Genetic mapping of male pheromone response in the European corn borer identifies candidate genes regulating neurogenisis

Fotini A. Koutroumpa\textsuperscript{a,b,1}, Astrid T. Groot\textsuperscript{a,b}, Teun Dekker\textsuperscript{c}, and David G. Heckel\textsuperscript{b}

\textsuperscript{a}Institute for Biodiversity and Ecosystems Dynamics, University of Amsterdam, 1098 XH Amsterdam, The Netherlands; \textsuperscript{b}Department of Entomology, Max Planck Institute for Chemical Ecology, D-07745 Jena, Germany; and \textsuperscript{c}Division of Chemical Ecology, Department of Plant Protection Biology, Swedish University of Agricultural Sciences, 23053 Alnarp, Sweden

Edited by John G. Hildebrand, University of Arizona, Tucson, AZ, and approved August 26, 2016 (received for review July 8, 2016)

The sexual pheromone communication system of moths is a model system for studies of the evolution of reproductive isolation. Females emit a blend of volatile components that males detect at a distance. Species differences in female pheromone composition and male response directly reinforce reproductive isolation in nature, because even slight variations in the species-specific pheromone blend are usually rejected by the male. The mechanisms by which a new pheromone signal–response system could evolve are enigmatic, because any deviation from the optimally attractive blend should be selected against. Here we investigate the genetic mechanisms enabling a switch in male response. We used a quantitative trait locus-mapping approach to identify the genetic basis of male response in the two pheromone races of the European corn borer, \textit{Ostrinia nubilalis}. Male response to a 99:1 vs. a 3:97 ratio of the E and Z isomers of the female pheromone is governed by a single, sex-linked locus. We found that the chromosomal region most tightly linked to this locus contains genes involved in neurogenesis but, in accordance with an earlier study, does not contain the odorant receptors expressed in the male antenna that detect the pheromone. This finding implies that differences in the development of neuronal pathways conveying information from the antenna, not differences in pheromone detection by the odorant receptors, are primarily responsible for the behavioral response differences among the males in this system. Comparison with other moth species reveals a previously unexplored mechanism by which male pheromone response can change in evolution.

\textit{Ostrinia nubilalis} | Z chromosome | pheromone response | sexual communication | QTL analysis

Few communication systems in the natural world can rival the sensitivity and elegance of moth sexual pheromones, which reliably bring together males and females of the same species for reproduction (1). The signal is a volatile blend of fatty-acid–derived compounds with various modifications, synthesized by the specialized pheromone gland of the female and released into the air. The detection system is the male antenna, bearing sensillar hairs with specialized neurons that express pheromone receptors that are activated by the binding of individual pheromone components. Species-specific blends activate a behavioral response in which the male flies upwind, following the odor plume, and eventually finds the female (2). Subtle differences in the pheromone compositions of closely related species are sufficient to block this attraction and to maintain reproductive isolation in nature (3, 4). The manifold variety of pheromone systems has been suggested as one of the factors promoting the evolution of the high species diversity of Lepidoptera (4, 5).

However, the origin of such variety poses an evolutionary dilemma. In such a finely tuned communication system, any deviation by the female from the optimally attractive pheromone blend or any preference by the male for an atypical blend would decrease mating success and should be selected against, maintaining the status quo (6–9). There is empirical evidence for such selection against deviance in moths (4, 10, 11). A coordinated change of both signal and response would seem necessary to overcome this problem but is hard to imagine because genes affecting female pheromone production and male response are different and reside on separate chromosomes (12–17). Moreover, despite significant advances in the biochemistry of pheromone synthesis and the physiology of pheromone detection, the identity and mode of action of most genes that shape these traits are unknown. Knowledge of the underlying genetic basis is a requirement for any theory invoking genetic mechanisms for evolutionary change.

The European corn borer (ECB), \textit{Ostrinia nubilalis} Hübner (Lepidoptera: Crambidae), is a well-studied model system of sexual communication in moths. This species consists of two strains, denoted “E” and “Z,” which use different ratios of the same two pheromone compounds, (E)-11-tetradecenyl acetate and (Z)-11-tetradecenyl acetate (henceforth “E11-14:OAc” and “Z11-14:OAc”). E-strain females produce E11-14:OAc and Z11-14:OAc in a ratio of 99:1, and Z-strain females produce these compounds in a nearly opposite ratio of 3:97 (18). E- and Z-strain males prefer the ratios produced by their respective females (19, 20). This preference leads to reproductive isolation (21), although hybrids are occasionally found in nature (22–25).

In the laboratory, hybrids are readily formed by matings; hybrid females produce an intermediate compound ratio of 65:35 (22), whereas hybrid males show a response centered around a 50:50 ratio (18, 26).

Significance

Most male moths find their mates by following species-specific pheromones released by females. Despite the importance of pheromone communication in reproductive isolation, much is still unknown about its genetic basis. We investigated male responses in the two pheromone races of the European corn borer. A reasonable hypothesis, that males of the two races differ in the genes encoding the receptor proteins that respond to the two pheromone components, was previously rejected without a convincing alternative. We found that instead male choice was correlated with genes affecting growth and differentiation of the nerve cells that may contain these receptors. This unexpected finding resolves the dilemma and points to another layer of complexity in the evolution of sexual pheromone communication systems.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

\textsuperscript{1}To whom correspondence should be addressed. Email: fotini.koutroumpa@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1610515113/-/DCSupplemental.
The genes controlling these strain differences have been mapped to different chromosomes by interstrain crosses (13). The gene causing variation in female pheromone production is autosomal and encodes a fatty-acyl reductase (27, 28). The main gene controlling the male behavioral response difference, \( \text{Resp} \), is sex-linked on the \( Z \) chromosome (13, 18) (males are ZZ and females ZW in the standard chromosomal nomenclature). The identities of the gene(s) controlling strain variation in male behavioral response in \( Ostrinia \) are still unknown.

So far, the most promising candidates for the control of male preference in moths are the genes encoding olfactory receptor proteins (ORs) that bind to specific pheromone components (29–33). In strong support of this idea, Gould et al. (14) showed that a cluster of OR genes expressed in the antenna is very tightly linked to an autosomal gene controlling the interspecific difference in male behavioral response among \( Heliothis virescens \) and \( H. subflexa \). Moreover, single-cell recordings showed that species-specific spike-amplitude responses to individual pheromone components were also linked to the OR cluster (14). In \( O. nubilalis \), five of the seven pheromone ORs are located within a large cluster on the \( Z \) chromosome (34). The pheromone ORs show similar spatial antennal expression patterns in males of both strains (35), but \( OnubOR6 \), which responds to \( Z11-14:OAc \) (36), is expressed at a higher level in \( Z \)-strain males (35), and \( OnubOR4 \), which binds to \( E11-14:OAc \) (32), is expressed at a higher level in \( E \)-strain males; expression levels of both \( OnubOR6 \) and \( OnubOR4 \) are intermediate in hybrids (35). Three \( Z \)-linked genes—\( OnubOR4, OnubOR6, \) and \( OnubOR1—\)show more sequence differences among the \( E \) and \( Z \) strains than do the two autosomal ORs; however, linkage mapping of the OR cluster on the \( Z \) chromosome placed it \( \sim 20 \) cM distant from the \( \text{Resp} \) gene (31). Although mapping of \( \text{Resp} \) was based on phenotypic data from 78 males (13), it is difficult to judge the significance of this result, because the accuracy of the behavioral scoring is unknown. Furthermore, independent replicate backcrosses using the \( E \) and \( Z \) strains by another group showed a high variability in recombination rates among \( Z \)-linked genetic markers (37).

To identify the position of the \( \text{Resp} \) locus more precisely, we conducted a new quantitative trait locus (QTL) analysis, using 470 male progeny from a series of backcrosses. We confirmed that \( \text{Resp} \) is separated from the OR cluster (31) with a more precise estimate of 15 cM. Moreover, we identified several genes that are much more tightly linked to \( \text{Resp} \) and which point to a site of action in establishing the connections between olfactory sensory neurons (OSNs) and the brain or deeper within the brain itself. This finding opens up a set of candidate genes acting at a deeper mechanistic level that could be responsible for variations in male response.

### Results

**Crossing Design.** Hybrid crosses were set up between the \( Z \) and \( E \) strains, and resulting male progeny were backcrossed to \( Z \)-strain females for seven consecutive generations. In each generation we scored behavioral responses to \( Z \), \( E \), and hybrid pheromone blends using artificial lures for \( 50 \)–\( 150 \) males (fewer for the third and fourth backcrossed generations). In each generation, one of the scored males with a hybrid phenotype was chosen for backcrossing to a \( Z \)-strain female to produce the next generation. In total, 649 males were phenotyped, using the wind tunnel setup described in Methods. After the last generation of backcrossing, DNA was isolated from all 649 phenotyped males, and genotypes were scored at nine sex-linked markers (\( \text{kettin} \), five ORs, \( \text{ldh} \), \( \text{bgi012356} \), and \( \text{bgi03892} \)). This preliminary scoring showed a loss of genetic polymorphism correlated with behavioral response in the last two backcross generations, probably because of the inadvertent selection of a homozygous male for continuation of the backcrossing. Therefore, only 470 males from the first five generations could be used for final mapping of 21 sex-linked genes and the QTL analysis. We also genotyped 143 females (~20 per generation) to aid in map construction.

**Behavioral Measurements.** The relative intensity of individual male behavioral responses to each of the pheromone blends—\( Z \), hybrid (\( H \)), and \( E \)—was summarized using a quantitative score that assigned numerical weights to two key behaviors previously shown to be correlated with directed flight toward the pheromone source (38). These behaviors are wing-fanning, a rapid vibration of the wings while remaining in a stationary position, and extrusion of the hair pencil scales at the tip of the abdomen, which release the male pheromone during courtship at close range (Fig. 1). Wing-fanning could be strong and continuous or weak and intermittent (on-off), and hair pencil extrusion could be present or absent. The latency time, i.e., the number of seconds between presentation of the stimulus and the initiation of wing-fanning, was also factored in. An initial set of numerical weights for these behaviors was first chosen arbitrarily, such that the scored response to a given pheromone blend increased with the occurrence and intensity of wing-fanning, increased more with a shorter latency time, and increased even more with the extrusion of hair pencil scales. Results from the three lures were combined by subtracting the \( H \) and \( E \) lure responses from the \( Z \) lure response, so that \( Z \)-strain males would have the highest score, because these are male-informative backcross progeny from a pure \( Z \)-strain mother. Phenotypic scores then were calculated for each male (see Fig. S1A for details). After an initial QTL analysis, the parameter set of numerical weights was optimized using a discriminant function approach, to discriminate more precisely between the responses of \( \text{Resp} \) locus homozygotes versus \( \text{Resp} \) locus heterozygotes (Fig. S1B).

**Linkage Maps.** To examine the contribution of each chromosome to male behavioral response, a preliminary map using amplified fragment length polymorphism (AFLP) markers covering all 31 chromosomes was constructed, using data from the 120 progeny of the first backcross generation (Fig. S2). We then mapped an additional 21 genes on the \( Z \) chromosome using the progeny of all backcrosses. These genes were chosen based on their \( Z \)-chromosomal location in \( B. mori \), and their orthologs in \( O. nubilalis \) were PCR-amplified using degenerate primers designed from multispecies sequence alignments (Table S1). Overall, the gene order in \( O. nubilalis \) between \( \text{kettin} \) at one end and \( \text{paraplegin} \) near the other end was the same as in \( B. mori \). However, we did not map any genes surrounding \( \text{tpi} \), where Wadsworth et al. (39) have detected a region of low recombination that probably corresponds to an inversion. Additionally, the order of the two end markers \( \text{pdp1} \) and \( \text{bgi03892} \) was inverted (Fig. 2). An inversion between \( \text{bgi03892} \) and another marker was also detected in \( O. nubilalis \) by Kroemer et al. (Fig. 2) (37). Finally, we found that the \( Z \)-linked marker
bg002071 in *B. mori* was autosomal in *O. nubilalis*, as was also shown for marker G23 just above kettin by Kroemer et al. (37). Thus, the segment of the Z chromosome above kettin in *B. mori* appears to be autosomal in *O. nubilalis* and therefore segregates independently from the behavioral response trait.

**QTL Analyses.** A QTL analysis using 198 AFLP markers that were mapped to the 31 chromosomes spread over 45 linkage groups showed that the Z chromosome was the only one with a significant effect on the male behavioral response. (Fig. S2). This result is in agreement with the findings of Dopman et al. (13), who used a qualitative flight assay to map the major Resp gene to the Z chromosome in a backcross to the E strain. Our quantitative measurement of male behavior additionally permits an estimate of the contribution of minor genes. In comparison with the Z chromosome with a significant logarithm of odds (LOD) score of 5.0, explaining 25% of the variance in behavioral score, the chromosome with the next largest effect is AFLP linkage group 31, explaining 11% of the variance, but this effect was below the *P < 0.05* significance threshold (Fig. S2). No other linkage group has an LOD score higher than 1.5 (Fig. S2).

We then conducted a more extensive QTL analysis using 12 genes on the Z chromosome, using data from all 470 phenotyped and genotyped males. Interval mapping produced a single broad peak of the log-likelihood function at position 84.2 cM with an LOD score of 8.7 covering an interval of 36 cM, containing one or more QTLs for the behavioral response. The OR cluster occurred near one edge of this broad interval (Fig. S3A). We used this large interval to optimize the behavioral scoring function (SI Materials and Methods), which increased the LOD score at the peak to 27.8 cM and narrowed its width to 12 cM (Fig. S3B). Finally, for the final QTL analysis we added genotype scores for an additional 12 markers, 6 of which were located within the 12-cM interval (Fig. 3).

The resulting log-likelihood function showed a single narrow peak at position 85.3 cM near the marker kon-tiki with an LOD score of 29.6. The 2-LOD confidence interval in which the LOD score exceeds 27.6 is 8 cM wide and covers positions 83 cM to 91 cM. The likelihood that Resp occurs at the peak is more than 100 times greater than the likelihood that it occurs outside the 2-LOD confidence interval. The log-likelihood function also showed a second peak at 108 cM with a maximum LOD score of 20.5, flanked by two markers (paraplegin and pdp1). This peak is approximately 107 times less likely to be the Resp locus than any position within the 2-LOD confidence interval, but it could represent a second QTL with a smaller effect on the behavioral score. To examine this possibility, we conducted fixed-effects two-way ANOVA using genotypes at the markers nearest the major and minor peaks as factors (Fig. S4). The major peak effect was highly significant (*P < 0.0001*); the minor peak effect and interaction effects were not (*P = 0.46* and *P = 0.37*, respectively for the marker to the left of the minor peak and *P = 0.49* and *P = 0.18* for the marker to the right of the minor peak), indicating that any possible contribution of the minor peak effect and interaction effect to the overall variance of the behavioral score is less than 10% of the contribution of the major QTL.

**Evaluation of Candidate Genes for the Behavioral Response.** We next evaluated the OR cluster as a candidate for the Resp locus. The cluster maps to position 72 cM; in scoring OnubOR5, OnubOR4,
and OnubOR6, we did not detect any crossovers within the cluster (Figs. 2 and 3). The OR genes occur at 85.3 cM, 13.3 cM away from the larger QTL peak, somewhat closer than the distance of 20 cM found by Lassance et al. (31) between the ORs and the Resp locus. The OR cluster is even further away (36 cM) from the smaller peak at 108 cM. The LOD score of the OR cluster is 20.5, making it very unlikely to be the Resp locus. The peak at 85.3 cM is 10^4 times more likely, and any location within the 2-LOD confidence interval is at least 10^2 times more likely to be the Resp locus than is the OR cluster. The slight hump in the LOD score at the OR cluster (Fig. 3) might indicate a smaller contribution to the behavioral response. To test the possibility that the ORs represent a third, minor QTL, we conducted a two-way ANOVA using the marker genotypes at the QTL peak and the ORs as factors (Fig. S5). As before, the main effect of the peak was highly significant (P < 0.0001), but neither the main effect of the OR genotype (P = 0.84) nor the interaction effect (P = 0.13) was significant. Moreover, the nonsignificant effect of the OR genotype depended on the genotype of the markers near the peak. Thus, sequence variation at the OR cluster makes no consistent contribution to male pheromone preference.

The 2-LOD confidence interval around the major peak contains two of the marker genes, kon-tiki and bric-a-brac (bab). Comparing our O. nubilalis map with the B. mori and Danaus plexippus genomes, a total of seven genes have been mapped between ldh and bg012356 (Fig. 3). Interestingly, six of these genes—highwire, pentraxin, trol, kon-tiki, bab, and CCR4-not—seem to be involved in neurogenesis (Table 1). These genes seem to control neuronal axonal growth and orientation processes specifically in insects and other animals (see references in Table 1). Furthermore, the genome comparisons to identify other candidates for Resp showed a total of 12 genes (including kon-tiki and bab) located within the Resp region. Eight have one or more identifiable protein domains, six have homologs in Drosophila with some functional information, and three are not similar to any protein of known function. Although none has an obvious connection to pheromone sensory physiology, several are implicated in neurogenesis; we discuss these possible roles below.

**Discussion**

By fine-scale mapping of the male response in O. nubilalis, we found candidate genes that had not been implicated in previous studies of male pheromone perception in moths. Moreover, confirming previous work (31), we ruled out variation in other candidate genes, i.e., the ORs that detect pheromone components and have been implicated in studies of other species. We will explain the justification for our approach in analyzing behavior, point out some aspects of our mapping methods, reconsider the role of the pheromone receptors, and discuss the possible roles of genes in the chromosomal region we have identified.

In mapping the genes underlying the male behavioral response, we deliberately chose to measure traits different from but correlated with the final, evolutionarily relevant outcome: a mating with a female of the same strain. This outcome has traditionally been measured by observing male behavior and flight toward a female or an artificial pheromone dispenser at the upwind end of a wind tunnel. The ultimate criterion of response is considered to be touching and attempting to copulate with the pheromone source, and this criterion has been used as a qualitative, dichotomous trait in previous mapping studies (13, 18). This approach has been justified by the observation that E-strain males may exhibit earlier signs of attention to the Z lure and the opposite applies to Z-strain males; such signs are unreliable predictors of the final outcome (38). However, recent studies have shown that early responses by Z-strain males can lock them in to subsequent behaviors that are maintained even if the pheromone lure is experimentally switched from Z to E while the moth is in midflight (40). Thus, variation in the early responses also may be a relevant indicator of genetic differences between the Z and E strains. Therefore, we scored male response as a quantitative trait by observing the preflight responses of males to different pheromone blends. This reduction in individual assay time greatly increased the number of males that could be scored against all three pheromone blends over the generations. Even though 68% of the variance in this quantitative trait was nongenetic or was caused by minor genetic factors, QTL analysis could define an interval on the Z chromosome that accounted for 32% of the variance. Our methods enabled the quantitative assessment of previously suspected candidate genes and identified candidates that had not previously been considered; this assessment would have not been possible if the response had been scored as a dichotomous trait. Moreover, our method avoids inaccuracies in fine-scale genetic mapping caused by errors in scoring male response as a dichotomous trait, which are inevitably confounded with recombination (see the discussion of mapping strategies in SI Materials and Methods).

The exclusion of the cluster of ORs from the QTL region means that sequence variation in those genes, whether in the coding sequence or in noncoding regions that affect their expression patterns, cannot be directly responsible for the difference in male response. However the possibility remains that differences in OR expression or function may be controlled by a trans-acting transcription factor located within the QTL region. We previously found no qualitative difference in OR expression: Both strains have the same ORs, localized at the same positions on the antenna, with the same number of sensilla and neurons housing them (35). However, we did find quantitative differences in the expression of OnubOR4 and OnubOR6 (35). These differences could be caused by a trans-acting transcription factor or could be affected by changes in neuron size and number. Nevertheless, the genetic factor underlying male response in the Ostrinia pheromone strains must be different from that of H. virescens and H. subflexa, in which the male behavioral
<table>
<thead>
<tr>
<th>Resp region (abbreviation)</th>
<th>Full name (abbreviation)</th>
<th>Homologs/synonyms</th>
<th>Accession number</th>
<th>Organism</th>
<th>Function</th>
<th>Expression</th>
<th>Ref. or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out Highwire (hiw)*</td>
<td>PHR protein family:</td>
<td>AF262977</td>
<td>*</td>
<td>Ubiquitin ligase, negative regulator of synaptic growth</td>
<td>Muscle neurons; CNS; retinal cone</td>
<td>(75–78)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Myc (PAM)†</td>
<td></td>
<td></td>
<td>Presynaptically; synaptic growth and axon guidance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rpm-1 gene‡</td>
<td></td>
<td></td>
<td>Postsynaptically: endocytosis of glutamate receptors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Esrom§</td>
<td></td>
<td></td>
<td>Innate immunity</td>
<td>Sensory neurons</td>
<td>(79)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lectins</td>
<td></td>
<td></td>
<td>Synapse formation and remodeling, neural crest cell migration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U18772</td>
<td>§</td>
<td>Signal transduction</td>
<td>Regulates neuromuscular junction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out terribly reduced optic lobes(tro)*</td>
<td>Heparan sulfate Proteoglycan§</td>
<td>XP_013464776.1</td>
<td>¶</td>
<td>Activates neuroblasts’ adhesion, growth migration and differentiation</td>
<td>Larval brain, imaginal discs, fat body, muscles; adult gonads</td>
<td>(43, 80, 81)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Galactose mutarotase†</td>
<td></td>
<td></td>
<td>Uncharacterized protein</td>
<td>unknown</td>
<td>GenBank</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KGM_08728§</td>
<td></td>
<td></td>
<td>Cellular retinaldehyde binding protein</td>
<td>Fat body; gut: hemocytes</td>
<td>(B3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KGM_088417</td>
<td></td>
<td></td>
<td>α-Tocopherol transfer protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In Aldose-1 epimerase*</td>
<td>Perlecan1,2</td>
<td></td>
<td></td>
<td>Target-derived signal initiating stable connection</td>
<td>Imaginal tissues: eye; imaginal disk: photoreceptor cell</td>
<td>(56, 57)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Galactose mutarotase‡</td>
<td></td>
<td></td>
<td>Uncharacterized protein</td>
<td>Uncharacterized protein</td>
<td>GenBank</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KGM_07611§</td>
<td></td>
<td></td>
<td>Chondroitin sulfate proteoglycan**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perdido*</td>
<td></td>
<td></td>
<td>Cell migration and differentiation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NG2/MCSP1,2</td>
<td></td>
<td></td>
<td>NG2/MCSP1,2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LOC101738019§</td>
<td></td>
<td></td>
<td>Uncharacterized protein</td>
<td>Testis</td>
<td>GenBank</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LOC101738910†</td>
<td></td>
<td></td>
<td>Uncharacterized protein</td>
<td>Testis</td>
<td>GenBank</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KGM_088417</td>
<td></td>
<td></td>
<td>α-Tocopherol transfer protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In Crabp*</td>
<td>RING finger protein 17</td>
<td>XP_013140855.1</td>
<td>††</td>
<td>Germine development</td>
<td>Myotube tips: Embryonic founder; Myoblats: CNS</td>
<td>(51, 52)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SANT domain</td>
<td>XP_004933184.1</td>
<td>#</td>
<td>Chromatin modification</td>
<td></td>
<td>(54)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>_transcription factor</td>
<td>XP_013184855.1</td>
<td>**</td>
<td>DNA binding</td>
<td></td>
<td>(55)</td>
<td></td>
</tr>
<tr>
<td>In Lim homebox</td>
<td></td>
<td></td>
<td></td>
<td>Neuronal patterning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Uncharacterized protein</td>
<td>Testis</td>
<td>GenBank</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Uncharacterized protein</td>
<td>Testis</td>
<td>GenBank</td>
<td></td>
</tr>
<tr>
<td>In bgi12353A</td>
<td>LOC101738771†</td>
<td>EHJ68261.1</td>
<td></td>
<td>Cyclin binding</td>
<td>Imaginal tissues: eye; imaginal disk: photoreceptor cell</td>
<td>(56, 57)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LOC101738910†</td>
<td>XP_013184860.1</td>
<td>**</td>
<td>Ubiquitin–protein transferase activity: negative regulator of cell growth, including axon guidance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In Archipelago (ago)</td>
<td>F-box protein</td>
<td>XP_013184917.1</td>
<td>**</td>
<td>Sex differentiation, i.e., abdominal color sexual dimorphism in D. melanogaster</td>
<td>Female gonad development</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cyclin binding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In Bric à brac (bab)*</td>
<td>BTB POZ domain</td>
<td>XP_013464829.1</td>
<td>¶</td>
<td>Leg and antenna segmentation</td>
<td>Leg and antenna developmental and imaginal disk: ovaries</td>
<td>(58–60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Eye–antennal disk morphogenesis and imaginal disk-derived leg morphogenesis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sex differentiation, i.e., abdominal color sexual dimorphism in D. melanogaster</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out CCR4-NOT*</td>
<td></td>
<td>XP_013184897.1</td>
<td>**</td>
<td>Ubiquitin ligase, negative regulator of synaptic growth</td>
<td>Regulation of synaptic growth of the neuromuscular junction</td>
<td>(61, 62)</td>
<td></td>
</tr>
</tbody>
</table>

* Drosophila melanogaster.
† Homo sapiens.
‡ Caenorhabditis elegans.
§ Mus musculus.
¶ Papilio machaon.
∥ Bombyx mori.
∥ Amyelois transitella.
** Papilio polytes.
†† Papilio polytes.

Koutroumpa et al. | PNAS | Published online October 3, 2016 | E6405
responses and most of the neurophysiological responses to pheromone components cosegregated with the autosomal OR cluster (14).

The candidate genes in the Resp QTL region (Table 1) suggest that other factors affect the structure, function, or connectivity of the OSNs or other neurons in the antennal lobe or elsewhere in the brain. Three candidates are similar to genes of unknown function in the sequenced genomes of other lepidopteran species (Table 1). Two other candidates could potentially play a role in gene regulation: One has a Tudor domain also found in proteins that bind to methylated histones (41), and the other has a SANT (switching-defective protein 3, adaptor 2, nuclear receptor co-repressor, and transcription factor IIIb) domain involved in histone modification (42). Two candidates, aldoase epimerase and a lipid-binding α-tocopherol transfer protein, have domains not necessarily specific to the nervous system. Four genes within the QTL region and two just outside it are similar to genes with roles in neurogenesis. Of the latter, trol (ternibly reduced optic lobes, also called perlecan) is a large multidomain heparan sulfate proteoglycan in the extracellular matrix with Laminin G domains that binds to and stores other signaling molecules controlling neuroblast proliferation (43). On the other side of the QTL region is CCR4-Not, coding for the catalytic subunit of the deadenylase complex (44, 45), which interacts with the conserved P body component HPatPat1 to regulate synaptic terminal growth in the Drosophila neuromuscular junction (46).

We consider that four genes mapping within the QTL region are most likely to play a relevant role in neurogenesis. In Drosophila melanogaster, kon-tiki encodes a transmembrane protein that promotes cell migration in muscle development (47, 48). Similar mechanisms have been proposed to generate directed migration and target recognition of myotubes as well as neuronal axon and dendrite growth toward their synaptic partners (49). The second likely candidate is a gene with LIM and homeobox domains characteristic of many transcription factors that has also been described as a neural-patternning gene in several organisms [i.e., Drosophila, mice, and Caenorhabditis elegans (50)]. The third likely candidate is archipelago (ago), which encodes an F-box protein and part of a ubiquitin ligase complex that interacts with the Notch signaling pathway and suppresses tissue growth in flies and tumor development in vertebrates (51, 52). Mutations in the fourth likely candidate, bab, have many concordant behavioral, disordered egg-laying patterns of eggs on ovaries in D. melanogaster (53). In both antennae and tarsi, bab causes distal segment fusion (54). The protein encoded by bab possesses a BTB domain also found in the Drosophila genes tramtrack and Broad-Complex (55). Moreover, bab was recently found to interact with other transcription factors in a network that patterns the developing olfactory tissue in D. melanogaster (56).

In addition to differing in their behavioral response to female pheromones, E- and Z-strain males also differ in the connections between the antennal OSNs and the antennal lobe of the brain. In the Z strain, the OSNs that respond to Z11-14:OAc project their axons onto the large, medial glomerulus of the antennal lobe, and the OSNs responding to E11-14:OAc project onto the smaller, lateral glomerulus (57). In the E strain, these connections are reversed (57). In F1 hybrids, the connections are similar to those in the E strain, but the lateral and medial glomeruli are more similar in size, and glomerular size appears to be Z-linked (58). A Z-linked mutant in B. mori shows a similar rewiring of neuronal connections. In wild-type B. mori males, OSNs expressing the bombykol receptor BmOR1 target their axons to the larger toroid glomerulus, whereas OSNs expressing the bombykalc receptor BmOR3 send their axons to the smaller cumulative (59, 60). In null mutants of the Z-linked acj6 gene, BmOR1-expressing OSNs are rare or absent, and BmOR3-expressing OSNs project instead to the toroid (61). Mutant males are attracted not by the main pheromone component bombykalc but instead by the minor component bombykal. We have confirmed that acj6 is also Z-linked in O. nubilalis, but it segregates independently from the QTL interval. Furthermore, acj6 maps distal to kettin on the Z chromosome of B. mori, far from the QTL interval (Fig. 2). Thus, despite intriguing similarities in the phenotype caused by mutation of acj6 in Bombus, acj6 does not seem to be responsible for male response in O. nubilalis.

Nevertheless, it is possible that one of the genes in the QTL interval could have a similar effect on axonal targeting. In Drosophila (unlike mammals) the expression of a given OR in a sensory neuron is not directly involved in glomerular targeting (62). Instead, upstream-acting transcription factors involved in neuronal differentiation separately specify both the downstream expression of a particular OR (63) and the axonal guidance of that neuron to its cognate glomerulus in the brain. In O. nubilalis, both strains possess the same cluster of ORs for pheromone detection. One of the newly identified candidate genes could control the expression of specific ORs in different OSNs [as inferred by Koutroupa et al. (35)] or the targeting of these OSNs to the antennal lobe, or both processes. A third possibility is a change in the antennal lobe itself, governing which OSNs will connect to which glomeruli. Such a change is suggested by the results of transplanting antennae of E-strain males onto Z-strain males: The chimeric males responded only to the Z blend (64). In this example, the Z genotype of the recipient bearing the antennal lobe had a greater influence than the E genotype of the OSNs in the transplanted antennae.

One other extensively studied trait is the spike amplitude of OSNs responding to the E or Z component in the two strains. In E-strain males, the E-responding OSN has bigger spike amplitude than the Z-responding OSN. In Z-strain males, the magnitudes of the spike amplitudes are reversed, and in hybrids they are more nearly equal. Spike amplitude is correlated with the diameter of the dendrite projecting into the sensillar lymph (65). Early studies found that spike amplitude was autosomally inherited (66), not sex-linked (18), and was not on the autosomal controlling female pheromone production (12). Later studies found that spike amplitude was not correlated with male behavioral response (67) but also may have a minor Z-linked component (15). There is a strain difference in electroantennogram responses that also appears to be Z-linked (58). Further discussion on spike amplitude measurements and the interpretation of these results in the light of sex-linkage can be found in Koutroupa et al. (35). We did not score spike amplitudes or whole-antennal responses in this study because of the prohibitive workload of neurophysiological preparations of each male after the behavioral assays, but further mapping of these traits would be worthwhile.

In conclusion, our QTL-mapping approach has revealed several genes involved in neurogenesis that may account for the differences in male behavioral response between pheromone strains in O. nubilalis. In addition to the previously established importance of changes in the ORs, as supported by phylogenetic analysis, functional expression, and genetics, other changes can shift preference without any change in the ORs themselves. Further advances will result from the application of QTL mapping to other recently diverged species pairs or species with pheromone races, as well as from functional studies of the candidate genes we have identified.

Materials and Methods

Insects. Laboratory colonies of ECB Z and E strains were used. The Z-strain colony derived from cornfield-collected adults in Kéty town, county of Tolna, Hungary in 2004. The E-strain colony was established from larvae extracted from maize stems collected by Magda Rak-Cizej of the Agriculture and Forestry Institute, Novo Mesto, Slovenia. The purity of the strains was monitored by GC analysis of female pheromone production following the protocol of Kárpáti et al. (68), and male response was evaluated with an electroantennogram (58). All insects were reared on a semiartificial diet (69).
Until pupation. Adults were fed a 5% honey–water solution throughout their adult lives. All animals were kept at 25 °C, relative humidity 70% under an 18-h:6-h light-dark photoperiod. The day of emergence was considered day 0.

**Backcrosses.** All crosses were single-pair crosses. Because there is crossing-over in male Lepidoptera (70), we conducted male-informative backcrosses to generate recombinant in the Z chromosome. Single-pair matings were set up hybridizing pure Z-strain females with E-strain males (ZE hybrid) and E-strain females with Z-strain males (EZ hybrid). The F1 hybrid males were backcrossed to pure Z-strain females. We selected one out of 30 hybrid families, namely family E22, because it was highly fecund. Sons of the E22 family that showed hybrid behavior in the wind tunnel bioassays were backcrossed with pure Z-strain females, giving rise to the first backcross generation (BC1: E22Z). Two BC1 families were continued to second-generation backcrosses, i.e., E22Z-4 and E22Z-7 families. Because only the E22Z-7 family generated enough offspring, we continued the backcross procedure with this family for six additional generations (BC2–BC7). One family for each of the subsequent BC2–BC7 families was phenotyped and genotyped.

**Winter Tunnel Phenotyping.** All 649 backcross males were tested separately in the wind tunnel for response to each of the Z, E, and H blends. Ourwind-tunnel assays were based on a series of behavioral male responses to each of the three pheromone blends: resting, wing fanning, and hair pencil extrusion. Each category was given a score (Fig. 1 and Fig. S1), so that we had a weighted and quantitative measurement for each male.

The males were 1–2 d old and were sexually and olfactorily virgin. All males were behaviorally tested at the second hour of their scotophase, with 79°C air temperatures and natural air humidity of 19 °C. Each male was kept in a cylinder 2 cm in diameter with gauze on both ends so that the air plume with the pheromone could pass through the tube. With this setup we were able to keep the males captured during the assay, avoiding the need to recapture them (and possibly damage them) at the end of the experiment.

We used the same males for subsequent matings.

The three pheromone blends were ordered from Pherobank and were separately presented to each of the males for 1 min maximally. During one experimental day, all males were tested first with the Z blend, followed successively by the H and E blends, with an interval of 30 min to ventilate the wind tunnel between the tests. After a male showed the hair pencil extrusion response or after 1 min, the male was removed from the wind tunnel. The latency of each reaction to the blend was recorded. If the behavioral response was ambiguous, the male was retested on the following day.

**Mapping and Genotyping.** We mapped the chromosomes using AFLP markers, using a protocol adapted from ref. 71. On polyacrylamide gels, the markers scored were present in the pure E-strain female, the F1 hybrid male, and the heterozygous backcross males and were absent in the pure Z-strain male, the F1 Z-strain female, and in the homozygous backcross offspring. The opposite pattern was scored as well. We scored 198 AFLP markers from 47 primer combinations in 123 samples (parents and F1 and BC1 males and females). Using Mapmaker 3.0 (72) we grouped 180 of these males into 45 linkage groups, with at least two markers from different primer pairs. The number of markers for the 35 linkage groups varied from 2 to 11; the average number was 5.14. For fine-scale mapping on the Z chromosome, we used intron-size polymorphisms or SNP variation in 21 Z-linked genes that were scored in all backcross males; then recombination rates were determined in Mapmaker as well. We verified that all polymorphic differences used in the mapping also differed among wild-type Z and E males in the laboratory strains.

**QTL Analysis.** With the phenotypic scores for all male behaviors, we first determined which linkage groups explained a significant portion of the variation, using R-studio using t test-based marker regression as implemented in RQTL (version 0.98.490). We established significance thresholds for LOD scores empirically by permutation tests using 10,000 permutations. After finding a significant LOD score for only the Z chromosome, we conducted our analysis with the Z chromosome and females. Using Mapmaker, we received further funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement 661322.

**ACKNOWLEDGMENTS.** The behavioral investigations were supported by Grant Vetenskaparapäät (The Swedish Research Council), Project 621-2014-4816 (to T.D.) and by a Linnaeus Chemical Ecology, Ethology, and Evolution Initiative (ICE3) grant [The Swedish Research Council for Sustainable Development (FORMAS)/Swedish University of Agricultural Sciences] (to F.A.K. and T.D.). The molecular and mapping analysis was supported by the Max Planck Institute for Chemical Ecology (D.G.H., A.T.G., and F.A.K.). F.A.K. received further funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement 661322.


