–1343.

Page RE, Amdam GV (2007). The making of a social insect: developmental architectures of social design. *Bioessays* **29**: 334–343.

Page RE, Fondrk MK (1995). The effects of colony level selection on the social organization of honey bee (*Apis mellifera* L) colonies-colony level components of pollen hoarding. *Behav Ecol Sociobiol* **36**: 135–144.

Page RE, Fondrk MK, Hunt GJ, Guzman-Novoa E, Humphries MA, Nguyen K. *et al.* (2000). Genetic dissection of honeybee (*Apis mellifera* L) foraging behavior. *J Hered* **91**: 474–479.

Page RE, Scheiner R, Erber J, Amdam GV (2007). The development and evolution of division of labor and foraging specialization in a social insect (*Apis mellifera* L). *Curr Topics Develop Biol* **74**: 253–286.

Palmer AR, Strobeck C (1986). Fluctuating asymmetry-measurement, analysis, patterns. *Annual Review of Ecology and Systematics* **17**: 391–421.

Pankiw T (2003). Directional change in a suite of foraging behaviors in tropical and temperate evolved honey bees (*Apis mellifera* L). *Behav Ecol Sociobiol* **54**: 458–464.

Pankiw T, Page RE (2000). Response thresholds to sucrose predict foraging division of labor in honeybees. *Behav Ecol Sociobiol* **47**: 265–267.

Pankiw T, Page RE (2001). Genotype and colony environment affect honeybee (*Apis mellifera* L) development and foraging behavior. *Behav Ecol Sociobiol* **51**: 87–94.

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.

Phillips TJ, Belknap JK (2002). Complex-trait genetics: emergence of multivariate strategies. *Nat Rev Neurosci* **3**: 478–485.

Ragoussis J, Elvidge GP, Kaur K, Colella S (2006). Matrix-assisted laser desorption/ionisation, time-of-flight mass spectrometry in genomics research. *Plos Genet* **2**: 920–929.

Robinson GE, Fernald RD, Clayton DF (2008). Genes and Social Behavior. *Science* **322**: 896–900.

Rodriguez A, Chen P, Oliver H, Abrams JM (2002). Unrestrained caspase-dependent cell death caused by loss of Diap1 function requires the *Drosophila* Apaf-1 homolog, Dark. *Embo J* **21**: 2189–2197.

Rozen S, Skaletsky HJ (2000). Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S and Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* Humana Press: Totowa, NJ, pp 365–386.

Rueppell O (2009). Characterization of quantitative trait loci for the age of first foraging in honey bee workers. *Behav Genet* **39**: 541–553.

Rueppell O, Chandra SBC, Pankiw T, Fondrk MK, Beye M, Hunt GJ *et al.* (2006). The genetic architecture of sucrose responsiveness in the honey bee (*Apis mellifera* L). *Genetics* **172**: 243–251.

Rueppell O, Hunggims E, Tingek S (2008). Association between larger ovaries and pollen foraging in queenless *Apis cerana* workers supports the reproductive ground-plan hypothesis of social evolution. *J Ins Behav* **21**: 317–321.

Rueppell O, Pankiw T, Nielson DI, Fondrk MK, Beye M, Page Jr RE (2004). The genetic architecture of the behavioral ontogeny of foraging in honey bee workers. *Genetics* **167**: 1767–1779.

Rüppell O, Pankiw T, Page Jr RE (2004). Pleiotropy, epistasis and new QTL: the genetic architecture of honey bee foraging behavior. *J Hered* **95**: 481–491.

Ruttner F, Hesse B (1981). Specific differences in the development of ovaries and egg-laying of queenless workers of several races of the honeybee, *Apis mellifera* L. *Apidologie* **12**: 159–183.

Scheiner R, Page RE, Erber J (2001). The effects of genotype, foraging role, and sucrose responsiveness on the tactile learning performance of honey bees (*Apis mellifera* L). *Neurobiol Learn Memory* **76**: 138–150.

Schmidt Capella IC, Hartfelder K (1998). Juvenile hormone effect on DNA synthesis and apoptosis in caste-specific differentiation of the larval honey bee (*Apis mellifera* L) ovary. *J Ins Physiol* **44**: 385–391.

Schmidt Capella IC, Hartfelder K (2002). Juvenile-hormone-dependent interaction of actin and spectrin is crucial for polymorphic differentiation of the larval honey bee ovary. *Cell Tissue Res* **307**: 265–272.

Schuelke M (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotech* **18**: 233–234.

Schug MD, Regulski EE, Pearce A, Smith SG (2004). Isolation and characterization of dinucleotide repeat microsatellites in *Drosophila ananassae*. *Genet Res* **83**: 19–29.

Sih A, Bell A, Johnson JC (2004). Behavioral syndromes: an ecological and evolutionary overview. *Trends in Ecology & Evolution* **19**: 372–378.

Sinha H, Nicholson BP, Steinmetz LM, McCusker JH (2006). Complex genetic interactions in a quantitative trait locus. *Plos Genet* **2**: 140–147.

Solignac M, Mougel F, Vautrin D, Monnerot M, Cornuet JM (2007). A third-generation microsatellite-based linkage map of the honey bee, *Apis mellifera*, and its comparison with the sequence-based physical map. *Genome Biology* **8**: R66.

Steller H (2008). Regulation of apoptosis in *Drosophila*. *Cell Death Differ* **15**: 1132–1138.

Steppan SJ, Phillips PC, Houle D (2002). Comparative quantitative genetics: evolution of the G matrix. *Trends in Ecology & Evolution* **17**: 320–327.

Thuller RHC, Malaspina O, Bueno OC, Chaud-Netto J (1996). Number of ovarioles in workers descendent from crossing between Africanized and Italian honeybees (*Apis mellifera* L): comparing stock, inbred and F1 colonies. *An Soc Entomol Brasil* **25**: 501–506.

Tsuruda JM, Page RE (2009). The effects of foraging role and genotype on light and sucrose responsiveness in honey bees (*Apis mellifera* L). *Behav Brain Res* **205**: 132–137.

van Ooijen JW, Boer MP, Jansen RC, Maliepaard C (2002). *MapQTL 4.0 Software for The Calculation of QTL Positions On Genetic Maps*. Plant Research International: Wageningen, the Netherlands.

Wang J, Koehler KJ, Dekkers JCM (2007). Interval mapping of quantitative trait loci with selective DNA pooling data. *Genetics Selection Evolution* **39**: 685–709.

Wang Y, Amdam GV, Rueppell O, Wallrichs MA, Fondrk MK, Kaftanoglu O *et al*. (2009). PDK1 and HR46 gene homologs tie social behavior to ovary signals. *PLoS ONE* **4**: e4899.

Wang Y, Mutti NS, Ihle KE, Siegel A, Dolezal AG, Kaftanoglu O *et al*. (2010). Down-regulation of honey bee IRS gene biases behavior toward food rich in protein. *Plos Genet* **6**: e1000896.

Wegener J, Lorenz MW, Bienefeld K (2009). Physiological consequences of prolonged nursing in the honey bee. *Insectes Sociaux* **56**: 85–93.

West-Eberhard MJ (1987). Flexible Strategy and Social Evolution. In: Ito Y, Brown JL and Kikkawa J (eds). *Animal Societies Theories and Facts*. Japan Scientific Societies Press: Tokyo, pp 35–51.

West-Eberhard MJ (1996). Wasp societies as microcosms for the study of development and evolution. In: Turillazzi S and West-Eberhard MJ (eds). *Natural History and Evolution of Paper Wasps*. Oxford University Press: New York, pp 290–317.

Whitfield CW, Behura SK, Berlocher SH, Clark AG, Johnston JS, Sheppard WS *et al*. (2006). Thrice out of Africa: ancient and recent expansions of the honey bee, *Apis mellifera*. *Science* **314**: 642–645.

Winston ML (1987). *The Biology of the Honey Bee*. Harvard University Press: Cambridge, Massachusetts, p296.

Wu Q, Brown MR (2006). Signaling and function of insulin-like peptides in insects. *Ann Rev Entomol* **51**: 1–24.

Xu JH, Xin SJ, Du W (2001). *Drosophila* Chk2 is required for DNA damage-mediated cell cycle arrest and apoptosis. *Febs Lett* **508**: 394–398.

Yin VP, Thummel CS, Bashirullah A (2007). Down-regulation of inhibitor of apoptosis levels provides competence for steroid-triggered cell death. *J Cell Biol* **178**: 85–92.

Zayed A, Whitfield CW (2008). A genome-wide signature of positive selection in ancient and recent invasive expansions of the honey bee *Apis mellifera*. *Proc Natl Acad Sci USA* **105**: 3421–3426.

Ziegler JO, Walther M, Linzer TR, Segelbacher G, Stauss M, Roos C *et al*. (2009). Frequent non-reciprocal exchange in microsatellite-containing-DNA-regions of vertebrates. *J Zool Syst Evol Res* **47**: 15–20.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)