Functional consequences of sequence variation in the pheromone biosynthetic gene *pgFAR* for *Ostrinia* moths

Jean-Marc Lassance[a,1,2], Marjorie A. Liénard[a,1], Binu Antony[a,3], Shuguang Qian[b], Takeshi Fujii[b], Jun Tabata[c], Yukio Ishikawa[b], and Christer Löfstedt[a]

[a]Department of Biology, Lund University, SE-22362 Lund, Sweden; [b]Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo 113-8657, Japan; and [c]Biodiversity Division, National Institute for Agro-Environmental Sciences, Tsukuba 305-8604, Japan

Pheromones are central to the mating systems of a wide range of organisms, and reproductive isolation between closely related species is often achieved by subtle differences in pheromone composition. In insects and moths in particular, the use of structurally similar components in different blend ratios is usually sufficient to impede gene flow between taxa. To date, the genetic changes associated with variation and divergence in pheromone signals remain largely unknown. Using the emerging model system *Ostrinia*, we show the functional consequences of mutations in the protein-coding region of the pheromone biosynthetic fatty-acyl reductase gene *pgFAR*. Heterologous expression confirmed that *pgFAR* orthologs encode enzymes exhibiting different substrate specificities that are the direct consequences of extensive nonsynonymous substitutions. When taking natural ratios of pheromone precursors into account, our data reveal that pgFAR substrate preference provides a good explanation of how species-specific ratios of pheromone components are obtained among *Ostrinia* species. Moreover, our data indicate that positive selection may have promoted the observed accumulation of nonsynonymous amino acid substitutions. Site-directed mutagenesis experiments substantiate the idea that amino acid polymorphisms underlie subtle or drastic changes in pgFAR substrate preference. Altogether, this study identifies the reduction step as a potential source of variation in pheromone signals in the moth genus *Ostrinia* and suggests that selection acting on particular mutations provides a mechanism allowing pheromone reductases to evolve new functional properties that may contribute to variation in the composition of pheromone signals.

chemical communication | Lepidoptera | speciation | genetics

Chemical signals such as pheromones play a central role in mediating the reproductive behavior of many animals (1). Also, population divergence often involves changes in the sex pheromone signals and/or their perception that cause modification of the associated behaviors and, subsequently, prezygotic reproductive isolation (1–3). Although some progress has been made in identifying the genetic architecture underlying pheromone communication systems (3, 4), the molecular bases of intraspecific variation or transitions in pheromone systems remain poorly documented. To unravel general patterns concerning the cause of divergence (5), it is fundamental to move beyond studies of individual species. Indeed, studies on intra- and interspecific differences among isolated taxa have the potential to reveal the causes of signal divergence that are associated with reproductive isolation (3).

In this context, moths (Insecta: Lepidoptera), which represent one of the largest group of insects with ~160,000 described species (6), provide highly relevant examples because subtle changes in their sex pheromone composition are often the initial trigger for population divergence and can lead to speciation (7, 8). Typically, female moths emit a sex pheromone that attracts males over long distances. Moth sex pheromones usually consist of a bouquet of structurally similar chemical components, and species specificity is in many cases achieved by the use of a narrow range of pheromone components combined in a specific ratio (1, 7, 8). The moth genus *Ostrinia* is an emerging model system ideal for addressing questions on the genetic bases of pheromone evolution and identifying genes that shape pheromone signals (9–12). The *Ostrinia* female sex pheromones are typically blends of monounsaturated tetradecenyl (C14) acetates that vary in double bond position (Δ^9, Δ^{11}, or Δ^{12}) and geometry [*cis* (*Z*) or *trans* (*E*); Fig. 1]. Like in many moths, specificity of the signals results from the makeup and proportion of pheromone components. A number of field and flight tunnel studies conducted using the European and Asian corn borer *O. nubilalis* and *O. furnacalis*, respectively, have demonstrated that males respond to a narrow range of pheromone ratios and that small changes in the pheromone ratio have a clear behavioral impact (13–16). The *Ostrinia* system has received much attention because of naturally occurring pheromone polymorphisms and the relatively simple nature of the pheromone components. These features have made the system amenable to reductionist approaches, although it has turned out to be surprisingly complex.

In *Ostrinia*, all of the pheromone precursors are synthesized de novo from palmitic acid (C16) through a few enzymatic steps including desaturation and chain-shortening (Fig. 2). Acyl precursors are subsequently converted into fatty alcohols, usually followed by acetylation to produce the acetate esters used as pheromone components. Earlier in vivo experiments demonstrated that the acetyltransferase, which is involved in the last step of the biosynthesis pathways leading to acetate pheromone components, has low substrate specificity in *Ostrinia* species (17–19). These results imply that pheromone blend ratio regulation occurs before the acetylation step. Previously, the intraspecific differences in pheromone blend ratio causing reproductive isolation between the diverging pheromone races of the European corn borer *O. nubilalis* were shown to result from mutations in the pheromone gland–specific fatty-acyl reductase gene *pgFAR*, which encodes the reductase catalyzing the specific reduction of the fatty-acyl pheromone precursors into fatty alcohols (10). To assess whether variation in the *pgFAR* gene parallels interspecific differences in pheromone phenotypes across the genus, we cloned and functionally characterized the orthologs of *O. nubilalis pgFAR*
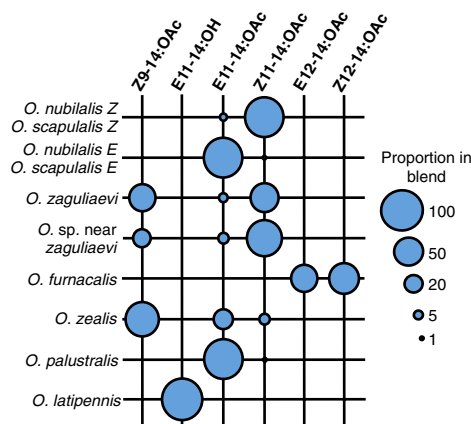
EVOLUTION

Fig. 1. Composition of female sex pheromones of *Ostrinia* moths. The size of a dot is proportional to the ratio of an individual pheromone component. Based on data reported elsewhere (14, 38, 48–50) and references therein.

from six additional *Ostrinia* congeners. Here our results establish that the pgFARs exhibit different substrate specificities that are the consequence of variation in the protein-coding regions of the *pgFAR* orthologs. Moreover, we show that pgFAR substrate preference has to be taken into account to explain the final ratios of the pheromone components. Finally, the results from site-directed mutagenesis experiments substantiate the idea that nonsynonymous amino acid substitutions can underlie subtle or drastic changes in pgFAR substrate preference that have the potential to impact the ratio of components found in the female sex pheromone.

Results

Functional Characterization of *Ostrinia* pgFARs. Coding sequences of the *pgFAR* orthologs showed on average 4.6% nucleotide divergence (0.4–8.4%) and 8.1% amino acid divergence (0.4–13.8%; Fig. S1; Table S1). These data show that there are extensive nonsynonymous substitutions compared with synonymous ones and, consequently, extensive amino acid variation in pgFARs, which may have functional consequences. In addition, the ratios of acyl precursors and pheromone components exhibit great disparities in most *Ostrinia* species (Fig. 3 *A* and *D*), which strongly suggests that the *pgFAR* orthologs encode reductases with different catalytic properties contributing to determine the final ratio used by a given species. To assess the effect of the observed amino acid differences on the reductase substrate specificity, we expressed the different *pgFARs* in yeast (*Saccharomyces cerevisae*) while supplementing the growth medium with the Δ9-, Δ11-, or Δ12-tetradecenoyl pheromone precursors. In a first set of experiments, the precursors were added individually, and their conversion into fatty alcohols was monitored by GC-MS analysis. The *pgFAR* orthologs conferred on the yeast the ability to convert acyl precursors into fatty alcohols, and more importantly, the enzymes varied in their ability to convert the different precursors (Fig. 3*B*). Based on the relative conversion efficiency of each precursor into the corresponding alcohol, we inferred the ratios of pheromone components that would result from combining the relative abundance of precursors and the affinity of pgFARs. Hence, the ratios of pheromone components predicted matched closely with the actual pheromone blend phenotypes produced by in *Ostrinia* female moths (Fig. 3 *C* and *D*).

In a second set of experiments, we supplemented the yeast growth medium with ratios of acyl precursors as observed in the *Ostrinia* female pheromone glands from which the *pgFARs* had been isolated. As anticipated, the final ratio of fatty alcohols produced by each pgFAR corresponded perfectly to the ratio of pheromone components found *in insecta* (Fig. 4).

pgFAR Polymorphisms and Tests of Selection. Our data show that there is extensive amino acid variation in pgFAR, raising the possibility that positive selection may have influenced the evolutionary trajectory of this gene. We investigated this in two complementary approaches: phylogenetic reconstruction and comparative analysis of the sequence data.

First, we examined the correspondence between the species phylogeny and the gene genealogy obtained with *pgFAR*. We obtained sequence information for five informative gene regions that have been successfully used to elucidate phylogenetic relationships among lepidopterans (20, 21), namely *elongation factor-1α* (*EF-1α*), *malate dehydrogenase* (*MDH*), *isocitrate dehydrogenase* (*IDH*), and *ribosomal protein S5* (*RpS5*) from the nuclear genome and *cytochrome oxidase subunit I* (*COI*) from the mitochondrial genome. Our molecular phylogenetic reconstruction is based on the combined nucleotide sequence dataset for these genetic markers totaling 4,804 bp. Interestingly, our phylogenetic analyses reveal a strong discordance between the *pgFAR* gene tree and the species tree (Fig. 5). The two trees have an overall topological similarity of 59.9%, and close examination of the trees reveals a shift of topological position across numerous sequences, which may have been caused by selection. One may speculate that the topology of the *pgFAR* gene tree depicts in part the functional relatedness between encoded enzymes rather than ancestry alone.

All proteins experience purifying selection due to the need to maintain structural and functional integrity, which narrows the range of possible amino acid substitutions, and the coding region of a gene is not uniform regarding the threshold of change acceptability (22). Consequently, positive selection may operate only on a subset of sites, with a large majority of codons experiencing purifying selection. We used a sliding-window analysis to uncover the ratio of nonsynonymous to synonymous substitutions (K_a/K_s) along the pgFAR protein sequence and detect regions of the protein that may have experienced positive selection. The analysis revealed that several regions in the protein exhibit a K_a/K_s ratio >1 (Fig. S2). Most of these regions are in the C-terminal part of the protein, which comprises the more variable Sterile pfam domain. The N-terminal region of FARs, which contains a Rossmann fold–like NADPH-binding pfam domain, harbors several regions that are highly conserved throughout the eukaryotes. To identify which sites may have experienced positive selection, we next inferred selection on individual amino acids using different detection tools. Results from the different evolutionary models supported the action of positive selection, and, of 465 codons, site-specific analyses showed strong evidence for positive selection at a total of 45 sites, 24 of which were
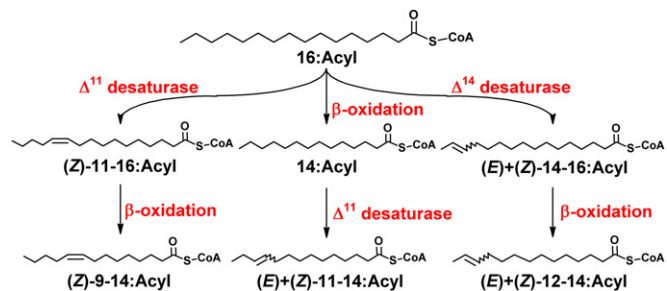
Fig. 2. Pathways toward pheromone precursors in *Ostrinia* moths. De novo biosynthesis of all precursors starts from palmitoyl-CoA (16:Acyl). A Δ11 desaturase catalyzes the production of (Z)-11-hexedecenoyl [(Z)-11–16:Acyl], which undergoes one cycle of β-oxidation to give (Z)-9-tetradecenoyl [(Z)-9–14:Acyl] (9, 51). The same Δ11 desaturase acts on myristoyl-CoA (14:Acyl) to produce (E)- and (Z)-11-tetradecenoyl [(E)+(Z)-11–14:Acyl] (9, 52). In *O. latipennis*, a different Δ11 desaturase uniquely catalyzes the stereo-specific production of (E)-11-tetradecenoyl [(E)-11–14:Acyl] (11). Finally, in *O. furnacalis*, (E)- and (Z)-12-tetradecenoyl [(E)+(Z)-12–14:Acyl] moieties are produced through chain-shortening of (E)- and (Z)-14-hexedecenoyl [(E)+(Z)-14–16:Acyl], respectively, the production of which is catalyzed by a Δ14 desaturase (9, 32).

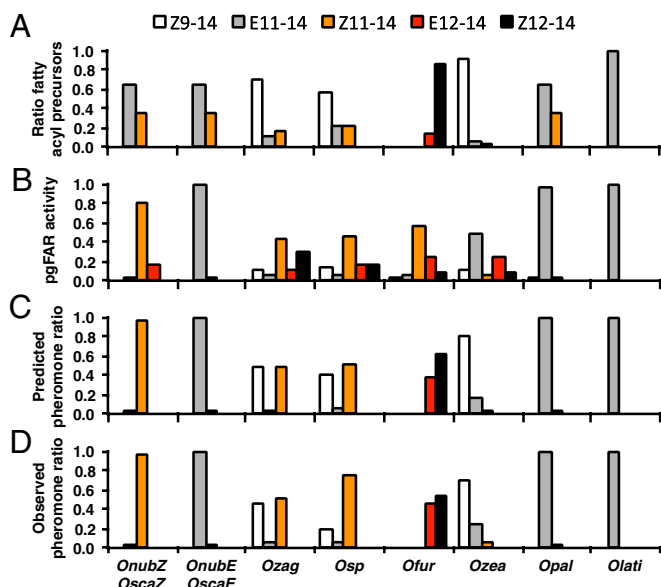**Fig. 3.** Activity of pgFARs in vitro compared with pheromone composition. (*A*) Schematic view of the relative abundance of acyl precursors in the female pheromone gland of *Ostrinia* species. (*B*) Relative abundance of the fatty alcohols accumulating in yeast expressing different *pgFARs* supplemented with methyl-ester precursors, one at a time (0.5 mM). (*C*) Predicted ratio of pheromone components calculated based on the relative abundance of precursors and the corresponding affinity of pgFARs. (*D*) Pheromone components and their ratio as reported previously (Fig. 1).

supported by the three methods used (Fig. S3; Table S2 and S3). Again, the C-terminal region contained a greater proportion of sites predicted to be under selection. Among the selected codons, 7 were associated with transitions in the charge of the encoded amino acid, 11 were associated with changes in the hydrophobicity profile, and 5 corresponded to changes in both charge and hydrophobicity at the particular site (Fig. S3). These predicted sites may mediate the biochemical specificity of the encoded enzyme and participate in enzyme-ligand interactions.

Site-Directed Mutagenesis of *O. furnacalis pgFAR*. The tertiary structure of fatty-acyl reductases cannot be modeled with accuracy, and structure–function relationships cannot be predicted either. Experimental approaches using methods such as site-specific mutagenesis can assist in teasing out the role of individual sites within a protein. Here we conducted mutagenesis experiments on two *pgFAR* orthologs displaying different substrate preferences while having the highest degree of sequence identity at the amino acid level: *Ofurnacalis*-pgFAR and *OnubilalisZ*-pgFAR. Both enzymes exhibit a strong substrate preference for Z11-14 acyl but differ in their biochemical properties: *OnubilalisZ*-pgFAR is unable to reduce the Z9-14 and Z12-14 acyl precursors that *Ofurnacalis*-pgFAR readily converts into the corresponding alcohols. The two enzymes also differ in the absolute amounts of primary fatty alcohols produced in our in vitro assay (Fig. S4). The two encoded proteins differ at 16 positions mainly found in the C-terminal region, 13 of which corresponded to sites predicted to have potentially experienced diversifying selection (Fig. S3). Five of these sites involving conservative substitutions (e.g., valine-leucine, isoleucine-leucine) were excluded a priori because they were unlikely to cause a change in function or structure (23). We mutated the remaining eight sites in *Ofurnacalis*-pgFAR individually to the corresponding *OnubilalisZ*-pgFAR residue in a loss-of-function approach. Using this strategy, we identified mutation F453C as affecting the enzyme's affinity for the Z9-14 and Z12-14 acyl precursors (Fig. 6). The *Ofurnacalis*-pgFAR^{F453C} mutant exhibited a substrate preference remarkably similar to WT *OnubilalisZ*-pgFAR. Conversely, the

chimeric enzyme with the reverse mutation C453F in *OnubilalisZ*-pgFAR had a substrate specificity highly similar to that of WT *Ofurnacalis*-pgFAR (Fig. 7). Two mutations, T137A and P146S, did not lead to any significant modification of the blend. Three of the other mutations, S431G, Y449F, and NY462/463D*, caused more minor modifications of enzymatic activity, resulting in the alteration of the relative ratios of the resulting alcohols (Fig. 6; Fig. S4). The substitutions E175K and H413R led to an overall decrease in alcohol production (approximately fourfold), without affecting the relative ratios of the alcohols (Fig. 6; Fig. S4). Subsequently, we produced double mutants by introducing each mutation individually in the *Ofurnacalis*-pgFAR^{F453C} background. Our results confirmed that these other mutations have the potential to modify the enzymatic activity and alter alcohol ratios (Fig. S5). For example, the introduction of S431G in either the *Ofurnacalis*-pgFAR^{WT} or *Ofurnacalis*-pgFAR^{F453C} background resulted in an increased proportion of E12-14:OH in the blend. Interestingly, a mutant carrying all five mutations T137A, P146S, S431G, Y449F, and NY462/463D* in an *Ofurnacalis*-pgFAR background was still capable of reducing the Z9-14 and Z12-14 acyl precursors, indicating that the F453C mutation is essential to generate the shift in substrate preference.

Discussion

Sex pheromones are one of the main features that define biological species within Lepidoptera (8, 24). After 50 y of research in this field, it is clear that the use of a narrow range of blend
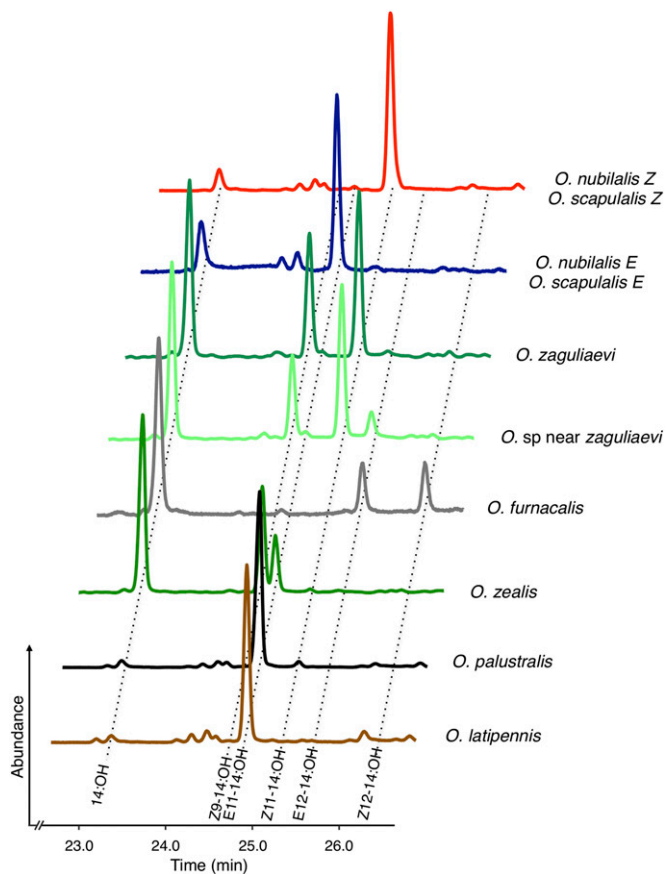
Fig. 4. Alcohol production of yeast expressing *Ostrinia* pgFARs when supplemented with pheromone precursors at natural ratios. The traces represent portions of the total ion chromatogram of extracts made from yeast grown in the presence of all methyl esters precursors (0.5 mM in total). The ratios of precursors supplemented correspond to the relative abundance of each fatty-acyl precursor naturally occurring in the pheromone gland (Fig. 3*A*).
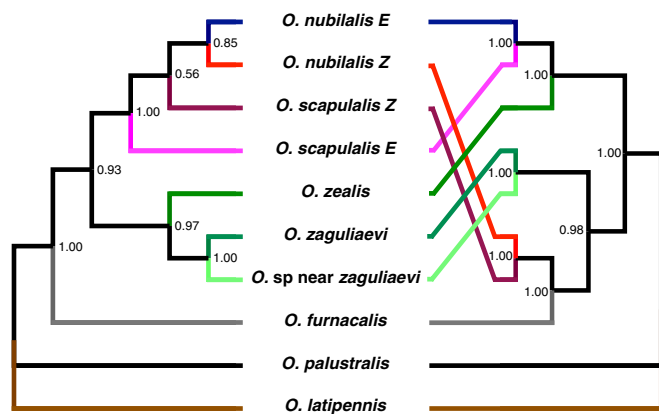
Fig. 5. Discordance between the phylogenetic trees of *Ostrinia* species and *pgFAR* orthologs. The Bayesian phylogenetic species tree (*Left*) was established based on five markers (*COI*, *EF-1α*, *MDH*, *IDH*, and *RpS5*) that are informative in the context of Lepidoptera phylogenetics (21), whereas the *pgFAR* tree (*Right*) was built using the coding region of the gene. Both nucleotide and amino acid datasets produced the same topology. Numbers correspond to Bayesian posterior probabilities.

ratios is one of the hallmarks defining the species specificity of moth pheromones (3, 8). The diversity of signals has its roots in the biochemical pathways that take place in the female pheromone gland, involving key enzymatic steps (25, 26). Although the sex pheromones of hundreds of lepidopteran species have been identified to date (27), our understanding of the molecular mechanisms underpinning variation and evolution within pheromone communication systems remains limited. To further address these questions, we investigated the molecular and biochemical differences in a fatty-acyl reductase gene involved in pheromone biosynthesis within the moth genus *Ostrinia*, revealing the functional consequences of coding variation in the *pgFAR* gene and showing how it correlates with observed differences in pheromone composition among congeners. Our data illustrate how reductases may evolve new functional properties as a consequence of accumulating nonsynonymous amino acid changes.

The composition of available acyl precursors may differ from species to species, which in itself contributes to variation in blend composition (26). Indeed, differential regulation of desaturase genes has been shown to be associated with evolutionary transitions in the pheromone communication systems of flies and moths including in *Ostrinia* (9, 11, 28–31). However, in most *Ostrinia* species, the precursor and pheromone component ratios exhibit significant disparities, and the pgFAR substrate preference

is in many cases necessary to explain how the final blend ratios are obtained. Our results show that structural mutations in *pgFAR* can alter the specificity for the location and isomeric nature of double bonds in the desaturated precursors, which has the potential to modify the isomeric composition of sex pheromone blends. In *Ostrinia*, pgFARs act in concert with other biosynthetic enzymes and catalyze an enzymatic step that shapes the final pheromone blend ratio, thus contributing to determine species-specific ratios, an essential feature of the pheromone signals. Taken together, our data support that coding variation in *pgFAR* may influence how pheromone blends evolve and thereby contribute to the specificity of the *Ostrinia* mating systems.

Some pgFARs have the ability to convert in vitro precursors that are not naturally present in the species from which the gene originated. Similar observations have been reported when the reductase system of *O. nubilalis* and *O. furnacalis* were studied in vivo using deuterium-labeled precursors (18, 19), substantiating the idea that these are not artifacts of the yeast system. We interpret these enzymatic properties as both ghosts from the evolutionary past of the species, i.e., the remains of ancestral features, as well as evidence for intrinsic preadaptations. Such preadaptations of the reductase system were necessary to allow the evolution of *O. furnacalis*, which uniquely possesses Δ12-tetradecenyl precursors as a consequence of the activity of a Δ14 desaturase (9, 32). Interestingly, although distantly related to *Ostrinia*, the *pgFAR* active in small ermine moths of the genus *Yponomeuta* (Lepidoptera: Yponomeutidae) also has the ability to reduce Δ12-tetradecenyl precursors (33). In addition, our data exemplify that different substitutions can produce similar functional effects, which may result from convergent evolution. Hence, *OnubilalisE*-pgFAR and *Opalustralis*-pgFAR have the same substrate preference and produce an identical ratio of alcohols despite showing 12.3% divergence in their amino acid sequences. In contrast, *OnubilalisZ*-pgFAR and *Ofurnacalis*-pgFAR show 96.1% identity at the amino acid level (98.2% at the nucleotide level). Although both enzymes exhibit a strong substrate preference for Z11-14 acyl, they differ in their biochemical properties. This result points to the necessity of performing tests of gene function, as computational analyses of sequence similarity alone may be misleading. We engineered chimeric *pgFAR* mutants to begin to disentangle sequence–structure–function relationships. Our mutagenesis data show that a single nucleotide mutation in the C-terminal region can determine the ability to catalyze the reduction of Z9-14 and Z12-14 acyl precursors. It is worth noting that, although remarkable in itself, the change in activity conveyed by the *Ofurnacalis*-pgFAR^{F453C} mutation is close but not identical to *OnubilalisZ*-pgFAR. Indeed, the fine-tuning of the reduction step and the precise *OnubilalisZ* ratio
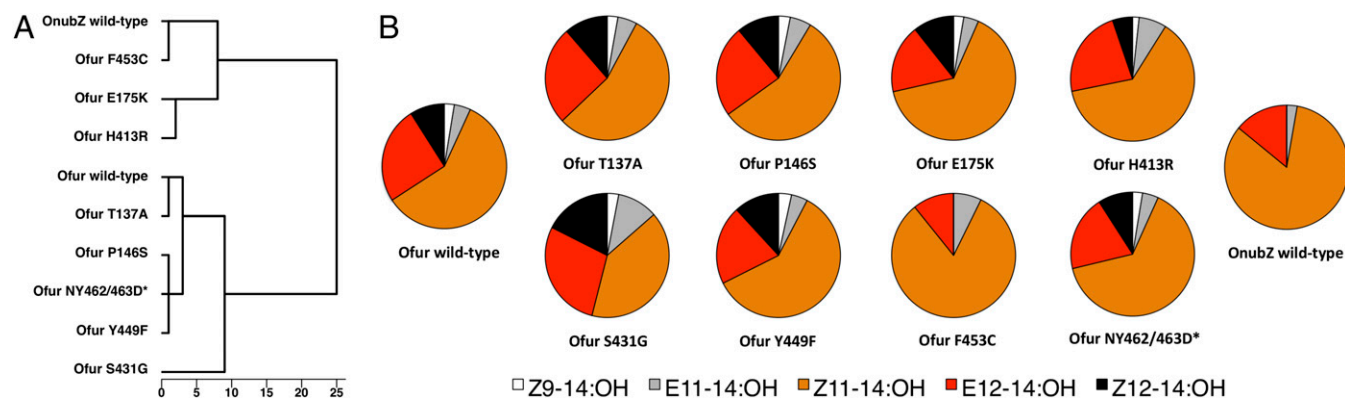

Fig. 6. Functional expression assays of parental and single-codon mutant pgFARs. (*A*) Hierarchical cluster dendrogram (Ward's grouping method, Euclidean distances) showing distances among the eight mutants and the WT pgFAR from *O. furnacalis* (parental background) and *O. nubilalis* Z. (*B*) Ratios of fatty-alcohol products for WT and mutant pgFARs.
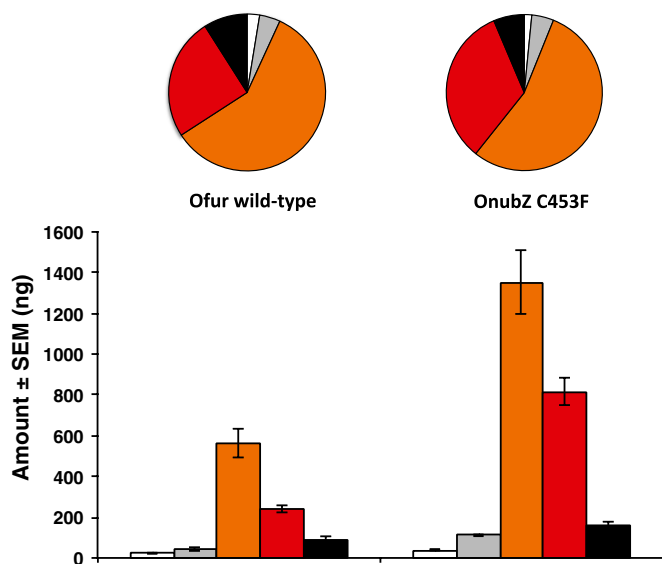
Fig. 7. Functional impact of the reverse mutation C453F on *OnubilalisZ*-pgFAR activity compared with the *Ofurnacalis* WT. The pie charts and graphs represents the relative and absolute abundances of alcohols, respectively (*n* = 4).

is likely the consequence of the subtle additive effects of additional mutations.

In conclusion, because mutations at the *pgFAR* locus have the potential to modify the proportion of individual pheromone components, our study identifies the enzymatic step catalyzed by *pgFAR* as a candidate source of variation in moth pheromone signals. Variation is necessary to provide the conditions for divergence in pheromone signals. It has been suggested earlier that natural selection on mate recognition may be a major component of the evolution of the mate recognition system of many taxa (34, 35). In moths, selection acting on the female pheromone signals is both direct and indirect as a consequence of the selection operating on males that are attracted to the female-produced pheromone blend (36). Because of the important role played by *pgFAR* in shaping the pheromone signal emitted by females, selection would have acted to maintain and enhance the fixation probability of those mutations, leading to new phenotypes whenever conferring a slight selective advantage. Hence, male moths may prefer a deviant pheromone ratio if it promotes privacy of the communication channel and limits interference with other species, which may in turn increase their fitness (36, 37). Chemical messages to recruit potential mates can act as a strong impediment to gene flow with inappropriate mates, and divergence in chemical signals is often an essential step in divergence between populations that may or may not lead to speciation. Whether a new signal will emerge depends on the ability of receivers to respond to it and the adaptive advantages inherent in the use of the new signal. Variation in the pheromone signal must therefore be paralleled by variation in male reception, as exemplified by the potential role played by a shift in pheromone receptor sensitivity in the evolution of *O. furnacalis* (12). Functional and comparative genomic studies of male olfactory receptor genes and genes defining the architecture of neurophysiological pathways will help us to better infer the evolutionary processes that have shaped pheromone communication systems and signal response coevolution.

Materials and Methods

Amplification of *pgFAR* Transcripts. We sampled and pooled per species 10–20 pheromone glands of 0- to 3-d-old virgin females of the Far Eastern knotweed borer *O. latipennis*, the dock borer *O. palustralis*, the adzuki bean borer *O. scapulalis* (E and Z pheromone strains), the butterbur borer *O. zaguliaevi*, the leopard plant borer *O. sp.* near *zaguliaevi* [a new member of *Ostrinia* for

which the taxonomic status remains to be determined (38)], and the burdock borer *O. zealis*, all of which were collected in Japan. Females of the Asian corn borer *O. furnacalis* originated from a Chinese laboratory strain, whereas females of both pheromone strains of the European corn borer *O. nubilalis* were from France. We extracted total RNA using the RNeasy isolation kit (Qiagen) and synthesized first-strand cDNA using Stratascript reverse transcriptase (Stratagene). An ~290-bp fragment was amplified using pgFAR_F and pgFAR_R primers as previously described (10, 39). Gene-specific primers were designed and used in PCRs to amplify the 5′ and 3′ cDNA ends with the SMART RACE amplification kit (Clontech) following the manufacturer's instructions. PCR products were gel-purified and ligated into the pGEM-T easy vector (Promega) or the PCR2.1 vector (Invitrogen). Ligation products were used to transform *Escherichia coli* TOPO10 chemically competent cells (Invitrogen). Plasmids were purified and served as templates for sequencing reactions with the BigDye terminator kit v1.1 (Applied Biosystems) followed by analysis on a ABI PRISM 3130xl genetic analyzer (Applied Biosystems). Primer walking sequencing was carried out to obtain information from long amplification products. Full-length sequences were compiled using the sequence information from the 5′ and 3′ RACE PCR products. Data were analyzed using Bioedit version 7 (40).

Phylogenetic Analysis. The sequence data used to build the species phylogeny corresponded to four protein-coding nuclear regions and one mitochondrial region amplified using pheromone gland cDNAs as template. Phylogenetic reconstruction was inferred using MrBayes v3.2 (41). Only the open-reading frame sequence was used for phylogenetic reconstruction of the *pgFARs*. The similarity in topology between the species and *pgFAR* trees was estimated following the method described in ref. 42. (*SI Materials and Methods*).

Sequence Analysis and Tests of Selection. Before testing for evidence of positive selection in the evolutionary history of the *pgFAR* gene, we explored whether recombination might have impacted our phylogenetic reconstruction and generated the pattern we observe. We tested for the presence of recombination breakpoints using GARD, a genetic algorithm for recombination detection available through the Datamonkey webserver (www.datamonkey. org) (43, 44). GARD predicted three breakpoints of which none showed significant topological incongruence using the Kishino-Hasegawa test, even at *P* = 0.1. Therefore, the breakpoints predicted by GARD are more likely due to spatial rate variation or heterotachy rather than recombination.

We performed a sliding window analysis of the K_a/K_s ratio of nonsynonymous over synonymous substitutions based on Li's method (45) by using the SWAAP 1.0.2 software. Next, we tested for evidence of positive selection using the random effects likelihood (REL) and the fast unbiased approximate Bayesian (FUBAR) tests of selection (both are implemented on the Datamonkey webserver) and the maximum likelihood methods implemented in Codeml, part of the PAML 4.4b package (46). In Codeml, we performed the analysis using the species tree and tested the specific site models implemented (M1, M2, M7, and M8). Likelihood ratio tests were used to compare the neutral models with the corresponding selection models (M1 vs. M2, M7 vs. M8, and M8a vs. M8). Results from the comparisons were consistent, all supporting the action of positive selection. The Bayes empirical Bayes procedure implemented under model M8 was used to identify sites under positive selection.

Analysis of Pheromone Precursors. Data were taken from previous published studies. When no data were available, tetradecenoyl precursors of pheromone components were analyzed in the form of methyl-esters (FAMEs) according to the method described earlier (47).

Functional Assays. Functional assays were carried out in a yeast expression system following the procedures described previously (10, 33). Briefly, the ORFs of the various *pgFAR* genes were cloned into the expression vector pYES2.1/V5-His TOPO (Invitrogen), and the resulting recombinant vectors were used to transform the yeast *Saccharomyces cerevisae* (INVSc1 strain; Invitrogen). Samples were analyzed by GC-MS (*SI Materials and Methods*).

Site-Directed Mutagenesis. Targeted mutagenesis was performed using a PCR-based approach and the Quickchange II Site-directed mutagenesis method (Stratagene). Mutagenic oligonucleotide primers were designed with the QuikChange Primer design program (www.stratagene.com). In most cases, a single nucleotide substitution was sufficient to introduce the targeted amino acid change. All clones were sequenced in both directions to confirm the introduction of the mutated nucleotide. The functional assays with individual precursors were carried out following the procedures described above. For each *pgFAR* mutant and substrate, the absolute amount of fatty alcohol produced was derived using the internal standard method (*n* = 2–4). Hierarchical

cluster analysis was used to define clusters based on the dissimilarities (distances) between the WT and single-codon mutant pgFARs. We used the absolute amount of each fatty-alcohol product and the proportion of each individual alcohol within the total amount of alcohol produced (standardized z scores). The analysis was carried out in SPSS Statistics version 20 (IBM).

1. Wyatt TD (2003) *Pheromones and Animal Behaviour: Communication by Smell and Taste* (Cambridge Univ Press, Cambridge, New York).
2. Coyne JA, Orr HA (2004) *Speciation* (Sinauer Associates, Sunderland, MA).
3. Smadja C, Butlin RK (2009) On the scent of speciation: The chemosensory system and its role in premating isolation. *Heredity (Edinb)* 102(1):77–97.
4. Symonds MRE, Elgar MA (2008) The evolution of pheromone diversity. *Trends Ecol Evol* 23(4):220–228.
5. Butlin R, et al.; Marie Curie SPECIATION Network (2012) What do we need to know about speciation? *Trends Ecol Evol* 27(1):27–39.
6. Kristensen NP, Scoble MJ, Karsholt O (2007) Lepidoptera phylogeny and systematics: The state of inventorying moth and butterfly diversity. *Zootaxa* 1668:699–747.
7. Löfstedt C (1993) Moth pheromone genetics and evolution. *Philos Trans R Soc Lond B Biol Sci* 340:167–177.
8. Cardé RT, Haynes KF (2004) Structure of the pheromone communication channel in moths. *Advances in Insect Chemical Ecology*, eds Carde RT, Millar JG (Cambridge Univ Press, New York), pp 283–332.
9. Roelofs WL, et al. (2002) Evolution of moth sex pheromones via ancestral genes. *Proc Natl Acad Sci USA* 99(21):13621–13626.
10. Lassance J-M, et al. (2010) Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. *Nature* 466(7305):486–489.
11. Fujii T, et al. (2011) Sex pheromone desaturase functioning in a primitive *Ostrinia* moth is cryptically conserved in congeners' genomes. *Proc Natl Acad Sci USA* 108(17):7102–7106.
12. Leary GP, et al. (2012) Single mutation to a sex pheromone receptor provides adaptive specificity between closely related moth species. *Proc Natl Acad Sci USA* 109(35):14081–14086.
13. Linn CE, Young MS, Gendle M, Glover TJ, Roelofs WL (1997) Sex pheromone blend discrimination in two races and hybrids of the European corn borer moth, *Ostrinia nubilalis*. *Physiol Entomol* 22(3):212–223.
14. Kochansky J, Cardé RT, Liebherr J, Roelofs WL (1975) Sex pheromone of the European corn borer, *Ostrinia nubilalis* (Lepidoptera: Pyralidae), in New York. *J Chem Ecol* 1(2):225–231.
15. Huang Y, et al. (1998) Geographic variation in sex pheromone of Asian corn borer, *Ostrinia furnacalis*, in Japan. *J Chem Ecol* 24(12):2079–2088.
16. Boo KS, Park JW (1998) Sex pheromone composition of the Asian corn borer moth, *Ostrinia furnacalis* (Guenée) (Lepidoptera: Pyralidae) in South Korea. *J Asia Pac Entomol* 1(1):77–84.
17. Jurenka RA, Roelofs WL (1989) Characterization of the acetyltransferase used in pheromone biosynthesis in moths: Specificity for the Z isomer in tortricidae. *Insect Biochem* 19(7):639–644.
18. Zhao CH, Lu F, Bengtsson M, Löfstedt C (1995) Substrate specificity of acetyltransferase and reductase enzyme systems used in pheromone biosynthesis by Asian corn borer, *Ostrinia furnacalis*. *J Chem Ecol* 21(10):1496–1510.
19. Zhu JW, Zhao CH, Bengtsson M, Löfstedt C (1996) Reductase specificity and the ratio regulation of *E/Z* isomers in the pheromone biosynthesis of the European corn borer, *Ostrinia nubilalis* (Lepidoptera: Pyralidae). *Insect Biochem Mol Biol* 26(2):171–176.
20. Wahlberg N, Wheat CW (2008) Genomic outposts serve the phylogenomic pioneers: Designing novel nuclear markers for genomic DNA extractions of lepidoptera. *Syst Biol* 57(2):231–242.
21. Mutanen M, Wahlberg N, Kaila L (2010) Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. *Proc Biol Sci* 277(1695):2839–2848.
22. Kwast KE, et al. (2002) Genomic analyses of anaerobically induced genes in Saccharomyces cerevisiae: Functional roles of Rox1 and other factors in mediating the anoxic response. *J Bacteriol* 184(1):250–265.
23. Betts MJ, Russell RB (2003) Amino acid properties and consequences of substitutions. *Bioinformatics for Geneticists*, eds Barnes MR, Gray IC (Wiley, Chichester, UK), pp 289–316.
24. Linn CE, Roelofs WL (1995) Pheromone communication in moths and its role in the speciation process. *Speciation and the Recognition Concept: Theory and Application*, eds Lambert DM, Spencer HG (Johns Hopkins Univ Press, Baltimore), pp 263–300.
25. Tillman JA, Seybold SJ, Jurenka RA, Blomquist GJ (1999) Insect pheromones—an overview of biosynthesis and endocrine regulation. *Insect Biochem Mol Biol* 29(6):481–514.
26. Roelofs WL, Rooney AP (2003) Molecular genetics and evolution of pheromone biosynthesis in Lepidoptera. *Proc Natl Acad Sci USA* 100(16):9179–9184.
27. El-Sayed AM (2011) The Pherobase: Database of insect pheromones and semiochemicals. Available at http://www.pherobase.com. Accessed February 2, 2012.
28. Sakai R, Fukuzawa M, Nakano R, Tatsuki S, Ishikawa Y (2009) Alternative suppression of transcription from two desaturase genes is the key for species-specific sex pheromone biosynthesis in two *Ostrinia* moths. *Insect Biochem Mol Biol* 39(1):62–67.
29. Albre J, et al. (2012) Sex pheromone evolution is associated with differential regulation of the same desaturase gene in two genera of leafroller moths. *PLoS Genet* 8(1):e1002489.
30. Shirangi TR, Dufour HD, Williams TM, Carroll SB (2009) Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. *PLoS Biol* 7(8):e1000168.
31. Dallerac R, et al. (2000) A delta 9 desaturase gene with a different substrate specificity is responsible for the cuticular diene hydrocarbon polymorphism in Drosophila melanogaster. *Proc Natl Acad Sci USA* 97(17):9449–9454.
32. Zhao CH, Löfstedt C, Wang XY (1990) Sex pheromone biosynthesis in the Asian corn borer *Ostrinia furnacalis* (II) - Biosynthesis of (*E*)-12-tetradecenyl and (*Z*)-12-tetradecenyl acetate involves delta14 desaturation. *Arch Insect Biochem Physiol* 15(1):57–65.
33. Liénard MA, Hagström ÅK, Lassance J-M, Löfstedt C (2010) Evolution of multicomponent pheromone signals in small ermine moths involves a single fatty-acyl reductase gene. *Proc Natl Acad Sci USA* 107(24):10955–10960.
34. Dobzhansky T (1951) *Genetics and the Origin of Species* (Columbia Univ Press, New York).
35. Butlin R (1995) Genetic variation in mating signals and responses. *Speciation and the Recognition Concept: Theory and Application*, eds Lambert DM, Spencer HG (Johns Hopkins Univ Press, Baltimore), pp 327–366.
36. Bengtsson BO, Löfstedt C (2007) Direct and indirect selection in moth pheromone evolution: Population genetical simulations of asymmetric sexual interactions. *Biol J Linn Soc Lond* 90(1):117–123.
37. Groot AT, et al. (2006) Experimental evidence for interspecific directional selection on moth pheromone communication. *Proc Natl Acad Sci USA* 103(15):5858–5863.
38. Tabata J, et al. (2008) Sex pheromone of *Ostrinia* sp newly found on the leopard plant *Farfugium japonicum*. *J Appl Entomol* 132(7):566–574.
39. Antony B, et al. (2009) Pheromone-gland-specific fatty-acyl reductase in the adzuki bean borer, *Ostrinia scapulalis* (Lepidoptera: Crambidae). *Insect Biochem Mol Biol* 39(2):90–95.
40. Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98.
41. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12):1572–1574.
42. Nye TMW, Liò P, Gilks WR (2006) A novel algorithm and web-based tool for comparing two alternative phylogenetic trees. *Bioinformatics* 22(1):117–119.
43. Pond SLK, Frost SDW (2005) Datamonkey: Rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21(10):2531–2533.
44. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW (2006) GARD: A genetic algorithm for recombination detection. *Bioinformatics* 22(24):3096–3098.
45. Li W-H (1993) Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J Mol Evol* 36(1):96–99.
46. Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24(8):1586–1591.
47. Tabata J, Hoshizaki S, Tatsuki S, Ishikawa Y (2006) Heritable pheromone blend variation in a local population of the butterbur borer moth *Ostrinia zaguliaevi* (Lepidoptera: Crambidae). *Chemoecology* 16(2):123–128.
48. Ishikawa Y, et al. (1999) *Ostrinia* spp. in Japan: Their host plants and sex pheromones. *Entomol Exp Appl* 91(1):237–244.
49. Cheng ZQ, et al. (1981) Sex pheromone components isolated from China corn borer, *Ostinia furnacalis* Guenée (Lepidoptera, Pyralidae), (*E*)-12-tetradecenyl and (*Z*)-12-tetradecenyl acetates. *J Chem Ecol* 7(5):841–851.
50. Lassance J-M (2010) Journey in the *Ostrinia* world: From pest to model in chemical ecology. *J Chem Ecol* 36(10):1155–1169.
51. Fu X, Fukuzawa M, Tabata J, Tatsuki S, Ishikawa Y (2005) Sex pheromone biosynthesis in *Ostrinia zaguliaevi*, a congener of the European corn borer moth *O. nubilalis*. *Insect Biochem Mol Biol* 35(6):621–626.
52. Roelofs WL, et al. (1987) Sex pheromone production and perception in European corn borer moths is determined by both autosomal and sex-linked genes. *Proc Natl Acad Sci USA* 84(21):7585–7589.

Supporting Information

Lassance et al. 10.1073/pnas.1208706110

SI Materials and Methods

Phylogenetic Analysis. We amplified the five loci by PCR using pheromone gland cDNAs as template. The amplified regions were as follows: *elongation factor-1α* (*EF-1α*), *malate dehydrogenase* (*MDH*), *isocitrate dehydrogenase* (*IDH*), and *ribosomal protein S5* (*RpS5*) from the nuclear genome and *cytochrome oxidase subunit I* (*COI*) from the mitochondrial genome. Primers were taken from Wahlberg and Wheat (1), and the PCR amplifications were performed using the following touchdown profile: 2 min at 94 °C, 15 cycles of 30 s at 94 °C, 30 s of primer annealing at an initial temperature of 55 °C dropped 0.3 °C per cycle, and 1 min at 72 °C, followed by 25 cycles of 30 s at 94 °C, 30 s at 55 °C, and 1 min at 72 °C. PCR products were treated with a mixture containing exonuclease I (EXO; Fermentas) and shrimp alkaline phosphatase (SAP; Fermentas) and sequenced in both directions using the T7 and T3 universal primers with the BigDye terminator kit v1.1 (Applied Biosystems), followed by analysis on an ABI PRISM 3130xl genetic analyzer (Applied Biosystems). Data were analyzed using the BioEdit sequence analysis software version 7 (2). Any heterozygous position was coded following the International Union of Pure and Applied Chemistry (IUPAC) ambiguity code. The combined nucleotide sequence dataset comprised 4,804 bp that were partitioned by gene. Phylogenetic reconstruction was inferred using MrBayes v3.2 (3). The substitution model (GTR+I +Γ) was selected following model testing using MultiPhyl (4). The Bayesian analysis was performed for 1 million generations, with parameter values estimated separately for each gene. Every 100th generation was sampled, and the first 1,000 were discarded as burn-in. We selected *Ostrinia latipennis* as an outgroup.

For the pheromone biosynthetic fatty-acyl reductase gene *pgFAR* gene phylogeny reconstruction, the sequence data used corresponded to the coding region of the gene only. We excluded the UTRs because no reliable alignment of those could be obtained. In particular, the 3′ UTRs are extensively diverged from one species to the other. The phylogenetic tree was inferred with MrBayes implementing the GTR+Γ model with the nucleotide sequences and the JTT+I+Γ model with the amino acid sequences. The tree topologies were the same irrespective of the dataset used. Again, *O. latipennis* was enforced as an outgroup.

Functional assays. Functional assays were carried out in a yeast expression system following the procedures described previously (5, 6). Briefly, the open reading frames of the various *pgFAR* genes were cloned into the expression vector pYES2.1/V5-His TOPO (Invitrogen) and the resulting recombinant vectors were used to transform yeast, Saccharomyces cerevisae (INVSc1 strain, Invitrogen). When testing the conversion of individual precursors, aliquots of yeast cultures were suspended in induction medium containing 0.5 mM of the FAME precursor diluted in ethanol. After 24 h of incubation at 30°C, alcohol products were extracted in n-hexane spiked with 150 ng of (Z)-11-tridecenol (Z11-13:OH) used as internal standard. Samples were analyzed by gas chromatography-mass spectrometry under the conditions described earlier (5). For each substrate, the alcohol production was calculated using the internal standard method. For each pgFAR, the relative abundance of each fatty alcohol was derived using the formula:

$$\text{ratio of alcohol } y \text{ in vitro} = \frac{\text{amount of } y \text{ (in ng)}}{\sum \text{amount of all alcohols (in ng)}}.$$

The predicted pheromone ratio was calculated using the following formula: predicted ratio of *y* in pheromone blend = (ratio of precursor of alcohol *y* in the gland) × (ratio of alcohol *y* in vitro) (results were scaled so that the sum of predicted ratios would equal 1).

When testing the alcohol production of yeast supplemented with precursor blends in the ratio as found in female pheromone glands, the total concentration of precursors was 0.5 mM and the culture medium was 4 ml, whereas all other parameters were kept unchanged. After concentration under a gentle stream of pure nitrogen, yeast extracts were analyzed on a Hewlett Packard HP 6890 gas chromatograph (Agilent Technologies) coupled to a mass-selective detector HP 5973 operated in electron impact mode (70 eV). The GC was equipped with a HP-INNOWax capillary column (30 m × 0.25 mm × 0.15 μm; Agilent Technologies). Helium was used as carrier gas (velocity 30 cm s^{-1}) and the injection port was maintained at 230°C. One μl of sample was injected in spitless mode using a HP 7683 automatic liquid sampler (Agilent Technologies). In order to obtain good separation of the different isomers, the oven temperature program was set as follows: 2 min at 100°C, then the temperature was raised to 150°C at a rate of 5°C min^{-1} then to 200°C at a rate of 2°C min^{-1}. A post-run of 10 min at 230°C was used to clean the column between runs.

1. Wahlberg N, Wheat CW (2008) Genomic outposts serve the phylogenomic pioneers: Designing novel nuclear markers for genomic DNA extractions of lepidoptera. *Syst Biol* 57(2):231–242.
2. Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98.
3. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12):1572–1574.
4. Keane TM, Naughton TJ, McInerney JO (2007) MultiPhyl: A high-throughput phylogenomics webserver using distributed computing. *Nucleic Acids Res* 35(Web Server issue):W33-7.
5. Lassance J-M, et al. (2010) Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. *Nature* 466(7305):486–489.
6. Liénard MA, Hagström ÅK, Lassance J-M, Löfstedt C (2010) Evolution of multicomponent pheromone signals in small ermine moths involves a single fatty-acyl reductase gene. *Proc Natl Acad Sci USA* 107(24):10955–10960.

```
OnubZ    1  MSANTMETDEQFT-YNSPIVNFYSGKSVFVTGATGFLGTVLVEKLLFSCKGINNIYILIKQTEDLTIEAR
OscaZ    1  MSANTMETDEQFT-DNSPIVNFYSGKSVFVTGATGFLGTVLVEKLLFSCKGINNIYILIKQTEDLTIEAR
Ofur     1  MSANTMETDEQFT-YNSPIVNFYSGKSVFVTGATGFLGTVLVEKLLFSCKGINNIYILIKQTEDLTIEAR
OnubE    1  MSANTMETDEQFT-YNSPIVNFYSGKSVFVTGATGFLGTVLVEKLLFSCKGINNIYILIKQTEDLTIEAR
OscaE    1  MSANTMETDEQFT-YNSPIVNFYSGKSVFVTGATGFLGTVLVEKLLFSCKGINNIYILIKQTEDLTIEAR
Ozea     1  MAANTMETDEQFT-DNSPIVNFYSGKSVFVTGATGFVGTVLVEKLLFSCKGINNIYILIKQTEDLTIEAR
Osp      1  MSANTVETDEQFTDNNSPIVNFYSGKSVFVTGATGFLGTVLVEKLLFSCKGINNIYILIKQTEDLTIETTR
Ozag     1  MSANTVETDEQFTDNNSPIVNFYSGKSVFVTGATGFLGTVLVEKLLFSCKGINNIYILIKQTEDLTIEAR
Olati    1  MAANTMETDEQFT-YNSPIVNFYSGKSVFVTGATGFLGTVLVEKLLFSCKGINNIYVLIKQTRDLTIEGR
Opal     1  MSANTMESDHQLINYNSPIVNFYSGKSVFVTGATGFLGTVLVEKLLFSCKGINNIYVLIKQSRDLTIEGR


OnubZ   70  ILNYLNSKAFHRVKNTNPELMKKIIPICGNLEDKNLGISDSDMKTLLEEVSIVFHLAAKLLFKMSIAAAV
OscaZ   70  ILNYLNSKAFHRVKNTNPELMKKIIPICGNLEDKNLGISDSDMKTLLEEVSIVFHLAAKLLFKMSIAAAV
Ofur    70  ILNYLNSKAFHRVKNTNPELMKKIIPICGNLEDKNLGISDSDMKTLLEEVSIVFHVAAKLIFKMSLTAAV
OnubE   70  ILNYLNSKAFHRVKNTNPELMKKIIPICGNLEDKNLGISDSDMKTLLEEVSIVFHVAAKLLFKMSLTAAV
OscaE   70  ILNYLNSKAFHRVKNTNPELMKKIIPICGNLEDKNLGISDSDMKTLLEEVSIVFHVAAKLLFKMSLTAAV
Ozea    70  ILNYLNSKEFHRVKNTNPELMKKIIPICGNLEDKNLGISDSDMKTLLEEVSIVFHLAAKLLFKMNLTAAV
Osp     71  ILNYLNSKAFHRVKNTNPELMKKIIPICGNLEDKNLGISDSDMKTLLEEVSIVFHLAAKLLFKMSLTAAV
Ozag    71  ILNYLNSKAFHRVKNTNPELMKKIIPICGNLEDKNLGISDSDMKTLLEEVSIVFHLAAKLLFKMSLTAAV
Olati   70  VLNYLNSKAFHRVKNTNPELMKKIIPICGNLEDKNLGISDSDIKTLQEEVSIVFHVAAKLLFKMSIATAV
Opal    71  ILNYLNSKAFHRVRNANPEFMKKIIPICGNLEPNNLGISDSDMKTLQAEVSIVFHVAAKLLFKMSIATAV


OnubZ  140  NINTKSTEQLIAICKKMRRNPIFIYVSSAYSNVNKQIIDEKVYSTGVPLETIYDTLDAKNTRIMDIFLDK
OscaZ  140  NINTKSTEQLIAICKKMRRNPIFIYVSSAYSNVNKQIIDEKVYSTGVPLETIYDTLDAKNTRIVDIFLDK
Ofur   140  NINTKPTEQLIAICKKMRRNPIFIYVSSAYSNVNEQIIDEKVYSTGVPLETIYDTLDAKNTRIMDIFLDK
OnubE  140  NINTKPTEQLIAICKKMRRNPIFIYVSSAYSNVNEQIIDEKVYNTGVPLETIYDTLDTENTRTTDIFLDK
OscaE  140  NINTKPTEQLIAICKKMRRNPIFIYVSSAYSNVNEQIIDEKVYNTGVPLETIYDTLDTENTRTTDIFLDK
Ozea   140  NINTKPTEQLIAICKKMRRNPIFIYVSSAYSNVNEQIIDEKVYSTRVPLETYYDTLDTENTRTTDIFLDK
Osp    141  NINTKPTEQLIAICKKMRRNPIFIYVSSAYSNVNKQIIDEKVYSTGVPLETIYDTLDTENTRTTDIILDK
Ozag   141  NINTKPTEQLIAICKKMRRNPIFIYVSSAYSNVNKQIIDEKVYSTGVPLETIYDTLNTENTRTTDIILDK
Olati  140  NINTKPTEQLIVICKKMRRNPIFIYVSSAYSNVNKQIIDEKVYSNGVPLETIYDTLBAENTRKMDKFLDK
Opal   141  NINTKPTEQLIAICKKMRRNPIFIYVSSAYSNVNKQIIDEKTYTGVPLKTIYDTLETENTRIIDIFLDK


OnubZ  210  RPNTYTYSKALAEVLVENEFDES-AAIVRPSIIASSIREPIPGWLSGSHGFPRVVEAACKGLLLRWHGDG
OscaZ  210  RPNTYTYSKALAEVLVEKEFDES-AAIVRPSIIASSIREPIPGWLSGSHGFPRVVEAACKGLLLRWHGNG
Ofur   210  RPNTYTYSKALAEVVVEKEFDES-AAIVRPSIIASSIREPIPGWLSGSHGFPRVVEAACKGLLLRWHGDG
OnubE  210  RPNTYTYSKALAEVVVEKEFDES-AAIVRPSIIVSSIREPIPGWLSGSHGFPRVVGAACKGLLLRWHGDG
OscaE  210  RPNTYTYSKALAEVVVEKEFDES-AAIVRPSIIVSSIREPIPGWLSGSHGFPRVVGAACKGLLLRWHGDG
Ozea   210  RPNTYTYSKALAEVVVEKEFNES-AAIVRPSIILSSIREPIPGWLSGSHGFPRVVGAACKGLLLRWHGDG
Osp    211  RPNTYTYSKALAEVVVEKEFDES-AAIVRPSIIVSSIREPIPGWLSGSHGFPRVVGAACKGLLLRWHGDS
Ozag   211  RPNTYTYSKALAEVVVEKEFDES-AAIVRPSIIVSSIREPIPGWLSGSHGFPRVVGAACKGLLLRWHGDS
Olati  210  RPNTYTYSKALAEBVVGKEFDESAAAIVRPSIILSSIREPIPGWLTGSHGFPQLVGAACRGFLLRWHGDG
Opal   211  RPNTYTYSKALAEBVVGKEFDESAAAIVRPSIILSSIREPIPGWLTGSHGFPQLVGAACRGLLLRWHGDG


OnubZ  279  TVAFGIIPVDHVANLIIAAAWESNERRLIGNKGVKVYNCCSGLRNPIDVSTVMNTCLKYRKYFGT-RTMS
OscaZ  279  TVAFDIIPVDHVANLIIAAAWESNERRLIGNKGVKVYNCCSGLRNPIDVSTVMNTCLKYRKYFGT-RTMS
Ofur   279  TVAFDVIPVDHVANLIIAAAWESNERRLMGNKGVKVYNCCSGLRNPIDVSTVMNTCKYRKYFGT-RTMS
OnubE  279  TVVCDLIPVDHVANLIIAAAWESNERRLMGNKGVKVYNCCSSLRNPIDVITVMKTCLKYRKYFGT-RTMS
OscaE  279  TVVCDLIPVDHVANLIIAAAWESNERRLMGNKGVKVYNCCSSLRNPIDVITVMKTCLKYRKYFGT-RTMS
Ozea   279  TVVCDVIPVDHVANLSIAVAWESNERRLMGNKGVKVYNCCSSLRNPIDVSTVLNTCLKYRKYFGT-RTMS
Osp    280  TVVFDVIPVDHVANLIIAAAWESNERRLMGNKGVKVYNCCSGLRNPIDVSTIMNTCLKYRKYFGT-RTLS
Ozag   280  TVVFDVIPVDHVANLIIAAAWESNERRLMGNKGVKVYNCCSGLRNPIDVSTIMNTCLKYRKYFGT-RTLS
Olati  280  TVVFDVIPVDHVANLIIAAAWETNERRLVGNKGVKVYNCCSSLRNPIDVITVMNTCLKYRKYFGTERTMS
Opal   281  TVVFDTIPVDHVANLIIAAAWETNERKLVGNKGVKVYNCCSSLRNPIDVITVMNTCLKYRKYFGTERTMS


OnubZ  348  IITPRFIMKKNYFIYKLLYFTYHTIPAAIIDGFFWLTGRTPMMLNTLHKLRKLSSVLEYFTLRQFLFLDS
OscaZ  348  IITPRFIMKKNYFIYKLLYFTYHTIPAAIIDGFFWLTGRTPMMLNTLHKLRKLSSVLEYFTLRQFLFLDS
Ofur   348  IITPRFIMKKNYFIYKLLYFTYHTIPAAIIDGFFWLTGRTPMMLNTLHKLRKLSSVLEYFTLHQFLFLDS
OnubE  348  IFTPRFIMKKNYFIYKLLYFTYHTIPAAIIDGFFWLTGRTPIMLKTLDKLGKLSSVLEYFTHHQFIFLDS
OscaE  348  IFTPRFIMKKNYFIYKLLYFTCHTIPAAIIDGFFWLTGRTPIMLKTLDKLSKLSSVLEYFTHHQFIFLDS
Ozea   348  IFTPRFIMKKNYFIYKLLYFTYHTIPAAIIDGFFWLTGRTPIMLKTLDKLRKLSSVLEYFTHHQFLFLDS
Osp    349  IFPPRFIYKKNYFIYKLLYFTYHTIPAAIVDGFFWLTRRTPMMLNTLHKLSKLSSVLEYFTLHQFLFLDS
Ozag   349  IFTPRFIYKKNYFIYKLLYFTYHTIPAAIVDGFFWLTRRTPMMLNTLHKLSKLSSVLEYFTLHQFLFLDS
Olati  350  IFTPRYIMKKNYFIYKLLYFTYHTIPAAIIDAFFWITGRTPMMLKTLNKLGKLSSVLEYFTHHQFLFLDS
Opal   351  IFTPRYIMKKNYFIYKLLYFTYHTIPAAIIDAFFWLTGRTPMMLKTLNKLSNLSSVLKYFTHHQFLFLDS


OnubZ  418  NVRGLLRRMEGTDRQTFNFDVTEIEWEPELQNCVRGIANNYD---
OscaZ  418  NVRGLLRRMEGTDRQTFNFDVTEIEWEPELQNCVRGIANNYD---
Ofur   418  NVRGLLRRMSTDRQTFNFDVTEIEWEPYLQNFVRGIANNYNYSM
OnubE  418  NVRGLLRRMEGTDRQTFNFDVTEIEWEPYLQNFVRGIANNYDYSM
OscaE  418  NVRGLLRRMEGTDRQTFNFDVTEIEWEPYLQNFVRGIANNYDYSM
Ozea   418  NVRGLLRRMEGTDRQTFNFDVTEIEWEPYLQNFVRGIANNYDNSM
Osp    419  NVRGLLRRMESTDRQTFNFDVTEIEWEPYLQNCVRGIANNCDYSM
Ozag   419  NVRGLLRRMESTDRQTFNFDVTEIEWEPYLQNCVRGIANNCDYSM
Olati  420  NVRGLLRRMDDTDRQTFNFDVTEIEWEMYLQNFVRGIANNYD-SC
Opal   421  NVRGLLGRMEDIDRKTFNFDVTEIEWEMYLQNFVRGIASNYD-SC
```

Fig. S1. Alignment of the *Ostrinia* pheromone biosynthetic reductase *pgFAR*. Identical amino acid residues and conservative substitutions are shaded in black and gray, respectively.
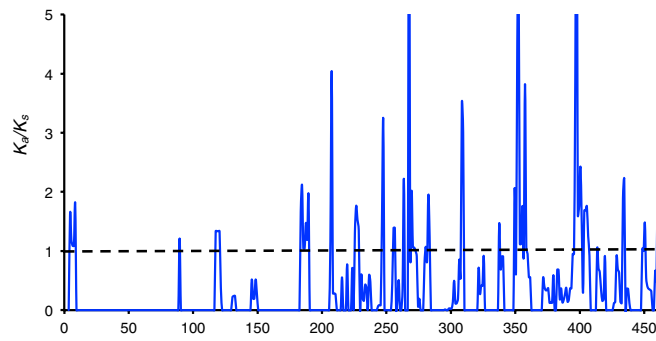
Fig. S2. Ratio of K_a/K_s calculated throughout the coding region of *pgFAR* estimated over a sliding window (window size: five codons; step size: one codon) using Li's synonymous/nonsynonymous method implemented in SWAAP 1.0.2 (1). Peaks above the dashed line indicate an excess of nonsynonymous substitutions over the neutral expectations ($K_a/K_s > 1$). Values are presented with an arbitrary cutoff of 5.

1. Li W-H (1993) Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J Mol Evol* 36(1):96–99.

Fig. S3. Sequence logo of the aligned *Ostrinia pgFAR* orthologs. The logo displays the frequencies of amino acids at each position, with the height of a letter reflecting the raw residue frequencies. Amino acids are colored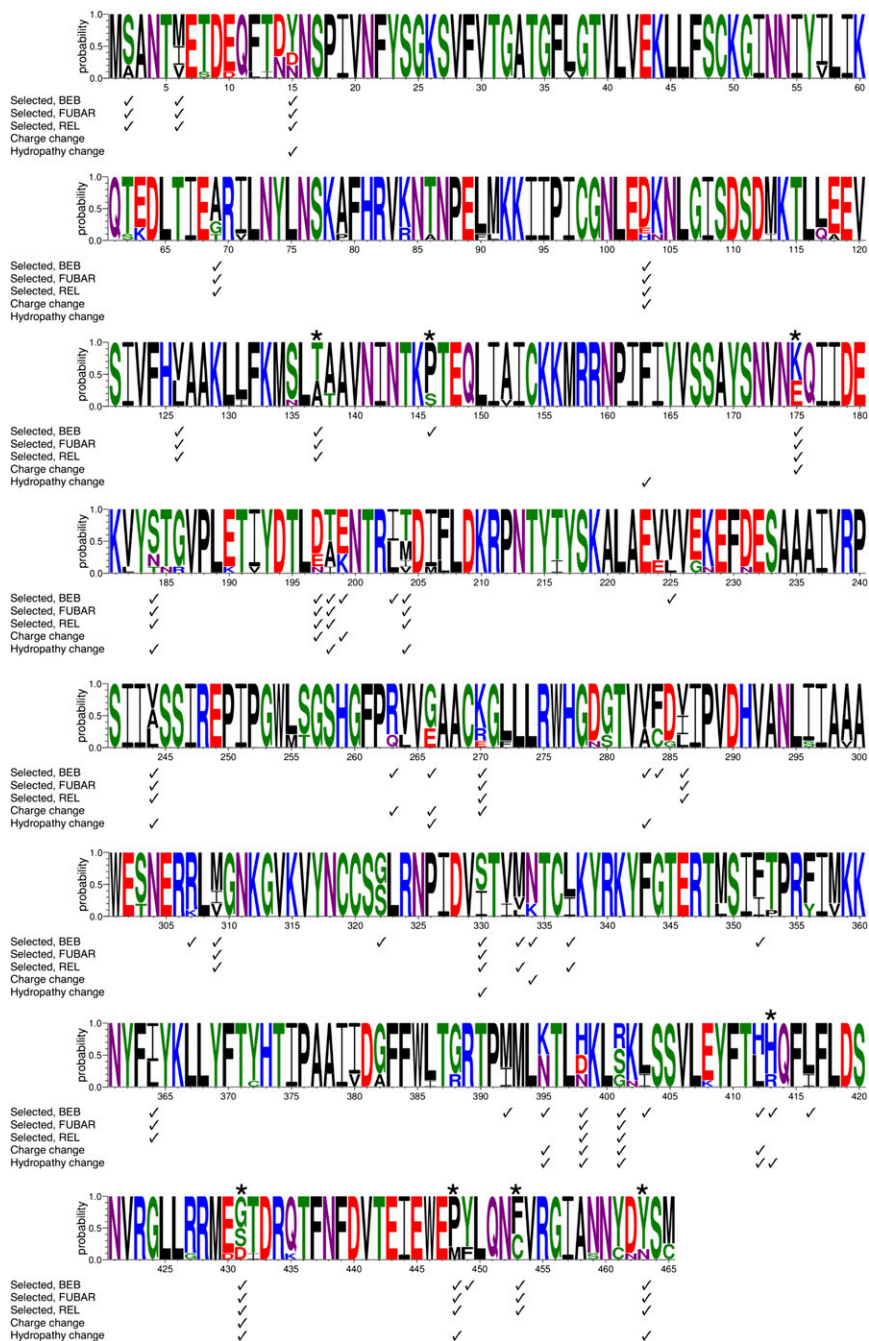 according to chemical properties (green: polar residues (G, S, T, Y, C); purple: neutral residues (Q, N); blue: basic residues (K, R, H); red: acidic residues (D, E); black: hydrophobic residues (A, V, L, I, P, W, F, M)). Positively selected sites detected using the Bayes empirical Bayes (BEB) method, the random effects likelihood method (REL), and/or the fast unbiased approximate Bayesian analysis (FUBAR) are indicated. Transitions in charge and/or hydrophobicity are indicated for the positively selected sites. Sites subjected to site-directed mutagenesis in this study are marked with asterisks. The logo was created using WebLogo 3.1 (1).

1. Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14(6):1188–1190.
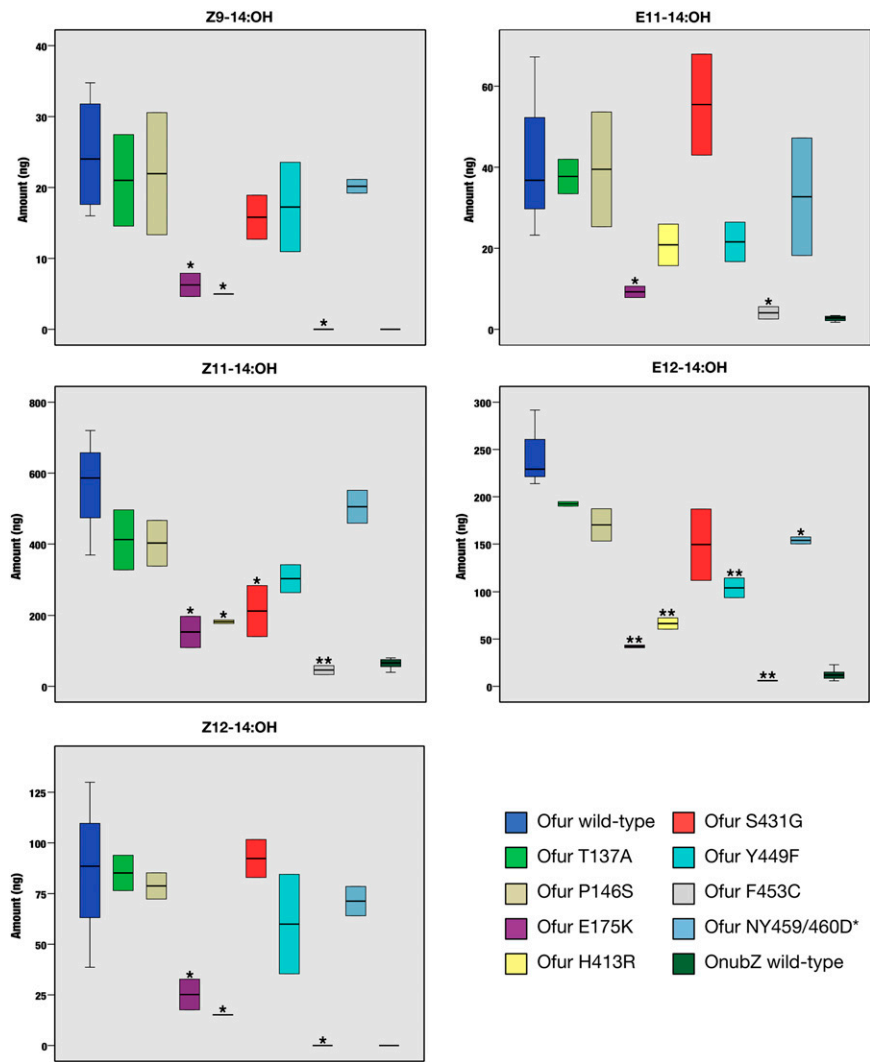
Fig. S4. Fatty alcohols produced by pgFAR WT and mutants expressed in *Saccharomyces cerevisae* supplemented with fatty-acyl precursors. Student *t* tests were used to compare the mean production of mutants with the mean production of the parental *Ofurnacalis* WT pgFAR ($n = 2–4$). The mean production of *OnubilalisZ* WT pgFAR is given for illustrative purposes. One or two asterisks denote significant difference at the 0.05 and 0.01 levels, respectively.
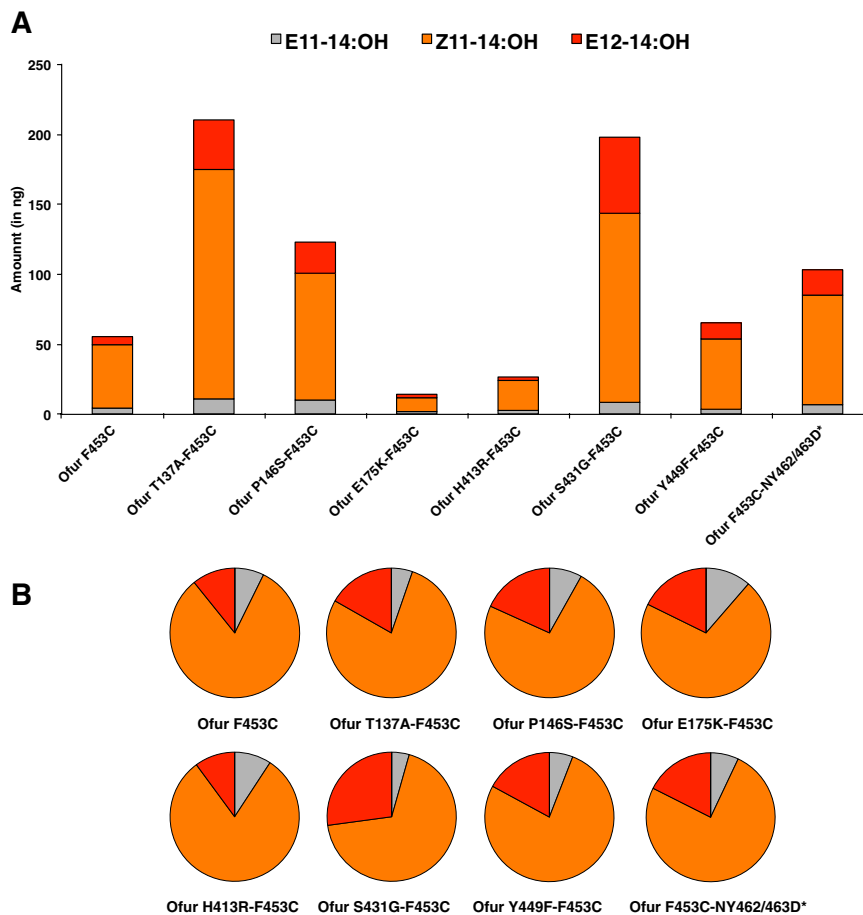
Fig. S5. Fatty alcohols produced by the *Ofurnacalis*-pgFAR^F453C mutant and double mutants. (*A*) Each bar in the stacked bar graph represents absolute amounts of fatty alcohol produced by yeast expressing each of the mutant when supplemented with fatty-acyl precursors (*n* = 2). (*B*) Relative proportions of fatty alcohol illustrating the substrate-preference of mutant enzymes.

Table S1. Pairwise distance matrix of uncorrected p-distance between the coding region of the *pgFAR* orthologs

Species	OnubZ	OscaZ	OscaE	OnubE	Ozag	Osp	Ozea	Ofur	Opal
Nucleotide									
OnubZ									
OscaZ	0.004								
OscaE	0.038	0.038							
OnubE	0.038	0.038	0.004						
Ozag	0.033	0.033	0.032	0.032					
Osp	0.030	0.030	0.031	0.031	0.004				
Ozea	0.040	0.040	0.019	0.020	0.034	0.032			
Ofur	0.018	0.018	0.033	0.033	0.030	0.028	0.036		
Opal	0.084	0.084	0.079	0.079	0.084	0.082	0.080	0.083	
Olati	0.071	0.071	0.069	0.069	0.071	0.070	0.066	0.070	0.053
Amino acid									
OnubZ									
OscaZ	0.009								
OscaE	0.079	0.077							
OnubE	0.077	0.074	0.004						
Ozag	0.070	0.068	0.066	0.066					
Osp	0.068	0.066	0.066	0.066	0.007				
Ozea	0.077	0.074	0.046	0.044	0.072	0.072			
Ofur	0.039	0.039	0.063	0.061	0.066	0.063	0.063		
Opal	0.138	0.136	0.123	0.123	0.136	0.138	0.127	0.138	
Olati	0.112	0.112	0.103	0.103	0.112	0.114	0.096	0.107	0.068

Table S2. Sites identified as having evolved under positive selection

Codon	Amino acid	BEB [posterior probability ($\omega > 1$)]*	FUBAR Posterior probability	FUBAR Empirical Bayes factor	REL Posterior probability	REL Bayes factor
2	S/A	0.838	0.831	7.004	0.756	22.187
6	M/T/V	0.840	0.796	5.574	0.904	67.572
15	Y/D/N	0.947	0.904	13.459	0.930	95.260
69	A/G/T	0.834	0.771	4.796	0.700	16.703
103	D/E/H	0.639	0.705	3.402	0.690	15.960
126	V/L	0.928	0.779	5.006	0.725	18.888
137	T/A	0.926	0.764	4.612	0.702	16.839
146	P/S	0.852	—	—	—	—
175	K/E	0.975	0.811	6.123	0.801	28.833
184	S/N/T	0.936	0.714	3.550	0.679	15.113
197	D/E/N	0.598	0.701	3.343	0.678	15.042
198	T/A/I	0.997	0.917	15.688	0.931	96.597
199	E/K	0.893	—	—	—	—
203	I/L	0.957	—	—	—	—
204	T/M/I/V	0.997	0.936	20.971	0.962	180.966
225	V/L	0.687	—	—	—	—
244	V/A/L	0.999	0.858	8.573	0.885	54.879
263	R/Q	0.739	—	—	—	—
266	G/E	0.938	—	—	—	—
270	K/R/E	0.725	0.756	4.421	0.705	17.096
283	V/A	0.951	—	—	—	—
284	F/C	0.952	—	—	—	—
286	V/I/L	0.986	0.912	14.812	0.932	98.537
307	R/K	0.511	—	—	—	—
309	M/I/V	0.869	0.712	3.520	0.660	13.921
322	G/S	0.989	—	—	—	—
330	S/I	0.944	0.738	4.011	0.705	17.131
333	M/V/L	0.978	—	—	0.671	14.615
334	N/K	0.606	—	—	—	—
337	L/I	0.955	—	—	0.706	17.205
352	F/I	0.925	—	—	—	—
364	I/L	0.991	0.787	5.258	0.738	20.146
392	M/I	0.769	—	—	—	—
395	K/N	0.982	—	—	—	—
398	H/D/N	0.990	0.731	3.864	0.702	16.837
401	R/S/G	0.999	0.940	22.251	0.971	236.251
403	L/I	0.819	—	—	—	—
412	H/L	0.992	—	—	—	—
413	H/R	0.798	—	—	—	—
416	L/I	0.819	—	—	—	—
431	G/S/D	0.950	0.862	8.861	0.921	83.162
448	P/M	0.790	0.761	4.523	0.690	15.951
449	Y/F	0.778	—	—	—	—
453	F/C	0.936	0.735	3.951	0.697	16.484
463	Y/N/*	0.805	0.797	5.586	0.754	21.990

*Positively selected sites identified using site model M8 in PAML and the Bayes empirical Bayes procedure.

Table S3. Likelihood ratio tests for positive selection using PAML site-specific models for *Ostrinia pgFAR* genes

Model	Parameter estimates	Parameters	lnL	Likelihood ratio test
M1a: nearly neutral	$p_0 = 0.64648$ $p_1 = 0.35352$ $\omega_0 = 0.02327$ $\omega_1 = 1.00000$	2	−3,604.518	
M2a: selection	$p_0 = 0.87127$ $p_1 = 0.00000$ $p_2 = 0.12873$ $\omega_0 = 0.18631$ $\omega_1 = 1.00000$ $\omega_2 = 3.59987$	4	−3,581.530	M1a vs. M2a 45.98*
M7: neutral, beta	$P = 0.00531$ $q = 0.00702$	2	−3,604.833	
M8: selection, beta + $\omega > 1$	$p_0 = 0.87170$ $p_1 = 0.12830$ $P = 22.86913$ $q = 99.00000$ $\omega = 3.60710$	4	−3,581.564	M7 vs. M8 46.54*
M8a: neutral, beta + $\omega = 1$	$p_0 = 0.64528$ $p_1 = 0.35472$ $P = 2.31757$ $q = 99.00000$ $\omega = 1.00000$	4	−3,604.537	M8a vs. M8 45.95*

The site-specific models implemented in Codeml (PAML 4.4 package) were used to test for positive selection acting on the *pgFAR* open-reading frames. The log likelihoods (lnL) of each model were compared in likelihood ratio tests. *P* values were determined from χ^2 distributions.
*Significant at the 1% level.