Four quantitative trait loci associated with low Nosema ceranae (Microsporidia) spore load in the honeybee Apis mellifera

Qiang Huang[1,2], Per Kryger[3], Yves Le Conte[4], H. Michael G. Lattorff[1,5], F. Bernhard Kraus[1,6], Robin F. A. Moritz[1,7,8]

[1]Institut für Biology/Zoologie, Molekulare Ökologie, Martin-Luther-Universität Halle-Wittenberg, 06099 Halle, Germany
[2]Honeybee Research Institute, Jiangxi Agricultural University, 330045 Nanchang, China
[3]Department of Agroecology, Section of Entomology and Plant Pathology, Aarhus University, Flakkebjerg, 4200 Slagelse, Denmark
[4]INRA, UR 406 Abeilles et Environnement, Laboratoire de Biologie et Protection de l'abeille, Site Agroparc, 84914 Avignon Cedex 9, France
[5]Institut für Biology/Zoologie, Tierphysiologie, Martin-Luther-Universität Halle -Wittenberg, 06099 Halle, Germany
[6]Department of Laboratory Medicine, University Hospital Halle (Saale), Ernst-Grube-Str. 40, 06120 Halle (Saale), Germany
[7]RoBeeTech, Universitatea de Stiinte Agricole si Medicina Vetereinaria Cluj-Napoca, Cluj-Napoca, Romania
[8]Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa

Abstract – *Nosema ceranae* has been recently introduced into the honeybee *Apis mellifera* as a novel microsporidian gut parasite. To locate the genetic region involved in *N. ceranae* infection tolerance, we fed *N. ceranae* spores to haploid drones of a F1 hybrid queen produced from a cross between a queen of a *Nosema*-resistant bred strain and drones of susceptible colonies. The spore loads of the infected F1 drones were used as the phenotype to identify quantitative trait loci (QTLs) associated with *N. ceranae* spore load. One hundred forty-eight infected drones were individually genotyped with microsatellite markers at an average marker distance of 20 cM along the genome. Four QTLs were significantly associated with low spore load, explaining 20.4 % of total spore load variance. Moreover, a candidate gene *Aubergine* (*Aub*) within the major QTL region was significantly overexpressed in drones with low spore loads than in those with high spore loads. Our results confirm the genetic basis of *Nosema* tolerance in the selected strain and show that both additive effects and epistatic interactions among the QTLs interfere with the tested phenotype.

***Apis mellifera* / drone / *Nosema* / QTL**

1. Introduction

Nosema apis and *Nosema ceranae* are two microsporidian gut parasites of the honeybee, *Apis mellifera*. *N. apis* is an evolutionarily old pathogen of the honeybee *A. mellifera* with a moderate virulence, and honeybee colonies can often cure themselves under favorable environ-

Corresponding author: Q. Huang,
qiang.huang@zoologie.uni-halle.de
Manuscript editor: Peter Rosenkranz

mental conditions (Zander 1909; Chen et al. 2009). In contrast, *N. ceranae*, which was originally found in the Asian honeybee *A. cerana* (Fries et al. 1996), is a newly established parasite of *A. mellifera* (Fries et al. 2006; Higes et al. 2006). Although there were reports suggesting that *N. ceranae* had a high virulence in *A. mellifera* at both the colony and the individual level (Higes et al. 2008, 2009), more recent studies suggest a moderate virulence of *N. ceranae* similar to that of *N. apis* (Forsgren and Fries 2010; Gisder et al. 2010).

In Denmark, honeybee colonies have been selected for the absence of *Nosema* infections (*Nosemosis*) for decades (Traynor 2008). Within the Danish honeybee breeding scheme, worker samples from colonies were used to determine the *Nosema* infection level. If the samples were found infected by *Nosema*, the queen of the colony would be replaced with one reared from a colony without *Nosema* infection. Although this mode of selection was at the colony level based on the presence of *Nosema* in the worker samples, the beekeepers physically exchanged queens. This breeding scheme resulted in a honeybee strain in which *Nosema* infections are rarely found and individual bees showed a high tolerance towards experimental *N. ceranae* infections (Huang et al. 2012). Since the biological mechanisms underlying the colony level tolerance are unknown so far, the identification of associated genes (quantitative trait locus, QTL) that had been selected by the beekeepers might help to better understand the actual tolerance mechanisms.

In the honeybee, QTL mapping has been widely used to address quantitative traits and complex social behavior (Hunt et al. 1998; Oxley et al. 2010; Behrens et al. 2011). In the case of *Nosema* tolerance, we can take advantage of haploid drones for QTL mapping because the trait is also expressed in the male sex. The male haploid genetic system is particularly suited for QTL mapping studies as interpretational problems resulting from dominance interactions between the alleles on different homologous chromosomes cannot occur. We here used the drone offspring of a single hybrid queen which resulted from a cross of the queen of the *Nosema*-resistant selected Danish strain with drones of an unselected French strain to identify QTLs associated with low *N. ceranae* spore load. In a subsequent step, we quantified the expression level of a candidate gene identified in the QTL analyses in the same mapping population.

2. Materials and methods

2.1. Instrumental insemination of queen bees

Nine virgin queens of the selected Danish strain were provided by the Department of Integrated Pest Management Research Centre Flakkebjerg, Denmark. The queens were all artificially inseminated with the same mixed sperm (Moritz 1984) of 30 drones of an unselected strain kept at the Laboratoire de Biologie et Protection de l'abeille, INRA Avignon, France. The inseminated queens were introduced into small colonies composed of ~4,000 freshly emerged workers at the apiary of the Martin-Luther-University Halle-Wittenberg, Germany. We reared F1 hybrid queens from the inseminated queens and treated the F1 queens with CO_2 to initiate ovary activation without mating, so that the F1 queens exclusively produced unfertilized eggs developing into drones. A single F1 queen, who first started to lay eggs, was chosen to produce the mapping population. The drones were reared in drone frames in full-sized colonies.

2.2. *N. ceranae* infection

Workers freshly infected with *N. ceranae* were provided by the Laboratoire de Biologie et Protection de l'abeille, INRA Avignon France, as the source of *N. ceranae* spores for subsequent infection. The abdomens of infected workers were homogenized in distilled water, filtered through filtering paper and centrifuged at 3,220×*g* for 10 min. The pellet was re-dissolved in distilled water and centrifuged at 8,700×*g* for 5 min to purify the *N. ceranae* spores. Spores were counted using a Fuchs–Rosenthal hemocytometer and the *Nosema* species was verified by a standard PCR protocol (Hamiduzzaman et al. 2010).

Frames of sealed drone brood of the single F1 hybrid queen were kept in an incubator (34±1 °C, 60 % rel. humidity). Freshly emerged drones were collected daily from the brood frames to provide age standardized individuals (0–24 h). Freshly emerged

workers from brood frames kept in the incubator served as nurse bees. Drones were individually fed with 2 μL sucrose solution containing ~10^5 *N. ceranae* spores. Drones that did not consume the entire solution were discarded. Infected drones and uninfected nurse workers were housed in a wooden cage (depth 13.0 cm×width 10.0 cm×height 11.5 cm) at 34±1 °C and 60 % rel. humidity. Drones receiving 2 μL sucrose solution without any *Nosema* spores were kept under the same conditions. These uninfected drones served as controls for the candidate gene expression analyses. Drones and workers were fed with 50 % sucrose solution ad libitum without pollen during the remaining time of the experiment.

2.3. Drones used in the QTL analysis

In order to select the drones for the QTL mapping, ~10^5 *N. ceranae* spores were individually fed to 319 drones of the single F1 queen. All drones were sampled on day six post infection to let the *Nosema* infection spread out of the ventriculi. Dead drones were daily removed and recorded. The guts of all sampled drones were individually removed and homogenized in 500 μL distilled H_2O. Ten microliters homogenized solution was again diluted by 90 μL distilled water to count the *N. ceranae* spores in a Fuchs–Rosenthal hemocytometer. By day six post infection, 64 out of the 319 infected drones had died. It is unclear whether the 64 drones were killed by the infection or the behavior of the "nurse" workers in the cage. Since new generation *N. ceranae* spores were detected approximately 4 days after the infection (Higes et al. 2007), all drones that were dead before day four post infection (n=55) had no meaningful phenotype with relation to *Nosema* reproduction and were discarded from further analyses. The guts of the remaining 255 drones were individually removed to determine the spore load (median and range, 5.6×10^5 and 2.4×10^6). No *N. ceranae* spores were found in the control groups. 76 drones with $<3.6\times10^5$ spores (low; median: 8×10^4, range: 3.4×10^5) and 72 drones with $>8.2\times10^5$ spores (high; median: 1.1×10^6, range: 1.7×10^6) were selected for the QTL mapping (Figure S1). The number of spores was significantly different between these two groups (Mann–Whitney U test, P<0.001).

2.4. Heterozygous marker selection

After the removal of the gut, the abdomen was immediately preserved in RNA-later® (Sigma-Aldrich) and stored at −80 °C for candidate gene expression analyses. The remaining thoraces and heads were preserved in 75 % ethanol and stored at −20 °C until DNA extraction using 5 % Chelex 100 (Walsh et al. 1991) for the genotyping. We genotyped the F1 queen using 732 fluorescence-labeled microsatellite markers by multiplex PCR (according to Behrens et al. 2011) to select heterozygous markers for the individual genotyping.

2.5. Phase determination

As the mother queen of the F1 queen was accidently killed and removed by the worker honeybees, we could not use her directly to determine the phase of the used markers. We sampled ten drones each from ten colonies of the selected Danish honeybee strain, and genotyped them to determine the phase (Danish alleles and French alleles). We identified all alleles for all heterozygous loci of the selected drones. As all the drones used for the QTL mapping were from a single F1 queen with a Danish and a French allele at every locus, Danish alleles could be unambiguously identified, whenever they were unique and different from the French strain. We used these unique markers as anchor loci to determine the phase of all other loci in the mapping population based on linkage analysis.

2.6. Single QTL analysis

Drones were individually genotyped with all heterozygous makers using the MegaBACE 1000 DNA analysis system (Amersham Biosciences, Germany) and scored with the MegaBACE Fragment Profiler Version 1.2. We used the number of spores in 1 μL of the diluted gut homogenates as the mapping phenotype. Interval mapping was performed to identify the significant QTL associated with low *N. ceranae* spores using Window QTL Cartographer 2.5 (Wang et al. 2011). The statistical significance of the putative QTL was calculated with a 5,000 times permutation test walking along the genome with a genotyping error probability of 0.01. We used

interval mapping to identify major QTLs without considering interaction effects in the first instance. If a locus showed a significant association ($P<0.05$) with the spore load variance, the nearest markers flanking (up-stream and down-stream) this significant locus were used as the criteria to define a QTL candidate region.

2.7. Interaction QTLs analyses

We searched for potential paired epistatic and additive QTLs using the R/qtl package (Broman et al. 2003; R development core team 2008) with haploid diploid genome type, a genotyping error probability of 0.01 and 1 cM walking steps along the genome. All loci used for the QTL mapping were analyzed pairwise for potential additive and epistatic interactions. The statistical significance of the interactive QTLs was assessed by an analysis of variance (Bonferroni adjusted for the multiply comparison). If the interaction effects of the paired loci showed a significant association ($P<0.05$) with the spore load variance, the nearest markers flanking each of the two loci involved were used to define the two interacting QTL regions.

2.8. Candidate genes and quantitative real time PCR analysis

All open reading frames within the mapped QTL region were classified as candidate genes (NCBI, Map viewer, Amel 4.5). We used the translated amino acid sequence of each candidate gene in a BLASTX search against those of *Drosophila* and human to assess their potential biological functions.

The expression levels of a candidate gene *Aub* (primer sequence forward: 5′~TTACCAACGCCTCTCAA CCAATG~3′; reverse: 5′~AGATATACCAATTCGG CTTGACCAG~3′) was quantified using quantitative real time PCR (qPCR) in 25 drones each from the high and the low spore load group of the mapping population. Ten uninfected drones of the same F1 queen were used as controls. RNA was individually extracted from the abdomen excluding the gut similar to Huang et al. (2012). The cDNA of five drones each was pooled, resulting in five pools with high and low spore load drones each and two pools for the controls. The gene expression in the uninfected control group was used as standard to quantify the expression level of the candidate gene in response to the infection. The genes GPDH-1 and EF-1α that were not regulated by the *N. ceranae* infection were used as reference genes. The qPCR procedures and relative gene expression level analyses were carried out according to Huang et al. (2012).

3. Results

3.1. Interval mapping

Of the 732 screened microsatellite markers, 216 were polymorphic and 60 anchor loci could unambiguously determine the phase of the marker alleles. The average distance of these polymorphic markers was 20 cM (Figure S2) with 90 % of the genome covered with a marker distance of less than 20 cM (99.5 % marker distance <50 cM). So the probability of missing a significant QTL was low due to the linkage gap. A total of 148 drones was individually genotyped using all 216 heterozygous markers. The linkage to the spore load was quantified for each marker by interval mapping using LOD scores. Only locus UN271 (LOD=2.5) on chromosome 14 showed a significant association with the spore load ($p<0.05$). Its nearest flanking loci K1418 and AT198 did not shown significant segregation, suggesting the candidate QTL region to be between these two loci. The significant QTL region spanned 1,598 kbp (between the locus K1418 at 5,355 kbp and AT198 at 6,953 kbp).

3.2. Fine mapping

To further narrow down the detected QTL region, we designed and genotyped the 148 drones with five additional heterozygous microsatellite markers flanking the locus UN271 in the significant QTL region (between the locus K1418 and AT198). The association between the markers and the spore load was re-calculated with interval mapping. The threshold of the significant QTL region was determined by the 5,000 times permutation test which walks along the entire genome for every 1 cM. The locus UN271 (LOD=2.6) again showed a significant association with the *Nosema* spore

load, confirmed by a permutation test along the genome (LOD=2.4, P=0.05) (Figure 1). By adding additional five markers, the statistically significant QTL (QTL_m) region was narrowed down to 338 kbp, spanning from locus HQ1414 at 6,071 kbp to BI103 at 6,409 kbp explaining 7.7 % of total variance. Looking at the actual segregation of the alleles at this locus, 46 out of 76 low spore load drones had the selected Danish allele and 49 out of 72 high spore drones had the unselected French allele. This is a significantly biased allelic distribution towards the predicted phenotype (P<0.001, χ^2 test). The locus UN271 showed significant segregation, but its nearest flanking loci HQ1411 and HQ1414 did not, suggesting the candidate QTL region to be between these two loci. The frequency of the selected Danish allele decreased with the increasing number of spores (Figure S3).

3.3. Epistatic and additive effects

Locus AC184 on chromosome 3 (QTL_{ep3}) significantly interacted with the locus AT129 (QTL_{ep10}) on chromosome 10 in a two-dimensional two QTLs scan, which explained 6.3 % of the total variance (LOD=3.6, P<0.01). Drones with either the Danish or the French alleles at both loci (QTL_{ep3} and QTL_{ep10}), had a higher spore load than those with a combination of two alleles at either locus (Figure 2, Figure S4). This suggests that the interactions between the Danish allele and French allele were important for a low spore load. Additionally, we identified an additive QTL on chromosome 6 (K0616, QTL_{ad}). The additive effect between the Danish allele of QTL_{ad} and the Danish allele of the major QTL_m was also associated with a low spore load explaining 6.4 % of the total variance (LOD=2.5, P<0.05) (Figure S4).

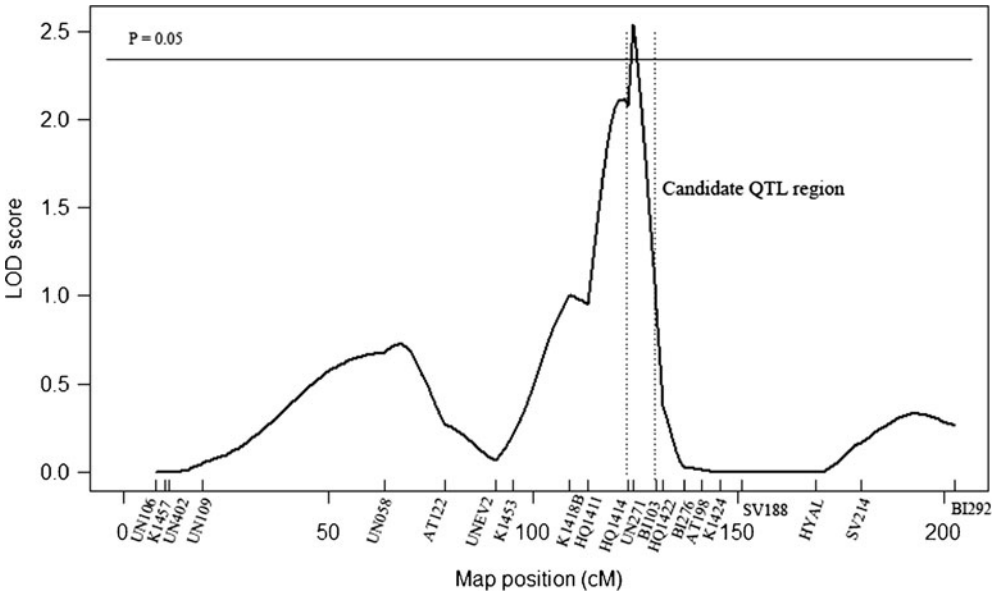

Figure 1. Significant QTL on Chromosome 14. The threshold for the significant QTL was calculated by 5,000 time permutation test along the genome with the genotyping error probability of 1 %. The QTL is statistically significant at P=0.05 level with LOD score of 2.4. By interval mapping, the marker UN271 showed a significant association with the *Nosema* spore load (LOD=2.6). The significant QTL region was located between the marker HQ1414 and BI103 (spanning 338 kbp) explaining 7.7 % of the total variance. The locus UN271 showed significant segregation, but its neighboring loci HQ1411 and HQ1414 did not, suggesting the target region is between these two loci. We used the interval between these two neighboring markers to search the candidate genes (represented by the *vertical dashed line*).

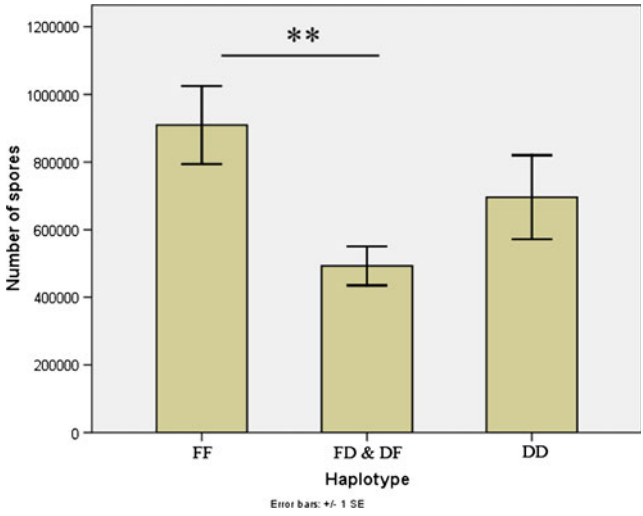

Figure 2. Observed spore load for two-locus genotype groups at the location of two epistatically interactive QTLs. AC184 which is in chromosome 3 significantly interacted with AT129 on chromosome 10 by explaining 6.3 % of total variance. D and F represent the allele originated from Danish and French strains respectively. The genotype "DD" represents the Danish allele at both AC184 in chromosome 3 and AT129 in chromosome 10. The genotype "DF" represents the Danish allele at AC184 and French allele at AT129. Drones with the genotype of "DF" and "FD" had significantly lower (P=0.002) spore load than the drones with "FF" allele at both loci which suggests the direction of the epistatic effect is to low spore load. **Represents the significant level at P<0.01, ANOVA, two-tailed test, Bonferroni adjusted for the multiply comparison.

From the actual allelic segregation, 86 % of the drones with a low spore load carried at least one Danish allele at either locus of QTL_{ad} or QTL_m; 79 % of the drones that carried the Danish alleles at both loci showed a low spore load, which was a strong and significant bias compared to the random combination of two alleles towards the predicted phenotype (P<0.01, χ^2 test).

3.4. Candidate genes and expression profiles

We searched for candidate genes within the QTL_m region (between the loci HQ1414 and BI103) and QTL_{ad} (between the loci K0618 and 5388); 31 and 11 genes respectively could be identified (Table S1, Table S2). However, the gene *Aubergine* (*Aub*) within QTL_m was of particular interest. *Aub* is a member of the *Argonaut* family containing the typical active catalytic piwi domain, which functions as a nucleic acid binding domain and is involved in RNA interference (Kawaoka et al. 2008; Liao et al. 2010). The expression level of *Aub* was upregulated after *N. ceranae* infections in drones with both high and low spore load in comparison to the controls. Hence, *Nosema* infection enhanced the *Aub* expression. This effect was however stronger in the low than high spore load group. *Aub* was significantly more overexpressed in drones with a low spore load than in those with a high spore load (two-tailed t test, P<0.05) (Figure 3).

4. Discussion

4.1. One major QTL on chromosome 14

We could confirm previous results that this breeding strain shows a strong tolerance towards the *Nosema* infections (Huang et al. 2012). The genetic basis of this trait is also reflected by the coefficient of variance for spore load within the selected (c_v=0.37) and unselected population (c_v=0.43) (Huang et al. 2012),

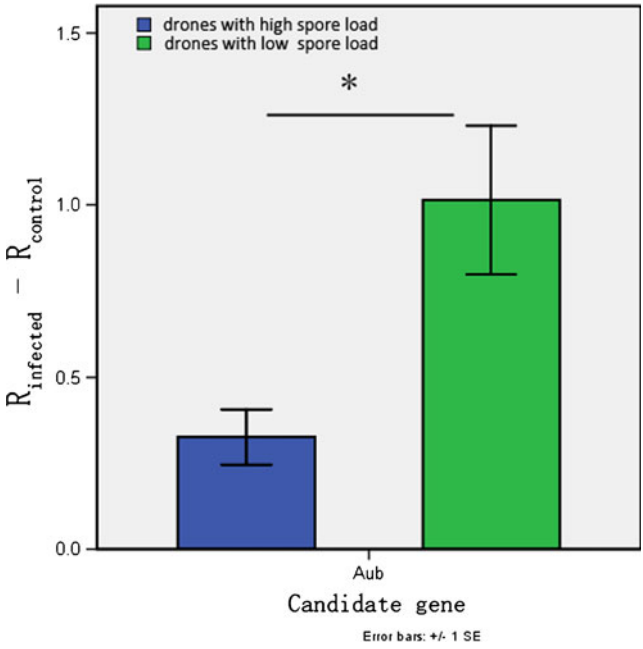

Figure 3. Candidate gene expression profiles between drones with high spore load and low spore load. The *Y*-axis is the relatively gene expression value. $R_{infected}$ and $R_{control}$ represent the relative gene expression value of the infected drones and control drones respectively. Comparing with the controls, the expression level of *Aub* was upregulated toward *N. ceranae* infection in drones with both high and low spore load. *Aub* was significantly higher expressed in drones with low spore load than drones with high spore load. *Significant level, $P<0.05$, two-tailed *t* test.

compared to the coefficient of spore load variance in the QTL mapping population (c_v=0.88) of this study. We here found four QTLs to be significantly associated with the *Nosema* spore load. With the average marker distance of ~20 cM, the power to identify a QTL explaining 7 % of the total variance exceeds 95 % (Rebai et al. 1995). Although QTLs with small effects might remain undetected, the chance is low that we have missed QTLs with larger effect than the identified significant QTLs. The QTL on chromosome 14 might play a major role (major QTL=QTL_m) for the low number of *Nosema* spores in the selected Danish honeybees, since it explained 7.7 % of the total variance. Even though the Danish allele of QTL_{ad} on chromosome 6 showed a significant association with the low spore load, this effect only occurred together with the Danish allele of the major QTL_m. Nevertheless, the combination of both loci might help to implement marker assisted breeding in the Danish population. Indeed a selective sweep could be detected in the QTL_m region (Huang et al. in revision) in the breeding population that independently confirms the importance of this locus for the tolerant phenotype.

4.2. Epistatic effects

In addition to the additive gene effects also the epistatic interactions between loci on chromosomes 3 and 10 interfered with the spore load. In this case, it was the combination of Danish and French alleles that caused a low spore load. Epistatic interactions had also been involved in *Varroa destructor* resistance (Behrens et al. 2011). In our case, the epistatic

interactions cannot be used for selective breeding because it is very likely impossible to maintain the specific allele combination at the two loci to conserve the same epistatic effects in a breeding line. Nevertheless, this example showed again the power of using drones for QTL studies as it is very likely we would have missed any epistatic effects if we had used diploid workers instead.

4.3. Candidate gene

Even though the *Aub* was the prominent candidate gene, we can of course not exclude that also other candidate genes might be involved in the tolerance towards *N. ceranae* infection. From a functional perspective, *Aub* is of particular interest, as it is involved in regulating via RNA interference machinery known to interfere with foreign RNA (Liao et al. 2010; Teo et al. 2011). In *Drosophila*, *Aub* has been shown to be involved in the silencing of retrotransposons via RNA interference (Kawaoka et al. 2008). Moreover, it was reported to be involved in the resistance towards the gut bacterium *Serratia marcescens* infection in the *Drosophila* (Cronin et al. 2009). A reduction in *Aub* transcript levels in *Drosophila* resulted in an increase of mortality due to bacterial infection. This fits to our findings of increased transcript levels in low spore load drones. Unfortunately, we do not know the exact mechanisms in which *Aub* is involved in the context of tolerance towards *N. ceranae* infections. *Aub* may trigger an apoptosis of the infected cells in order to prevent the production of spores, potentially with the help of immune genes. Indeed six innate immune genes from three different immune pathways had significantly higher expression levels in drones of the selected strain that in those of unselected strain (Huang et al. 2012). We also found the expression level of *Aub* to be higher in the selected strain, albeit not statistically significant which might be due to a low sample size (unpublished data). Nevertheless, this might reflect a general enhanced immune response of bees of the selected strain against *Nosema* infections which may have considerably contributed to the selection success of the Danish bee breeders.

Acknowledgments

We would like to thank Dieter Behrens and Petra Leibe for technical assistance; Holger Scharpenberg and Hans-Hinrich Kaatz for the assistance with bee keeping; Didier Crauser for the artificial insemination. This work was supported by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement BEE DOC FP7-KBBE-2009-3 244956CP-FP, the Deutsche Forschungsgemeinschaft Project DFG SPP 1399(MO 373/26-1) and the China Scholarship Council.

Quatre locus quantitatifs associés à une faible charge de spores de *Nosema ceranae* (Microsporidia) chez l'abeille *Apis mellifera*

***Apis mellifera* / mâle / *Nosema* / QTL**

Vier QTL die mit einer geringen Sporenbelastung von *Nosema ceranae* (Microsporidia) bei Honigbienen (*Apis mellifera*) assoziiert sind

***Apis mellifera* / Drohnen / *Nosema* / QTL**

References

Behrens, D., Huang, Q., Geßner, C., Rosenkranz, P., Frey, E., et al. (2011) Three QTL in the honey bee *Apis mellifera* L. suppress reproduction of the parasitic mite *Varroa destructor*. Ecol. Evol . doi:10.1002/ece3.17

Broman, K.W., Wu, H., Sen, S., Churchill, G.A. (2003) QTL mapping in experimental crosses. Bioinformatics **19**, 889–890

Chen, Y.P., Evans, J.D., Zhou, L., Boncristiani, H., Kimura, K., Xiao, T.G., et al. (2009) Asymmetrical coexistence of *Nosema ceranae* and *Nosema apis* in honey bees. J. Invertebr. Pathol. **101**, 204–209

Cronin, S.J.F., Nehme, N.T., Limmer, S., Liegeois, S., Pospislik, J.A., et al. (2009) Genome-wide RNAi screen identifies genes involved in intestinal pathogenic bacterial infection. Science **325**, 340–343

Forsgren, E., Fries, I. (2010) Comparative virulence of *Nosema ceranae* and *Nosema apis* in individual European honey bees. Vet. Pathol. **170**, 212–217

Fries, I., Feng, F., Silva, A., Slemenda, S.B., Pieniazek, N.J. (1996) *Nosema ceranae* n sp (Microspora, Nosematidae), morphological and molecular charac-

terization of a microsporidian parasite of the Asian honey bee *Apis cerana* (Hymenoptera, Apidae). Eur. J. Protistol. **32**, 356–365

Fries, I., Martin, R., Meana, A., García-Palencia, P., Higes, M. (2006) Natural infections of *Nosema ceranae* in European honey bees. J. Apic. Res. **45**, 230–233

Gisder, S., Hedtke, K., Möckel, N., Frielitz, M.C., Linde, A., et al. (2010) Five-year cohort study of *Nosema* spp. in Germany: does climate shape virulence and assertiveness of *Nosema ceranae*? Appl. Environ. Microbiol. **9**, 3032–3038

Hamiduzzaman, M.M., Guzman-Novoa, E., Goodwin, P.H. (2010) A multiplex PCR assay to diagnose and quantify *Nosema* infections in honey bees (*Apis mellifera*). J. Invertebr. Pathol. **105**, 151–155

Higes, M., Martin-Hernandez, R., Garrido-Bailón, E., González-Porto, A.V., García-Palencia, P., et al. (2009) Honeybee colony collapse due to *Nosema ceranae* in professional apiaries. Environ. Microbiol. Rep. **1**, 110–113

Higes, M., García-Palencia, P., Martín-hernández, R., Meana, A. (2007) Experimental infection of *Apis mellifera* honeybees with *Nosema ceranae* (Microsporidia). J. Invertebr. Pathol. **94**, 211–217

Higes, M., Martín, R., Meana, A. (2006) *Nosema ceranae*, a new microsporidian parasite in honeybees in Europe. J. Invertebr. Pathol. **92**, 93–95

Higes, M., Martin-Hernandez, R., Botias, C., Bailon, E.G., Gonzalez-Porto, A.V., et al. (2008) How natural infection by *Nosema ceranae* causes honeybee colony collapse. Environ. Microbiol. **10**, 2659–2669

Huang, Q., Kryger, P., Le Conte, Y., Moritz, R.F.A. (2012) Survival and immune response of drones of a Nosemosis tolerant honey bee strain towards *N. ceranae* infections. J. Invertebr. Pathol. **109**, 297–302

Hunt, G.J., Guzmán-Novoa, E., Fondrk Jr., M.K., Page, R.E. (1998) Quantitative trait loci for honey bee stinging behavior and body size. Genetics **148**, 1203–1213

Kawaoka, S., Minami, K., Katsuma, S., Mita, K., Shimada, T. (2008) Developmentally synchronized expression of two *Bombyx mori* Piwi subfamily genes, SIWI and BmAGO3 in germ-line cells. Biochem. Biophys. Res. Commun. **367**, 755–760

Liao, Z., Jia, Q.D., Li, F., Han, Z.J. (2010) Identification of two Piwi genes and their expression profile in honeybee, *Apis mellifera*. Arch. Insect Biochem. Physiol. **74**, 91–102

Moritz, R.F.A. (1984) The effect of different diluents on insemination success in the honeybee using mixed semen. J. Apic. Res. **23**, 164–167

Oxley, P.R., Spivak, M., Oldroyd, B.P. (2010) Six quantitative trait loci influence task thresholds for hygienic behaviour in honeybees (*Apis mellifera*). Mol. Ecol. **19**, 1452–1461

R Devolopment Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Rebai, A., Goffinet, B., Mangin, B. (1995) Comparing power of different methods for QTL detection. Biometrics **51**, 87–99

Teo, C.H., Pui, H.P., Othman, R.Y., Harikrishna, J.A. (2011) Comparative analysis of Argonaute gene sequences in bananas (*Musa* sp.) shows conserved species-specific Ago-7 PIWI domains. Genet. Resour. Crop. Evol. **58**, 713–725

Traynor, K. (2008) Bee breeding around the world. Am. Bee J. **148**, 135–139

Walsh, P.S., Metzger, D.A., Higuchi, R. (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques **10**, 506–513

Wang, S., Basten, C.J., Zeng Z.B. (2011) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm)

Zander, E. (1909) Tierische Parasiten als Krankenheitserreger bei der Biene. Münchener Bienenzeitung **31**, 196–204

Supplementary Material

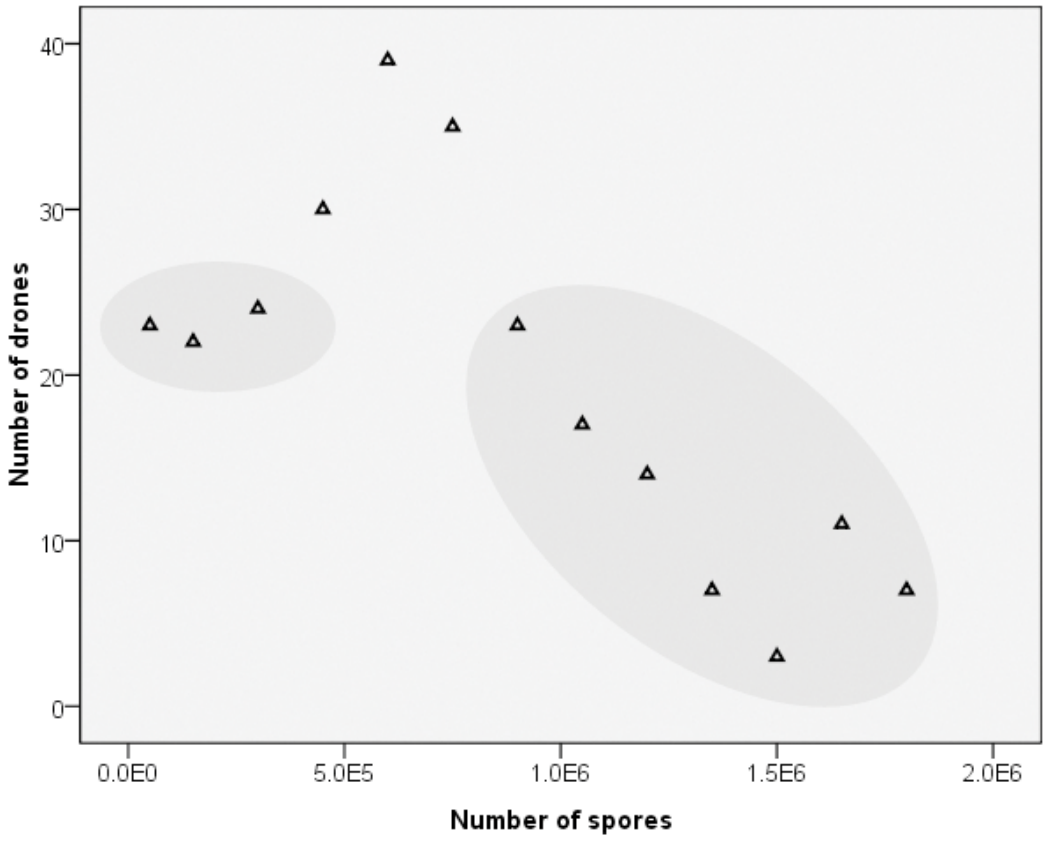

Figure S1 Spore load of the infected drones. The Y axis represents the number of the infected drones. The guts of all drones were removed on day six post infection and the spore load was counted using Fuchs-Rosenthal haemocytometer. We chose drones with 72 highest spore load (ellipse on the right side, $>8.2*10^5$ spores) and 76 lowest spore load drones (ellipse on the left side, $< 3.6*10^5$ spores) as mapping population.

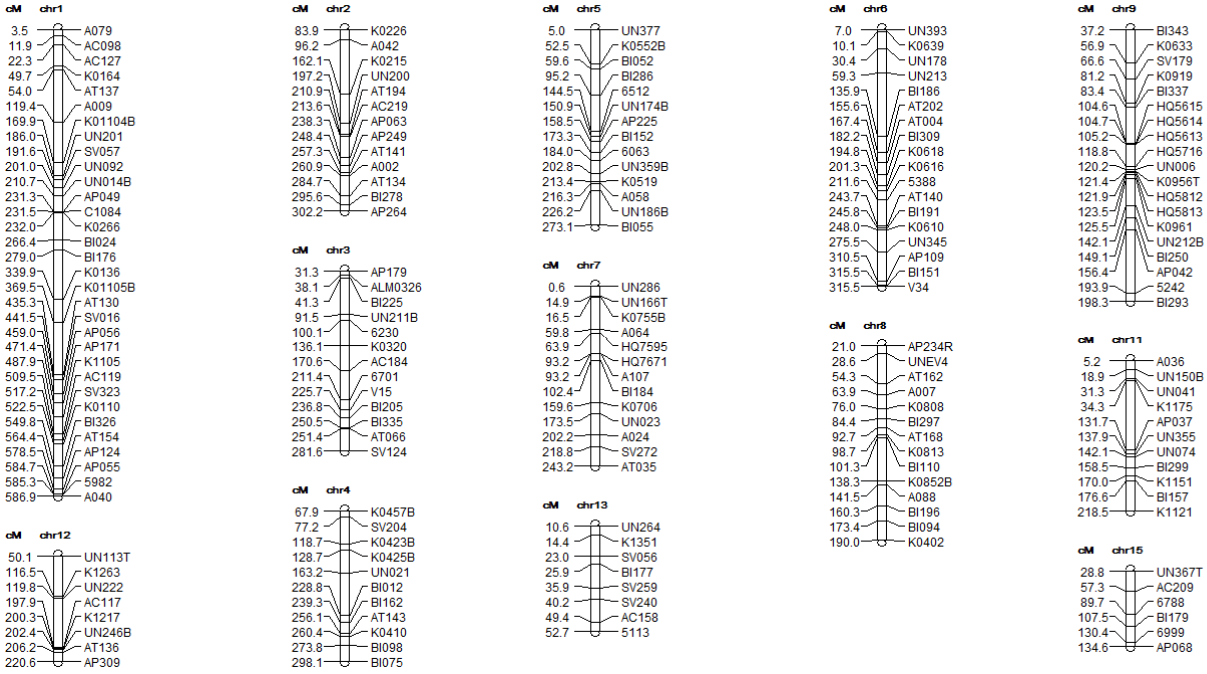

Figure S2 216 heterozygous genetic marker trees along the genome. We screened 732 fluorescence labeled microsatellite markers with the F1 queen to search heterozygous markers by multiplex PCR. The microsatellite markers' names are on the right side of each chromosome and the genetic distance (cM) is one the left side.

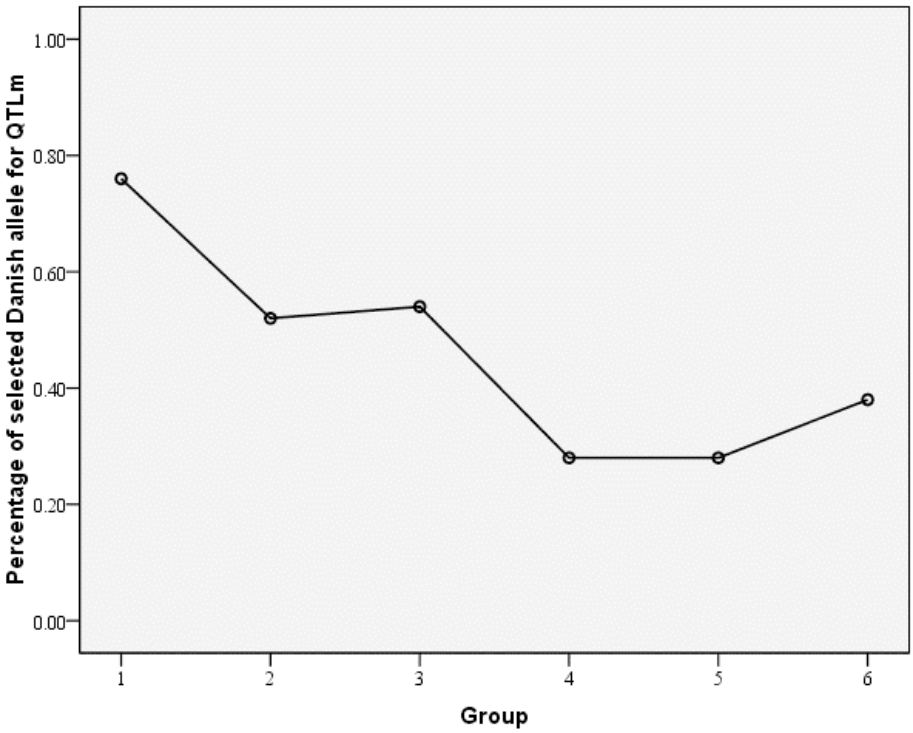

Figure S3 Percentage of the selected Danish alleles at the QTLm. The mapping population was divided into 6 groups according to the spore load from low to high. Group 1 (n=25), group 2 (n=25) and group 3 (n=22) represent three groups of drones with low spore load. Group 4 (n=25), group 5 (n=25) and group 6 (n=26) represent three groups of drones with high spore loads. Group 1 had the lowest spore load and group 6 had the highest spore load. The allele was significantly biased distributed within the drones with low and high spore load.

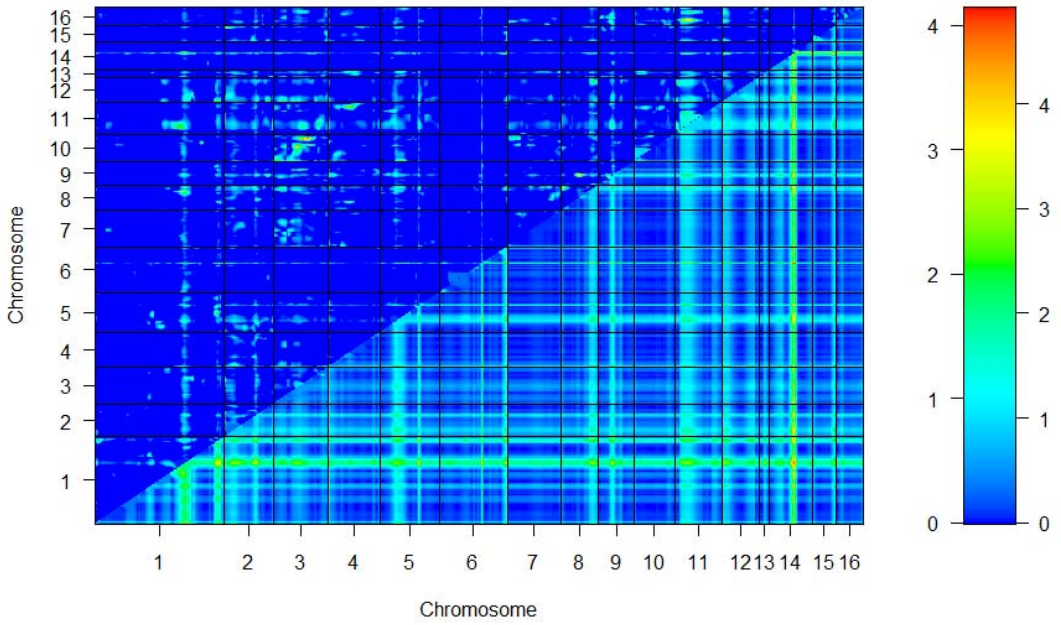

Figure S4 LOD scores along the genome for the evidence of the second QTL by two-dimensional two-QTL scan. Epistatic effect displayed in the upper left triangle. Additive effect is displayed in the lower right triangle. The numbers of left and right in the color scale on the right correspond to LOD score for epistatic effect and additive effect respectively.

Table S1 candidate genes, fly orthologs and human orghologs identified within the statistically significant major QTL region in chromosome14.

Gene name	Description / function in the honey bee	Orthologous in the fly	Orthologous in human
LOC100577208	hypothetical protein LOC100577208	Unknown	Unknown
LOC 724365	UPF0539 protein CG14977 - like	Unknown	Unknown
LOC551191	coiled-coil domain-containing protein 130 homolog	Protein structure motif; gene expression regulation	regulation of germ cells differentiation during spermatogenesis
LOC724506	heat repeat containing protein 3-like	binding	Involved in ribosome biosynthesis
Aub aubergine	RNA interference	DNA/RNA binding; expression interference	RNA-mediated gene silencing; Translation regulation
LOC 4099300	hypothetical LOC 4099300	Microtubule binding; microtubule cytoskeleton organization	assembly of kinetochore proteins, mitotic progression and chromosome segregation.
LOC100577050	Hypothetical protein LOC100577050	Zinc ion binding	regulation of spermatogenesis
LOC724688	Pallidin-like	Unknown	intracellular vesicle trafficking
O-fut2	GDP fucose protein O-fucosyltransferase 2	GDP-fucose protein O-fucosyltransferase	Carbohydrate metabolism; Fucose metabolism
LOC409902	Liprin-alpha-2-like	Sterile alpha motif domain;	cell-matrix adhesion; regulating the disassembly of focal adhesions
LOC100576983	Hypothetical protein LOC100576983	Unknown	extracellular matrix structural constituent; transferase activity, transferring acyl groups
LOC724846	Hypothetical protein LOC724846	Unknown	Unknown
LOC100577029	Hypothetical protein LOC100577029	Ionotropic glutamate receptor activity; extracellular-glutamate-gated ion channel activity	glycine catabolic process; aminomethyltransferase activity; aminomethyltransferase activity
Shrb	Transport and catabolism; Endocytosis	Protein transport	Protein transport
LOC552492	Dual specificity protein phosphatase 3-like	Protein tyrosine / serine / threonine phosphatase activity; protein tyrosine phosphatase activity; protein dephosphorylation	tyrosine-protein phosphate and serine-protein phosphate
Kinesin-6A	Kinesin motor domain; ATPase sctivity and involved in mitosis.	ATP binding; microtubule motor activity; microtubule-based movement; Kinesin, motor domain	Cell division; ATP-binding; microtubule-based movement
RpL13	Ribosomal protein; involved in translation	structural constituent of ribosome; involved in translation	SRP-dependent cotranslational protein targeting to membrane; Ribonucleoprotein; structur al constituent of ribosome
LOC552532	probable leucyl-tRNA synthetase, mitochondrial-like	leucine-tRNA ligase activity; ATP binding; leucyl-tRNA aminoacylation	leucyl-tRNA aminoacylation; tRNA aminoacylation for protein translation
LOC725055	gamma-glutamylcyclotransferase -like	gamma-glutamylcyclotransferase activity; Butirosin biosynthesis, BtrG-like; Gamma-glutamylcyclotransferase	glutathione homeostasis process; gamma-glutamylcyclotransferase activity
LOC552569	Butirosin biosynthesis,	transferase activity,	transferase activity, transferring

	BtrG-like; Gamma-glutamylcyclotransferase	transferring phosphorus-containing groups; ABC-1; Protein kinase-like domain	phosphorus-containing groups
LOC100576113	Hypothetical protein LOC100576113	kinetochore assembly; mitotic spindle organization; neurogenesis	histone acetyltransferase complex;
LOC100576082	Hypothetical protein LOC100576082	Unknown	ATP binding; protein serine/threonine kinase activity; signal transduction
LOC725200	IQ and AAA domain-containing protein 1-like	ATP binding; nucleoside-triphosphatase activity	ATP-binding
LOC551378	slit homolog 1 protein-like	Cysteine-rich flanking region; Leucine-rich repeat; Leucine-rich repeat, typical subtype	protein stabilization; enzyme regulator activity
LOC725292	homocysteine-responsive endoplasmic reticulum-resident ubiquitin-like domain member 2 protein-like	Ubiquitin; Ubiquitin supergroup	response to unfolded protein; integral to membrane
Aprt	Aprt adenine phosphoribosyltransferase	adenine phosphoribosyltransferase activity; adenine salvage; nucleoside metabolic process	Catalyzes a salvage reaction resulting in the formation of AMP; Purine salvage
LOC724206	WD repeat-containing protein 67-like	Rab GTPase activator activity; regulation of Rab GTPase activity	positive regulation of Rab GTPase activity
LOC409984	DET1- and DDB1-associated protein 1-like	Unknown	Transcription regulation; proteasomal degradation; interaction with CUL4A and WD repeat proteins
LOC724158	Hypothetical protein LOC724158	Unknown	DNA damage response, signal transduction resulting in transcription; double-strand break repair
LOC100576385	Hypothetical protein LOC100576385	Unknown	regulation of Rho protein signal transduction

Table S2 candidate genes, fly orthologs and human orghologs, identified within the statistically significant additive QTL region in chromosome 6. The QTL region is defined between the neighboring markers (K0618 at 11293Kbp and 5388 at 12262Kbp) of the statistically significant additive locus K0616.

Gene name	Description / function in the honey bee	Orthology in the fly	Orthology in human
LOC100576573	Hypothetical protein LOC100576573	Acetylation, Phohphorprotein	unknown
LOC100576598	Hypothetical protein LOC100576598	calcium ion binding, zinc ion binding	unknown
LOC551803	Spalt-like	system development; sensory organ development; neuron differentiation	inductive cell-cell signaling; olfactory bulb development; positive regulation of neuron differentiation
LOC100576253	Hypothetical protein LOC100576253	Retinitis pigmentosa GTPase regulator, nucleosomal DNA binding	unknown
LOC100576206	Hypothetical protein LOC100576206	Hemophilic cell adhesion, calcium ion binding	Importing protein into nucleu
LOC552589	ENSANGP00000002131-like	establishment or maintenance of actin cytoskeleton polarity	interacting between neurites derived from specific subsets of neurons during development
LOC409263	glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1-like	Protein modification, protein glycosylation	lycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase activity
LOC410159	mediator of RNA polymerase II transcription subunit 13-like	a coactivator involved in the regulated transcription of nearly all RNA polymerase II-dependent genes	regulation of transcription from RNA polymerase II promoter
LOC724516	Hypothetical protein LOC724516	cytokine activity	unknown
LOC100577622	Hypothetical protein LOC100577622	ATP binding, nucleotide binding	unknown
mam	neurogenic protein mastermind	methyltransferase activity, zinc ion binding	transition metal ion binding; oxidoreductase activity