PNAS RESEARCH ARTICLE | NEUROSCIENCE

Exitron splicing of odor receptor genes in *Drosophila*

Xueying Shang[a], Gaëlle J. S. Talross[a], and John R. Carlson[a,1]

Contributed by John R. Carlson; received November 17, 2023; accepted February 6, 2024; reviewed by Christopher J. Potter and Michael Rosbash

Proper expression of odor receptor genes is critical for the function of olfactory systems. In this study, we identified exitrons (exonic introns) in four of the 39 *Odorant receptor* (*Or*) genes expressed in the *Drosophila* antenna. Exitrons are sequences that can be spliced out from within a protein-coding exon, thereby altering the encoded protein. We focused on *Or88a*, which encodes a pheromone receptor, and found that exitron splicing of *Or88a* is conserved across five *Drosophila* species over 20 My of evolution. The exitron was spliced out in 15% of *Or88a* transcripts. Removal of this exitron creates a non-coding RNA rather than an RNA that encodes a stable protein. Our results suggest the hypothesis that in the case of *Or88a*, exitron splicing could act in neuronal modulation by decreasing the level of functional *Or* transcripts. Activation of *Or88a*-expressing olfactory receptor neurons via either optogenetics or pheromone stimulation increased the level of exitron-spliced transcripts, with optogenetic activation leading to a 14-fold increase. A fifth *Or* can also undergo an alternative splicing event that eliminates most of the canonical open reading frame. Besides these cases of alternative splicing, we found alternative polyadenylation of four *Ors*, and exposure of Or67c to its ligand ethyl lactate in the antenna downregulated all of its 3′ isoforms. Our study reveals mechanisms by which neuronal activity could be modulated via regulation of the levels of *Or* isoforms.

olfaction | *Drosophila* | splicing | exitron | post-transcriptional regulation

Significance

We found exitrons (exonic introns) in four *Odorant receptor* (*Or*) genes of *Drosophila*. Exitrons are sequences that can be spliced from within a protein-coding exon, thereby altering the encoded protein. We are unaware of previous reports of exitrons in *Drosophila*. Removal of the exitron from the pheromone receptor gene *Or88a* creates a non-coding transcript. The exitron has been conserved for 20 My. Activation of *Or88a*-expressing olfactory receptor neurons via either optogenetics or pheromone stimulation increased the level of these non-coding transcripts, suggesting that exitron splicing could act in neuronal modulation. Furthermore, our investigation unveiled the presence of alternative polyadenylation in four *Drosophila Ors*. This study reveals a previously unrecognized layer of complexity in the regulation of *Or* gene expression.

Odor receptors signal the presence of food, mates, and predators. In *Drosophila*, a family of 60 *Odorant receptor* genes encodes seven-transmembrane-domain proteins, of which one is typically expressed per olfactory receptor neuron (ORN) of the antenna (1–4). The Or forms a heteromer with the co-receptor Orco and imparts response characteristics to the ORN.

The regulation of *Or* gene expression underlies the functional organization of the olfactory system, and by extension, odor coding. Transcriptional regulation of *Or* genes has been studied in detail, largely to elucidate the process by which an individual ORN selects a single receptor to express from among the repertoire of 60 (5–9). Post-transcriptional regulation of *Or* genes, by contrast, has received very little attention.

Post-transcriptional regulation can generate multiple RNAs and proteins from a single gene. It can also modulate the level, localization, or translation of mRNAs (10, 11). One mode of post-transcriptional regulation, alternative splicing, has been reported in *Or* genes (3), but in only two of the 60 *Or* genes of *Drosophila* (*SI Appendix*, Fig. S1*A*). Another mechanism of post-transcriptional regulation, alternative polyadenylation, has not been reported in *Or* genes of *Drosophila* to our knowledge.

Exitron splicing is a particularly intriguing mode of alternative splicing. Exitrons (exonic introns) are sequences that can be spliced from within a protein-coding exon, thereby altering the encoded protein (12). Exitrons are removed from only a subset of transcripts. Exitrons were first discovered in plants and have subsequently been found in mice and humans (12–14), but exitrons have not been reported in *Drosophila* or other insects to our knowledge.

We have capitalized on the recent availability of a wealth of antennal RNA-Seq data (15) to explore post-transcriptional regulation of *Or* expression. Although alternative splicing of *Ors* was believed to be rare, we identified exitron splicing in four of the 39 *Ors* expressed in the adult antenna. Particularly interesting is *Or88a*, which encodes a pheromone receptor (16). An exitron is removed from 15% of *Or88a* transcripts, removing coding sequences and severely disrupting their coding potential. This exitron has been conserved for 20 My of evolution. Epitope-tagging experiments suggest that exitron removal creates a non-coding RNA rather than an mRNA encoding a stable protein. Activation of *Or88a*-expressing ORNs via either optogenetics or pheromone stimulation increased the level of exitron-spliced transcripts. These results suggest the possibility that in some cases exitron splicing could act in neuronal modulation. A fifth *Or* can undergo a different kind of alternative splicing event that removes most of the open reading frame (ORF). We also found alternative polyadenylation of four *Ors*, and exposure of Or67c to its ligand ethyl lactate downregulated all of its 3′ isoforms. Finally, we discuss how these mechanisms could provide a means of tuning neuronal activity in the olfactory system.

Author affiliations: [a]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06511

Author contributions: X.S., G.J.S.T., and J.R.C. designed research; X.S. and G.J.S.T. performed research; G.J.S.T. contributed new reagents/analytic tools; X.S. and G.J.S.T. analyzed data; and X.S., G.J.S.T., and J.R.C. wrote the paper.

Reviewers: C.J.P., Johns Hopkins University School of Medicine; and M.R., Howard Hughes Medical Institute.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: john.carlson@yale.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2320277121/-/DCSupplemental.

Published March 20, 2024.

Results

Identification of Exitrons in *Ors*. During close examination of strand-specific rRNA-depleted antennal RNA-Seq data (15) we identified reads in which regions of protein-coding exons were unexpectedly missing from *Or* genes (red box, $rRNA^-$ in Fig. 1*A*). Such reads were found in four *Or* genes, *Or88a, Or82a, Or92a, and Or23a* (Fig. 1*B*). To investigate whether these putative alternatively spliced transcripts are polyadenylated, we carried out a strand-specific $poly(A)^+$-selected antennal RNA-Seq experiment. We identified reads lacking the exact same exonic region for all four *Or* genes, suggesting that these transcripts are indeed polyadenylated [red box, $poly(A)^+$ in Fig. 1*A*].

To test whether these reads might represent exitrons as opposed to artifacts generated during the construction of the different RNA-Seq libraries, we first performed RT-PCR analysis using primers that were designed to specifically amplify transcripts from which exitrons had been spliced (Fig. 1*C* and *SI Appendix*, Fig. S1*B*). In all four cases, amplification from antennal cDNA yielded products of length and sequence expected of an exitron splicing event. As a control, products of these sizes were not amplified when genomic DNA (gDNA) was used as template (Fig. 1*C*).

We note that an annotated antisense RNA, *asRNA:CR44237*, overlaps with *Or88a*. Three lines of evidence indicate that the unexpected reads we have found derive from *Or88a* and not *asRNA:CR44237*: i) *Or88a* and *asRNA:CR44237* are in opposite orientations, and the orientation of the unexpected reads is the same as that of *Or88a* (Fig. 1*A*); ii) we carried out a 5′RACE experiment to determine the 5′ end of *Or88a* using an oligonucleotide primer downstream of the putative exitron (*SI Appendix*, Fig. S1*C*). We identified two cDNA products: one canonical and one spliced (*SI Appendix*, Fig. S1*C*), consistent with splicing of *Or88a* transcripts. iii) *asRNA:CR44237* lacks canonical splice sites at the corresponding sites.

We next considered other possible sources of artifacts. Reverse transcriptase (RT) can skip over a stable structured RNA sequence (18) such as a hairpin (inverted repeat) structure, which can be exacerbated by the presence of direct repeats at the base of the structure (17) (*SI Appendix*, Fig. S2 *A* and *B*). This skipping, or "slippage," generates cDNA in which a sequence of the transcript is removed, a sequence known as a false intron or "falsitron" (17). We analyzed the sequences removed from the four *Or* genes and found that they share characteristics with exitrons (12) rather than falsitrons (17), in that:

(i) In all four genes, we observed precise boundaries between the constitutive exonic sequence and the removed sequences, as shown by all of 79 reads (Fig. 1*A*; 11, 21, 18, and 29 reads for *Or88a*, *Or82a, Or92a, and Or23a*, respectively). By contrast, the boundaries of falsitrons are often imprecise, showing variation in the locations of their junctions (17).

(ii) All four of the sequences are removed at the canonical splicing site GT/AG (Fig. 1*B*), unlike many falsitrons [*SI Appendix*, Fig. S2*B* and (17)]. The mean splicing site prediction score of the four removed sequences is much higher than that for a collection of 29 analyzed falsitrons (17), most of which were scored as 0 (Fig. 1*D*). Moreover, the mean splicing score of the four removed *Or* sequences is somewhat lower than that of the constitutive *Or* introns within these four genes, again reminiscent of exitrons (12) (Fig. 1*D*).

(iii) All four removed sequences lack predicted hairpin structures and direct repeats (≥4 nucleotides in length) that overlap or abut their stems (*SI Appendix*, Fig. S3).

The simplest interpretation of these data taken together is that the removed *Or* sequences are exitrons. Several additional lines of evidence that support the identity of these sequences as exitrons are described below.

In addition to these four genes, we found that a fifth *Or* gene, *Or35a*, undergoes splicing at an alternative 5′ splice site, thereby removing most of the canonical open reading frame (ORF). This event is distinct from the other four in that the 3′ splice acceptor site is at the 5′ end of a downstream exon, as opposed to lying within the same exon as the 5′ donor site (Fig. 1*B*). Thus we do not refer to this removed sequence as an exitron. Also this splicing uses the non-canonical splice site GC/AG, which has been found in the first intron of a number of non-coding genes (19). However, this splicing of *Or35a* is unlike the previously reported alternative splicing of *Or46a* and *Or69a* (*SI Appendix*, Fig. S1*A*), in that it removes sequences from the canonical open reading frame (ORF), as does the splicing of the four exitrons (Fig. 1*E*).

In three cases, *Or88a, Or82a,* and *Or92a,* the length of the exitron is not divisible by three, and its removal leads to a frame shift. For example, splicing of the *Or88a* exitron (82 nucleotides) leads to a frameshift that creates a stop codon 283 bases downstream of the ATG (Fig. 1*E*). Thus the resulting ORF, ORF1, encodes a predicted polypeptide of 67 amino acids, of which the first 41 agree precisely with those of the canonical Or88a receptor and represents its intracellular N terminus, and the remaining 26 amino acids would differ from those of Or88a. In addition, the exitron-spliced transcript of *Or88a* contains another in-frame ORF, ORF2, that starts 80 nucleotides after the exitron 3′ splice site and stops at the canonical stop codon (Fig. 1*E*). ORF2 is predicted to encode a protein of 306 amino acids that agrees with amino acids 96 to 401 of Or88a, which encode the last five transmembrane domains of the receptor and the extracellular C terminus.

Exitron-splicing of two other *Ors*, *Or82a* and *Or92a*, also leads to a frameshift and two ORFs, one of which would again encode a relatively short N-terminal fragment and one of which would again encode the last five transmembrane domains and the C terminus (Fig. 1*E*). In the other two cases, *Or23a* and *Or35a*, the removed sequences are much longer and their length is divisible by three, indicating that the splicing is predicted not to cause a frameshift but to generate shorter proteins than the canonical mRNAs (i.e., the mRNAs that have not been spliced in this manner) (Fig. 1*E*).

To further investigate exitron splicing in *Drosophila Ors,* we focused on *Or88a*, a particularly interesting *Or* that encodes a pheromone receptor. To determine the relative level of the *Or88a* exitron-spliced transcript, we performed RT-qPCR experiments with primers that specifically amplify the *Or88a* canonical transcript and with primers that amplify both transcripts (Fig. 2 *A*, *Left*). We determined the amount of exitron-spliced transcript by subtracting the amount of canonical transcript from the amount of both transcripts. Our results revealed that the exitron was spliced from 15% of *Or88a* transcripts (Fig. 2 *A*, *Right*; see additional data below).

The *Or88a* Exitron Has Been Conserved for 20 My. We next asked whether splicing of the *Or88a* exitron is conserved across *Drosophila* species. If so, this would suggest that exitron splicing serves a valuable function. We tested four other species, *D. sechellia, D. simulans, D. erecta,* and *D. suzukii,* all of which contain sequences that are similar to that of the *D. melanogaster* exitron (Fig. 2*B*). To determine whether these sequences are in fact removed from some transcripts of the other species, we carried out two experiments.

First, we hand-dissected antennae from each species, extracted RNAs, and performed RT-PCR using primers that flanked the exitron sequence (Fig. 2*B*). We cloned the amplified products,

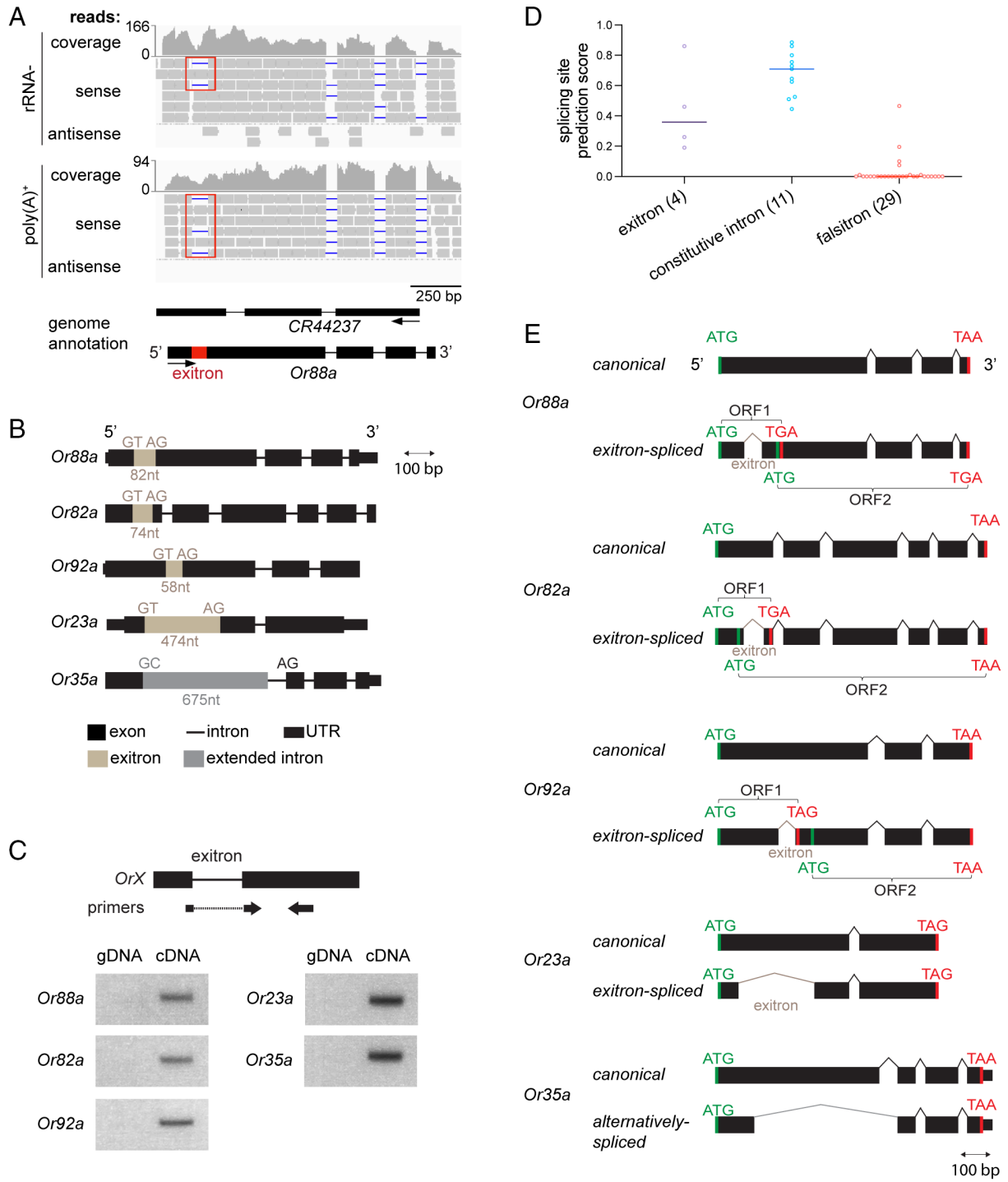

Fig. 1. Identification of exitron splicing in *Odor receptors* (*Ors*) in *Drosophila*. (*A*) RNA-Seq of *Drosophila* antennae demonstrating splicing of an exitron in *Or88a*. Representative reads that reflect exitron splicing are highlighted in the red boxes. These reads were detected both by strand-specific rRNA-depleted RNA-Seq (top tracks) and strand-specific poly(A)$^+$-selected RNA-Seq (bottom tracks). (*B*) Gene models for the *Ors* in *D. melanogaster* antennae that we found to be alternatively spliced based on RNA-Seq results. Exitrons, exitron splice sites, and exitron lengths are shown in sand color. The extended intron of *Or35a* is shown in gray. (*C*) RT-PCR results using exitron-spliced transcript-specific primers for *Or88a, Or82a, Or92a,* and *Or23a*. The forward primers are designed to anneal to two regions in the first exon: a region immediately upstream of the exitron 5′ splice site and a region immediately downstream of the exitron 3′ splice site. An analogous experiment is shown for *Or35a*. gDNA: genomic DNA extracted from whole flies. cDNA: cDNA from antennal RNA. (*D*) The splicing site prediction scores of exitrons, constitutive introns from *Or88a, Or92a, Or82a, Or23a,* and 29 falsitrons (17) were analyzed. Splicing site prediction scores were calculated as the average of the 5′ and 3′ splice site scores determined with NNSPLICE 0.9. Among the 29 falsitrons, 22 showed a prediction score of 0 for both the 5′ and 3′ splice sites, and the other 7 had a prediction score of 0 for either the 5′ or 3′ splice site. The mean of the falsitron scores is 0.03 and cannot be seen in the graph. (*E*) Canonical ORFs of the *Ors* and the predicted ORFs of their exitron-spliced transcripts. Start codons and stop codons are shown in green and red, respectively.

sequenced them and in each case identified at least four out of 20 to 50 sequenced plasmids that contain products consistent with the removal of an exitron (i.e., the products lacked the nucleotides highlighted in yellow in Fig. 2*B*). The position and length of exitrons were conserved, except that *D. erecta* had a shorter exitron (59 bp instead of 82 bp).

To confirm these results, and to safeguard against artifacts from the reverse transcription step, we designed primers specific to the

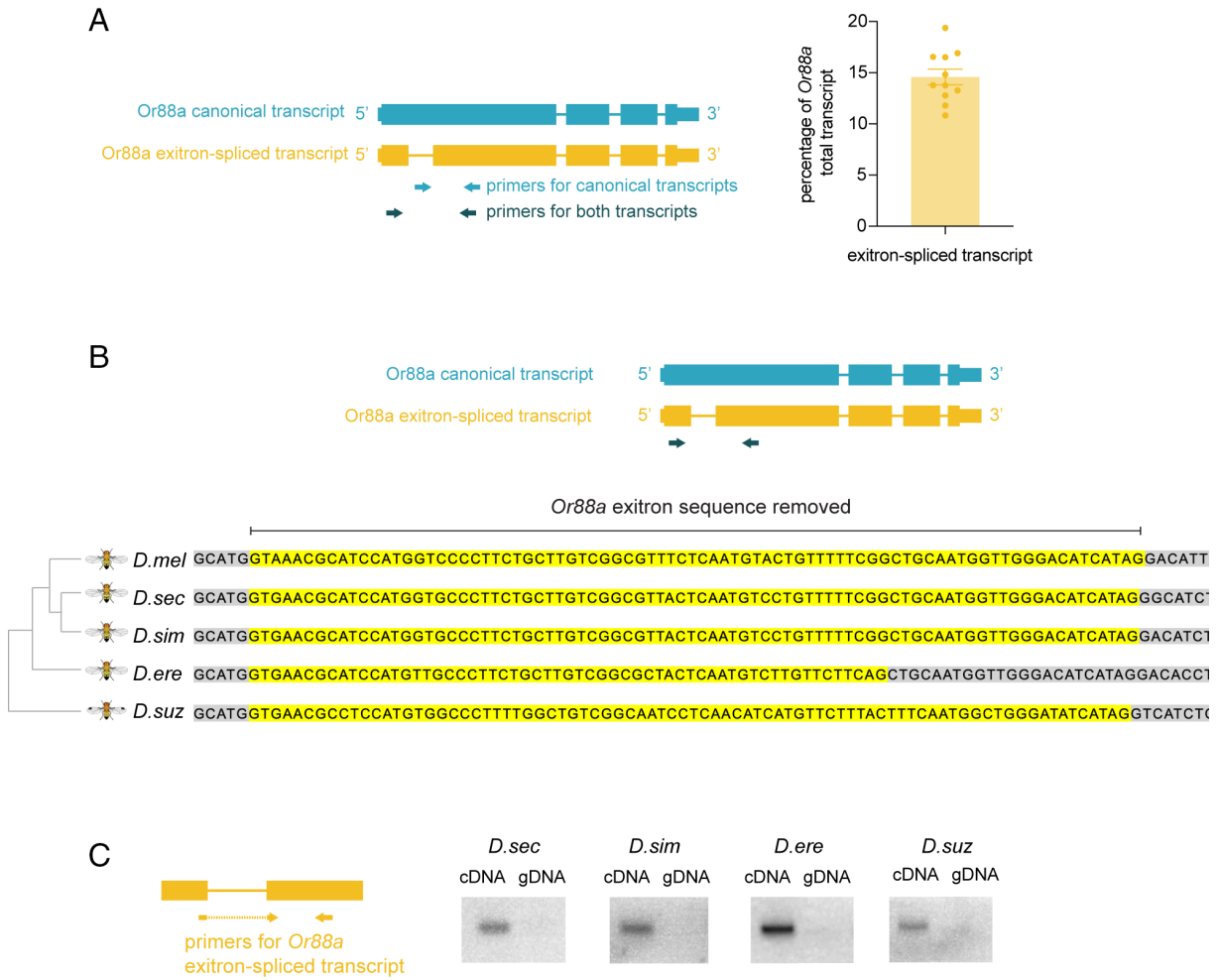

Fig. 2. Exitron splicing of *Or88a* is conserved in five *Drosophila* species**.** (*A*) *Left*: gene models of the *Or88a* canonical transcript and the *Or88a* exitron-spliced transcript and primers designed to distinguish the *Or88a* transcripts. *Right*: percentage of *Or88a* transcripts from which the exitron has been removed. n = 11 biological replicates. Error bar is SEM. The forward primer for the canonical transcript is designed within the exitron sequence. Primers for total *Or88a* transcripts are designed within sequences shared by both the canonical transcript and the exitron-spliced transcript. (*B*) Antennal cDNA from five *Drosophila* species was amplified using primers that target both the *Or88a* exitron-spliced transcript and the canonical transcript. The resulting fragments were sequenced after ligation to a plasmid. At least four out of 20 to 50 sequenced plasmids from each species were identified that revealed removal of the exitron. The exitron sequences from all five species are highlighted in yellow, and the flanking sequences are shown in gray. (*C*) RT-PCR analysis of antennal RNA samples extracted from *Drosophila* species using primers specific for the exitron-spliced transcript of *Or88a*. Primer design is shown on the left. cDNA: cDNA from antennal RNA. gDNA: genomic DNA extracted from whole flies.

exitron-spliced transcript of each species (Fig. 2*C*). Using these primers in an antennal RT-PCR experiment, we again found transcripts from which the exitrons had been removed, with lengths consistent with those of the products found with the first set of primers used in Fig. 2*B*. As a further control, these products were not amplified from genomic DNA (Fig. 2*C*).

We note that in all species, the exitron is removed at the canonical splicing site GT/AG—including *D. erecta*, in which the 3′ end of the exitron is different from those of the other species. This conserved use of the GT/AG splice site in *D. erecta* and the other species lends further support to the identification of the removed sequences as exitrons.

In summary, exitrons have been conserved at the same position within the genes of all five species, and their removal by splicing has been conserved for ~20 My (20), suggesting that they serve an important biological function.

The Exitron-Spliced Transcript of *Or88a* Is a Non-Coding RNA. We next asked whether the *Or88a* transcript from which the exitron has been removed is translated into a stable protein. We used a homozygous $Or88a^{GAL4}$ knock-in construct to drive expression of two *UAS-Or88a* constructs. In the first construct, the AG from the splice acceptor site was mutated to TG to prevent exitron splicing (Fig. 3*A*). We refer to the resulting transcript as the "canonical" transcript for convenience. To this first construct, an HA tag was added, initially to the 3′ end, immediately upstream of the last stop codon of *Or88a*. In the second *UAS-Or88a* construct, the exitron was deleted (Fig. 3*C*). We refer to this transcript as the "exitron-spliced" transcript. To investigate whether the longest putative ORF, ORF2, is translated to produce a receptor containing five of the seven transmembrane domains, we added an HA tag to its 3′ end. We then generated $Or88a^{GAL4}$*; UAS* lines for both *UAS* constructs and confirmed via RT-PCR that they produced transcripts of the expected sizes (*SI Appendix*, Fig. S4*A*). We then carried out a double-labeling experiment, using in situ hybridization to detect the transcripts and immunofluorescence to detect the HA tags.

The canonical *Or88a* transcript was found in a subset of ORNs, as expected (Fig. 3*B*). Expression of the HA tag was also detected, in the same cells that were labeled with the *Or88a* in situ hybridization probe (probe design shown in *SI Appendix*, Fig. S4*B*). The transcript was localized in cell bodies, whereas the HA tag was observed in both cell bodies and dendrites, as found for anti-Or antibodies (21).

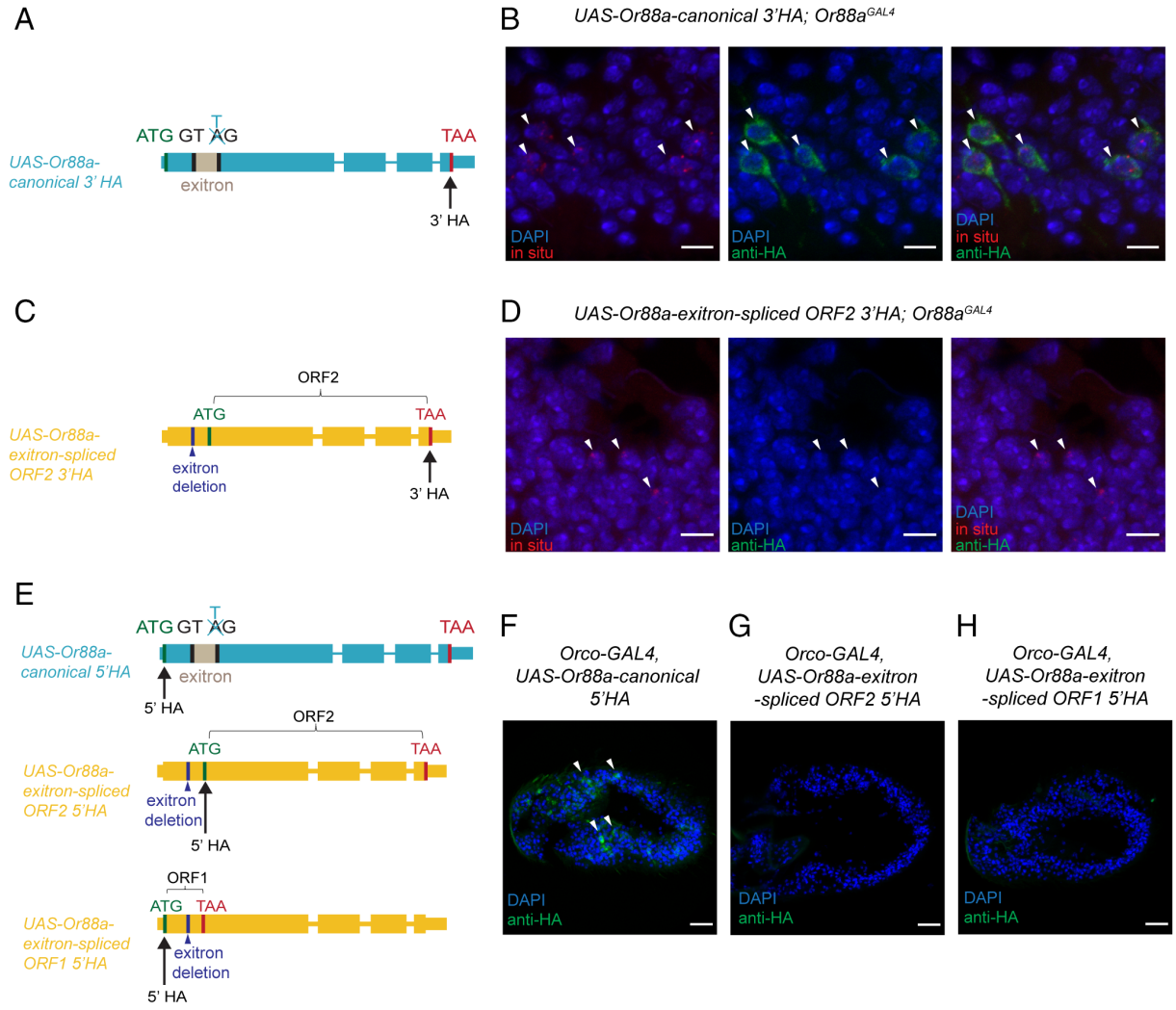

Fig. 3. The exitron-spliced transcript of *Or88a* is a non-coding RNA. (*A*) Diagram of the *UAS-Or88a-canonical 3′HA* construct. The exitron 3′ acceptor site is mutated from AG to TG while preserving the canonical amino acid sequence of Or88a. An HA tag is inserted at the 3′ end immediately upstream of the stop codon. (*B*) Dual labeling of the antenna of a homozygous *Or88a* knock-in *GAL4* construct (*Or88a*^{GAL4}) driving the expression of the *UAS-Or88a-canonical 3′HA* construct. White arrows point to in situ hybridization signals and anti-HA signals. Red: in situ hybridization using a ~500-nt probe targeting the common region (*SI Appendix*, Fig. S4*B*) of the *Or88a* exitron-spliced transcript and canonical transcripts; Blue: DAPI staining showing localization of nuclei; Green: immunofluorescent labeling of HA tags. (Scale bars: 5 μm.) (*C*) Diagram of the *Or88a UAS-exitron-spliced ORF2 3′HA* construct. In this construct, the exitron region is deleted, and an HA tag is inserted at the 3′ end immediately upstream of the stop codon of the predicted ORF2 of the *Or88a* exitron-spliced transcript. (*D*) Dual labeling of *Or88a*^{GAL4} driving the *UAS-exitron-spliced ORF2 3′HA*. White arrows point to in situ hybridization signals and anti-HA signals. Red: in situ hybridization using the same probe as in Fig. 3*B*; Blue: DAPI staining showing localization of nuclei. Green: immunofluorescence on HA tags. (Scale bars: 5 μm.) (*E*) Diagrams of the *UAS-Or88a-canonical 5′HA* construct and two different *UAS-exitron-spliced 5′HA* constructs, each containing an HA tag at the 5′ end of a different predicted ORF. For the *UAS-Or88a-canonical 5′HA* construct, the exitron 3′ acceptor site is mutated from AG to TG, and an HA tag is inserted at the 5′ end immediately upstream of the start codon (ATG-HA-endogenous ATG). For the *UAS*-exitron-spliced 5′HA constructs, the exitron region is deleted, and an HA tag is inserted at the 5′ immediately upstream of the predicted ORF2 or ORF1. (*F–H*) Immunofluorescence of antennae in which *Orco-GAL4* drives the *UAS-Or88a-canonical 5′HA* and the *UAS-Or88a-exitron-spliced* constructs. White arrows indicate the anti-HA signals. (Scale bar: 10 μm.)

The exitron-spliced transcript was also detected in the cell bodies of a subset of ORNs (Fig. 3*D*). However, in this case, no anti-HA labeling was observed. We also used an *Orco-GAL4* driver to express both constructs in a broad population of ORNs. Again, expression of the HA tag was detected in the case of the canonical *Or88a* transcript but not the exitron-spliced transcript (*SI Appendix*, Fig. S4 *C* and *D*).

These results using the 3′ HA-tagged constructs suggested that the exitron-spliced transcript is not translated. However, it seemed possible that its C-terminal HA tag is not accessible to labeling, or perhaps is removed as part of a post-translational modification. Therefore, we also generated constructs in which the HA tag was added to the 5′ end of the canonical *Or88a* transcript, and the 5′ end of ORF2 of the exitron-spliced transcript (Fig. 3*E*). When we expressed these constructs with the *Orco-GAL4* driver, we detected anti-HA labeling with the *UAS-Or88a-canonical 5′HA*, but not the 5'HA tagged ORF2 of the exitron-spliced construct (Fig. 3 *F* and *G*).

We also investigated by HA-tagging whether ORF1, the shorter ORF of the *Or88a* exitron-spliced transcript, is translated into a stable peptide. We limited our analysis to the addition of a 5′ HA tag, since experiments with a 3′HA tag could be complicated to interpret: The C termini of many non-canonical peptides promote their degradation, and thus adding a 3′ HA tag could lead to artificial stabilization of the peptide (22). Accordingly, we expressed the 5′ HA-tagged construct with *Orco-GAL4* and did not observe anti-HA labeling with the construct (Fig. 3*H*).

The simplest interpretation of these results, taken together, is that the exitron-spliced transcripts are not translated into stable proteins; rather, they are non-coding RNAs.

Neuronal Activity Increases the Level of *Or88a* Exitron-Spliced Transcript. We wondered whether the level of exitron-spliced transcript might be affected by neuronal activity. There is precedent for modulation of alternative splicing patterns by neuronal activity in other systems (23, 24), and levels of some exitron-spliced transcripts are affected by stress in *Arabidopsis* (12). Accordingly, we used two different means to activate the ORNs that express *Or88a* and then measured the levels of the exitron-spliced transcript.

First, we expressed the channelrhodopsin CsChrimson in at4C neurons, the neurons that express *Or88a*, using the *Or88a-GAL4* driver. We then exposed flies to two different intensities of red light and carried out RT-qPCR experiments (Fig. 4*A*). As controls, we also measured the expression level of *Or88a* transcripts in flies that were not exposed to red light or that were not fed the all-*trans*-retinal (ATR) co-factor that is essential for CsChrimson function (25). We note that the *Or88a-GAL4* driver, like all other constructs used in this study, was backcrossed at least five generations to w^{1118} *Canton-S* before using, in order to minimize genetic background effects.

We used three sets of primers to determine the level of *Or88a* exitron-spliced transcripts, canonical transcripts, and both transcripts, relative to transcripts of the *Or* co-receptor *Orco* (Fig. 4*B*). The specificity of the primers was confirmed by the PCR using as templates plasmids that contain sequences representing either the canonical transcript or the exitron-spliced transcript (*SI Appendix*, Fig. S1*B*).

Interestingly, we detected increased levels of exitron-spliced RNA in flies exposed to red light (Fig. 4*C*). The mean level of exitron-spliced RNA was ~14-fold greater in flies illuminated with 0.14-mW/mm^2 red light, compared to flies not fed ATR or not exposed to red light. Levels of canonical *Or88a* transcripts, however, remained unaffected by red illumination. The level of total *Or88a* transcripts, i.e., both types combined, as determined with primers that recognize both transcripts, increased moderately following illumination with 0.14-mW/mm^2 red light. We note that each data point in this experiment is based on ~130 hand-dissected antennae (i.e., the entire experiment is based on ~10,000 hand-dissected antennae).

In control flies that were not fed the ATR co-factor, the exitron-spliced transcript accounted for 14.6% of total *Or88a* transcripts, in good agreement with the value of 15% found in the experiment shown in Fig. 2*A*. By contrast, in flies exposed to 0.14-mW/mm^2 red light, the exitron-spliced transcript accounted for 54% of total *Or88a* transcripts.

Longer exposure to red light (24 h) led to a comparable increase (8×) in the level of exitron-spliced transcripts (*SI Appendix*, Fig. S5*A*). In an independent experiment, we again exposed flies to 0.14 mW/mm^2 red light for 5 h and found a 12-fold increase in levels of exitron-spliced *Or88a* (Fig. 4*D*) similar to the results shown in Fig. 4*C*. In parallel, we exposed flies to the same red light and then allowed them to recover in the dark for 24 h. The level of the exitron-spliced transcript then appeared to have declined back to control levels (Fig. 4*D*).

We next switched from optogenetic stimulation of ORNs to odorant stimulation, which activates the ORNs via odor receptors, as in a natural context. *Or88a* is activated by the pheromone methyl myristate (16). Exposure to high odor concentrations has previously been found to either increase or decrease levels of *Or* expression, depending on the odorant and the *Or* (26, 27). Using conditions established in these prior studies, we exposed flies to 5% methyl myristate for 5 h, and compared levels of *Or88a* transcripts to those of flies exposed to the solvent only (Fig. 4*E*; we note that methyl myristate has a molecular weight of 242 g/mol and is of limited volatility.) We again found increased levels of the exitron-spliced *Or88a* transcript (Fig. 4*F*). There was a decrease in levels of the canonical *Or88a* mRNA. As a control, we exposed flies to an odorant that does not activate *Or88a*-expressing ORNs, geranyl acetate (GA), and found no effect on the level of either *Or88a* transcript (*SI Appendix*, Fig. S5*B*). Nor did methyl myristate have an effect on the expression of transcripts of another *Or* gene, *Or82a* (*SI Appendix*, Fig. S5*C*).

Taken together, these results suggest that activation of *Or88a*-expressing ORNs increases the level of the exitron-spliced *Or88a* transcript, but not the canonical transcript.

Transcript Levels and Olfactory Physiology. We asked whether the exitron-spliced transcript might actively downregulate the level of the canonical transcript. First, we overexpressed the exitron-spliced transcript (Fig. 5*A*) in at4C neurons with an *Or88a-GAL4* driver. Although the level of the exitron-spliced transcript was 12-fold higher than in an *Or88a-GAL4* parental control, the level of the canonical transcript was unaffected (Fig. 5*B*). Reciprocally, when we overexpressed the canonical RNA, the level of exitron-spliced transcript was unaffected. These results argue against a model in which the exitron-spliced transcript negatively regulates the level of the Or88a receptor by actively downregulating the level of the mRNA that encodes it.

We next wondered whether expression of the exitron-spliced transcript might regulate the translation of the canonical *Or88a* transcript, or affect the function of at4C ORNs by other mechanisms. We carried out electrophysiological recordings from at4 sensilla in a series of genotypes to measure the response of the at4C neuron to the pheromone methyl myristate. The at4 sensillum contains three ORNs (Fig. 5*C*). In principle, the spikes of the at4C neuron are distinguishable by their small size from those of the other two ORNs in this sensillum, but in practice, we were not able to distinguish them with high confidence in all cases; accordingly, we summed all the spikes in each of our recordings.

First, we tested the homozygous $Or88a^{GAL4}$ knock-in construct with no *UAS* constructs and confirmed that in this mutant there was little if any response to methyl myristate (Fig. 5*D* and *SI Appendix*, Fig. S6*A*). When we introduced into this background two copies of the *UAS* construct containing the endogenous *Or88a* sequence (i.e., identical to that of the canonical transcript except that the AG splice acceptor site was not mutated), there was a strong response to methyl myristate (the "++" in Fig. 5*D* indicates that two copies were introduced). When we introduced two copies of the *UAS-exitron-spliced* construct, there was no response, as expected. When we introduced a copy of the *UAS-canonical* construct, a strong response was observed. We then used $Or88a^{GAL4}$ to drive copies of both a *UAS-canonical* and a *UAS-exitron-spliced* construct together to see whether the exitron-spliced transcript negatively affected the function of the canonical transcript or the neuron. We found that the *UAS-canonical* construct conferred strong responses despite the presence of the *UAS-exitron-spliced* construct, arguing against an interfering function for the exitron-spliced transcript (Fig. 5*D*).

We also considered the possibility that removal of the exitron affects the neuron passively, i.e., by effectively reducing the level of remaining RNA whose exitron has not been removed. As an initial test of this hypothesis, we asked whether the response of the neuron was sensitive to the dosage of the *Or88a* gene. We found that flies with two functional wild-type copies of *Or88a* (left bar in Fig. 5*E*, "++" indicates two copies) gave a larger response to methyl myristate than flies with one functional copy ($Or88a^{Gal4}$/+, center bar in Fig. 5*E*). Correspondingly, flies with two functional copies of *Or88a* have a greater *Or88a* expression level than those with a single copy (*SI Appendix*, Fig. S6*B*). We also tested flies that contain both of

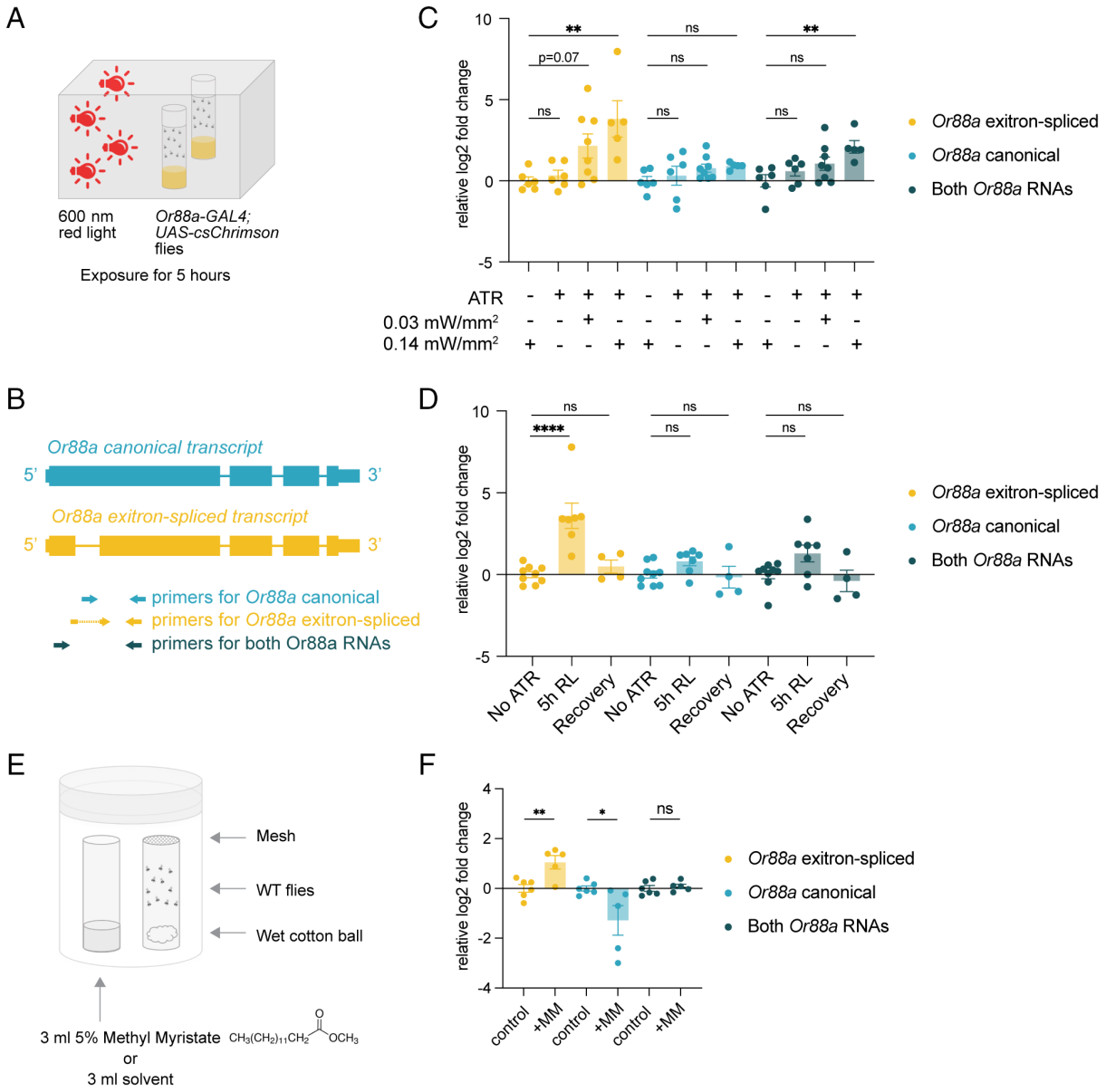

Fig. 4. Optogenetic and ligand activation of *Or88a*-expressing neurons increase *Or88a* exitron-spliced transcript levels. (*A*) Diagram of optogenetics experiments. One-day-old *Or88a-GAL4; UAS-CsChrimson* flies were collected and raised on fly food with ATR in the dark for 3 d. Then, fly vials were placed next to 600-nm red light LED panels for optogenetics exposure. (*B*) Primers used for RT-qPCR presented in panels (*C*), (*D*), and (*F*). (*C*) Effect of optogenetic activation of *Or88a*-expressing ORNs on *Or88a* transcript levels in *Or88a-GAL4; UAS-CsChrimson* flies. Two types of control flies were used in this experiment: flies that were not fed on the ATR supplement and that were exposed to red light, and flies that were fed ATR but were not exposed to red light. Red light exposure was at intensities of 0.03 mW/mm^2 or 0.14-mW/mm^2 for 5 h. The levels of *Or88a* RNAs are normalized to the control levels of flies that did not receive the ATR supplement. n = 5 to 8 biological replicates. (*D*) Effect of 5 h 0.14-mW/mm^2 red light followed by 24-h recovery on *Or88a* transcript levels. n = 4 to 9 biological replicates. (*E*) Diagram of ligand exposure experiments. Five-day-old flies are transferred to two vials (of which one is depicted) that contain wet cotton balls and are covered with mesh and then placed in a sealed container with a vial containing 3 mL solvent (DMSO) with or without methyl myristate for 5 h. (*F*) *Or88a* RNA levels after 5-h methyl myristate exposure. The levels of *Or88a* RNAs are normalized to the control levels when flies are only exposed to DMSO for 5 h. n = 5 to 6 biological replicates. For (*C*), (*D*), and (*F*), one-way ANOVA followed by Dunnett's multiple comparisons test was used. *P < 0.05; **P < 0.01; ****P < 0.0001; ns, not significant. Error bars are SEM.

these endogenous *Or88a* copies and in addition contained an *Or88a-GAL4* driver and a *UAS-Or88a-canonical* construct (right bar in Fig. 5*E*). Flies containing these transgenes showed greater responses than flies that did not.

Alternative Polyadenylation of *Ors*. In addition to finding unexpected alternative splicing of *Ors*, we also noticed through examination of RNA-Seq data (15) that some *Ors* have longer 3′ UTRs than indicated in their genomic annotations (BDGP version 6.48). We hypothesized that these apparent discrepancies could reflect alternative polyadenylation.

An interesting example of such a discrepancy is *Or67c* (Fig. 6*A*). According to the genomic annotation, there is a lncRNA *(CR44541)* downstream from *Or67c*, but our RNA-Seq data suggested that in the antenna, the annotated lncRNA sequences could rather be part of an alternative 3′ UTR of *Or67c*. To investigate the possibility of alternative polyadenylation, we used 3′ Rapid Amplification of cDNA Ends (3′RACE) and confirmed that *Or67c, Or59b, Or85f,* and *Or83c* transcripts have at least two 3′ UTR isoforms (Fig. 6 *B* and *C*).

Since there has been little if any previous documentation of alternative polyadenylation in *Drosophila* odor receptor genes, we investigated further using *Or67c* as an example. First, we asked whether its polyadenylation was conserved in evolution. Using antennal RNA from *D. simulans* and *D. sechellia*, we performed 3'RACE and found that all three species have three

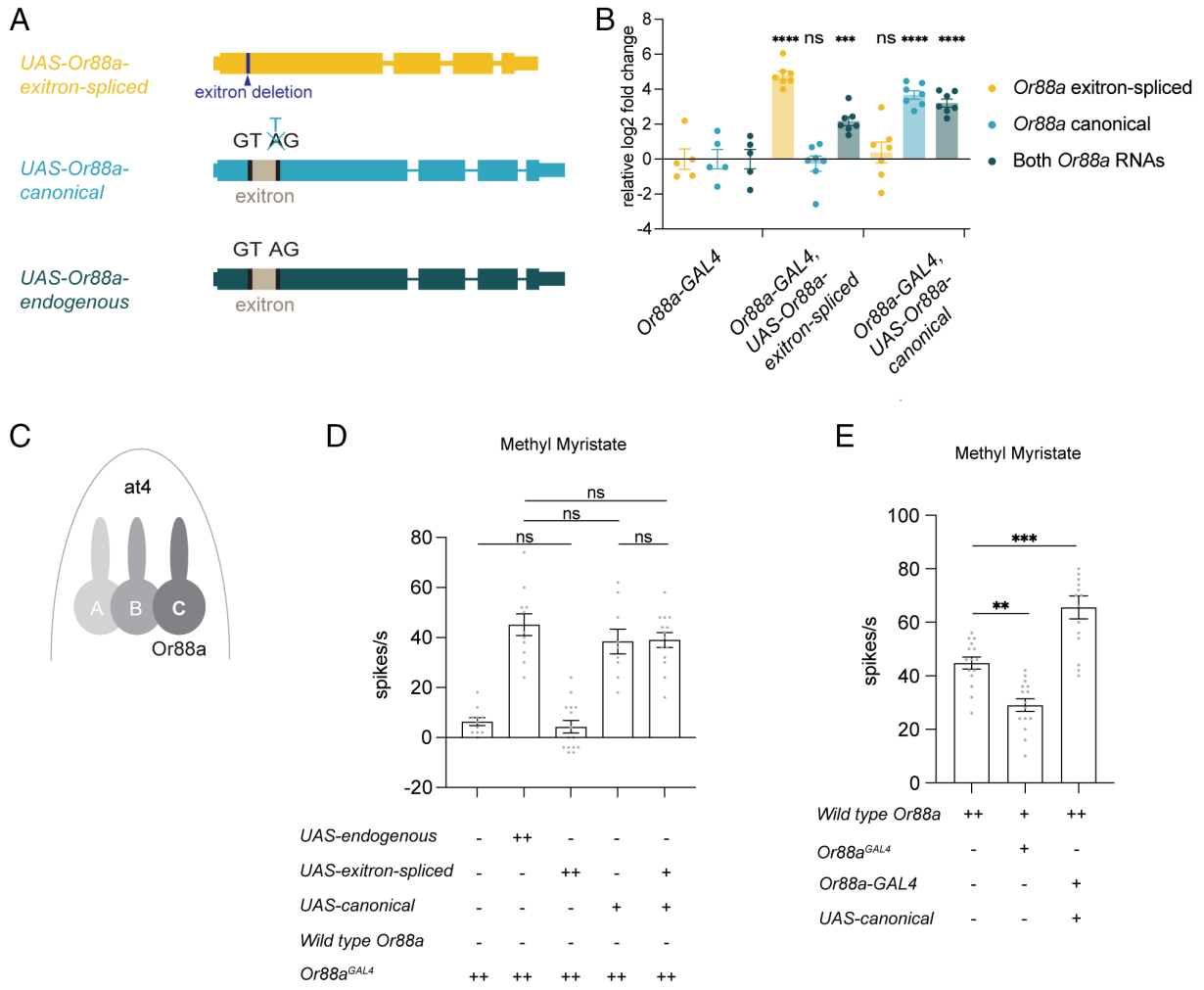

Fig. 5. The *Or88a* exitron-spliced transcript does not interfere with the function of the *Or88a* canonical transcript. (*A*) Diagram of the *UAS-Or88a* constructs used in this experiment. The *UAS-Or88a-exitron-spliced* and the *UAS-Or88a-canonical* are described in Fig. 3. The *UAS-Or88a-endogenous* is the same as the *UAS-Or88a-canonical* but preserves the AG exitron splice acceptor site. (*B*) The level of each indicated type of *Or88a* transcript (key at right) when overexpressing either exitron-spliced transcripts or canonical transcripts in *Or88a*-expressing ORNs. The transcript levels are normalized to the level of each transcript type in the control genotype (*Or88a-GAL4*). n = 5 to 7 biological replicates. (*C*) Diagram of at4 sensilla. Each at4 sensillum houses three ORNs, and Or88a is expressed in the at4C neuron. (*D*) Responses of at4 to methyl myristate in flies in which $Or88a^{GAL4}$ drives the indicated *UAS-Or88a* constructs. The number of "+" indicates the number of constructs or alleles of each type in each genotype. n = 9 to 15 sensilla from five to seven animals. (*E*) Responses of at4 to methyl myristate in flies that contain two (Wild type *Or88a*) versus one ($Or88a^{GAL4}$/+) endogenous copies of functional *Or88a*, and in flies that overexpress *Or88a (Or88a-GAL4; UAS-canonical)*. n = 14 to 15 from 5 to 10 animals. For (*B*), (*D*), and (*E*), one-way ANOVA followed by Dunnett's multiple comparisons test was used. $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$; ns, not significant. Error bars are SEM.

major 3′UTR isoforms, with lengths that were reasonably well conserved across species (Fig. 6*D*).

We next quantitated the relative levels of each *Or67c* isoform in *D. melanogaster* via RT-qPCR. We used three primer pairs, designed to amplify one, two, or three of the isoforms, from which we could deduce the relative expression levels of each (Fig. 6*E*). Approximately 20% of the transcripts were of the short form (isoform 1), 40% were of isoform 2, and 40% were of isoform 3 (Fig. 6*F*).

We examined a transposable element insertion mutant that carries a 7-kb insertion in the 3′UTR that is shared by isoforms 2 and 3 of *D. melanogaster Or67c* (Fig. 6*E*; see *Materials and Methods*). The two transcripts with longer 3′ UTRs are expected to be disrupted in this mutant. *Or67c* encodes a receptor that is expressed in the ab7B ORN and that responds to ethyl lactate. Electrophysiological recordings showed that the mutant had similar responses at two different ethyl lactate concentrations as the control strain, to which it had been backcrossed for at least five generations (Fig. 6*G* and *SI Appendix*, Fig. S6*C*). These results suggest that a single polyadenylation isoform is sufficient to confer response to the ORN under the stimulation conditions we used.

A previous study found that the level of *Or67c* mRNA was downregulated following exposure to ethyl lactate (26). This study did not discriminate among 3′UTR isoforms of *Or67c*. Using our primer sets, we determined that all three isoforms are downregulated by ethyl lactate exposure (Fig. 6*H*).

The presence of alternative 3′ UTRs offers several means of regulating the expression of a protein. Different 3′ UTRs may confer different mRNA stability, localization, or translation rates. The 3′ UTR can also influence post-translational modification and the formation of protein complexes (28). Our identification of alternative polyadenylation thus provides further opportunities for exploring the fine-tuning of ORN function.

Discussion

We have found that among 39 *Or* genes in the *D. melanogaster* antenna, four (10%) produce transcripts lacking sequences from within protein-coding exons. Multiple lines of evidence support the conclusion that these transcripts arise via removal of *Or* exitron sequences by the splicing machinery, as opposed to via an artifactual mechanism:

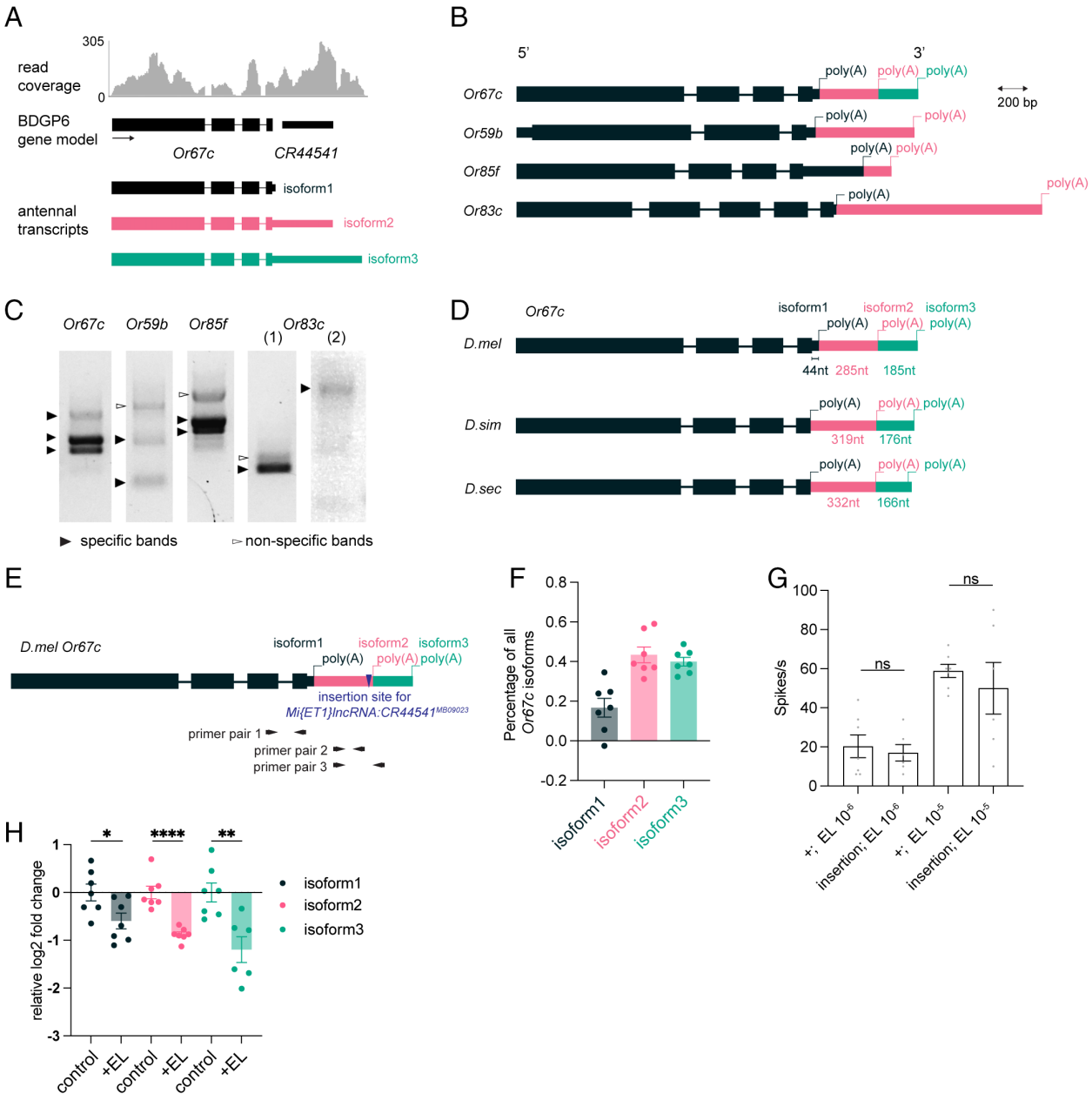

Fig. 6. Alternative polyadenylation of *Drosophila Or* genes. (*A*) The read coverage of *Or67c* and *CR44541* in our antennal RNA-Seq analysis. Below is the BDGP6 (version 6.48) gene model of *Or67c* and our updated model of *Or67c* with different 3′UTR isoforms. (*B*) Four *Ors* were found to undergo alternative polyadenylation based on our RNA-Seq analysis. Different forms of *Or* transcripts are shown in black, pink, and green. (*C*) 3′RACE results of the four *Ors*. Black triangles: specific DNA bands that can be aligned to regions of the respective *Or* genes. White triangles: non-specific bands that when sequenced align to other, apparently unrelated genes due to non-specific binding of primers. According to the RNA-Seq data, one *Or83c* 3′UTR isoform is ~1 kb longer than the other one and because of its greater length is shown on a distinct gel. Specifically, 3′RACE using a primer near the 3′ end of the coding region only shows the short transcript due to limitations of amplifying long fragments with this technique (Track 1). Therefore, for 3′ RACE, we used a primer ~1 kb downstream of the *Or83c* coding region near the 3′ end of the long UTR (Track 2). We also ligated the fragments to plasmid backbones and sequenced them. The sequencing results were in good agreement with the results of Fig. 6*B*, except that we did not obtain enough sequence data from *Or85f* to confirm conclusively. (*D*) Summary of 3′RACE results of *Or67c* in three different species. (*E*) Diagram of *Or67c* primers used in this study. Primer pair 1 is used to detect the total amount of *Or67c* isoforms. Primer pair 2 is used to detect the amount of *Or67c* isoform2+isoform3. Primer pair 3 is used to detect the amount of isoform3. The site of the insertion mutation is indicated (*Materials and Methods*). (*F*) Percentage of *Or67c* isoforms. Primers are shown in Fig. 6*E*. n = 7 biological replicates. (*G*) Responses recorded from ab7 sensilla of control flies (+), which are a *wCS* control for genetic background, and an insertion mutant (insertion site shown in Fig. 6*E*) using single-sensillum recording. Two concentrations of ethyl lactate, diluted in paraffin oil, (EL 10^{-5}, EL 10^{-6}) were used for recording. n = 6 to 7 sensilla from 5 animals. (*H*) Effect of ethyl lactate exposure on *Or67c* transcript levels. Control (*wCS*) flies were exposed to ethyl lactate, a strong ligand for Or67c, for 5 h using the paradigm shown in Fig. 4*E*. In the central vial, 3 mL of 5% ethyl lactate was added. n = 6 to 7 biological replicates. For (*F–H*), one-way ANOVA followed by Dunnett's multiple comparisons test was used. *$P < 0.05$; **$P < 0.01$; ****$P < 0.0001$; ns, not significant. Error bars are SEM.

i) All four of the removed *Or* sequences have precise boundaries (79/79 cases);

ii) Splice site motifs, including the canonical splice donor and acceptor sites GT and AG, are present at the boundaries of the removed sequences in all four *Or* genes;

iii) In the case of *Or88a*, the sequence is removed at the canonical splicing site GT/AG in all species tested including *D. erecta*, in which the 3′ end of the sequence is different from those of the other species;

iv) The sequences lack a hairpin structure and direct repeats that overlap or abut their stems, which could have promoted RT slippage;

v) Following mutation of the splice acceptor site AG to TG in *Or88a,* few if any of the spliced transcripts were detected

(*SI Appendix*, Fig. S4, *UAS-Or88a-canonical; Or88a*^{Gal4}; Fig. 1*C* shows abundant product in wild type);

vi) Overexpression of the *Or88a* canonical transcript, containing the AG-to-TG mutation, did not lead to an increased abundance of the exitron-spliced transcript (Fig. 5*B*). (Note that the mutation does not change the predicted secondary structure of the exitron, as shown in *SI Appendix*, Fig. S3);

vii) Optogenetic activation leads to an increase in the level of exitron-spliced transcript but not the canonical transcript (Fig. 4*C*), and exposure of flies to a pheromone resulted in different changes in the levels of exitron-spliced transcript and canonical transcript (Fig. 4*F*). It is simpler to explain these differences in levels of the two splice forms in terms of alterations in levels of spliced transcripts in vivo than in terms of alterations in reactions in vitro after antennal RNA extraction.

To our knowledge, there are no other reports of exitrons in *Drosophila*. Why have exitrons not previously been reported in *Ors* or other *Drosophila* genes? One reason is that these exitron-spliced transcripts are present at low levels: The exitron-spliced *Or88a* transcript, for example, constituted ~15% of *Or88a* transcripts in each of two independent experiments. Exitron-spliced transcripts may have been previously misinterpreted as artifacts of RNA-Seq library construction or as transcription errors (17). Our identification of them has relied upon high-quality antennal-specific RNA-Seq data containing many reads (15) and careful analysis of the data. The low level of exitron-spliced *Or* transcripts makes it difficult to identify them by direct long-read RNA sequencing or northern blot analysis, but in the future, improved sequencing technology may facilitate the discovery of additional exitrons in sensory organs, or other organs, of *Drosophila*.

We found that the exitron of *Or88a* is present in five *Drosophila* species. The most recent common ancestor of these species lived 20 Mya (20), suggesting that these exitrons have been conserved and contribute to the fitness of the species—presumably by enhancing the ability of the flies to interpret their olfactory environment.

In principle, the exitron-spliced transcripts could be translated. The exitron-spliced transcript of *Or23a* is predicted to encode an internally deleted protein that lacks roughly half of the amino acids of its canonical counterpart. In *Or88a, Or82a,* and *Or92a,* the removal of the exitron creates a frameshift, and in each case, two polypeptides could in principle arise. One polypeptide would be short and would include N-terminal sequences of the receptor; the other would be longer, containing the last five membrane domains and the C terminus of each receptor.

One can imagine a variety of intriguing functions for such partial receptor proteins. They could bind to full-length Ors and inhibit their assembly into functional receptors, which are multimeric; they could bind and inhibit other proteins that interact with Ors; they could even bind to small molecules—perhaps even odorants—and influence Or signaling. However, extensive analysis of a variety of epitope-tagged *Or88a* exitron-spliced transcripts, including transcripts tagged at both 5′ and 3′ ends, did not identify translation products. The simplest interpretation of these results is that *Or88a* exitron-spliced transcripts are non-coding. We cannot exclude the possibility that *Or88a* exitron-spliced transcripts are translated but into unstable peptides. Moreover, we have not carried out analogous experiments with alternatively spliced transcripts of the other four *Ors*, leaving open the possibility of their translation. However, our findings with *Or88a* prompted us to turn our attention from potential functions of shortened receptor proteins to potential functions of shortened receptor transcripts.

To explore the roles of the non-coding, exitron-spliced transcripts of *Or88a*, first, we asked whether the level of the *Or88a* exitron-spliced transcript was affected by neuronal activity. We found that optogenetic activation of the neurons that express *Or88a* caused a 14-fold increase in the level of the exitron-spliced transcript, but no increase in the level of the canonical transcript (Fig. 4*C*). The fraction of *Or88a* transcripts that were exitron-spliced increased from 15% to 54%. Consistent with these results, activation of the neurons by another means, exposure to the pheromone that activates them, also increased the level of *Or88a* exitron-spliced transcripts (Fig. 4*F*). Together, these results provide direct evidence that the activity of *Drosophila* ORNs can regulate the levels of specific transcript isoforms.

We considered ways in which exitron-spliced transcripts, and an increase in their level, might affect the neurons in which they are expressed. Cytoplasmic long non-coding RNAs (lncRNAs) can act as competitors for RNA binding proteins and miRNA and thereby influence the stability and translation of mRNAs. Alternatively, lncRNAs can modulate protein function; for example, a lncRNA can regulate post-translational modification or localization of a given protein (29). Engineered overexpression of the exitron-spliced transcript did not affect the level of the endogenous canonical transcript (Fig. 5*B*). Nor did expression of the exitron-spliced transcript interfere with the functional expression of the canonical transcript (Fig. 5*D*).

We next considered the possibility that exitron-splicing could affect the neuron simply by virtue of reducing the level of canonical transcript, i.e., when primary transcripts undergo the removal of an exitron, there are fewer transcripts to encode full-length receptor proteins. Consistent with this notion, we found that the response of the neuron to odorant is dependent on the dosage of *Or88a* genes (Fig. 5*E*), suggesting that the level of functional *Or88a* mRNA may also influence the magnitude of the neuron's response. Our identification of exitron splicing thus invites further analysis of its effects on neuronal function, and in particular, whether it may be a mechanism of neuronal modulation. For example, it could provide a means of olfactory habituation.

We note the possibility that the *Or88a* exitron-spliced transcript could also have an independent function. In mammals, *Or* mRNAs have been found to have non-coding functions: They can inhibit the transcription of other *Ors,* acting as a critical component in the intricate regulatory mechanism of *Or* gene choice (30). In addition, there is ample precedent for functional non-coding transcripts derived from other protein-coding genes (31, 32).

There is prior evidence that activation of ORNs can change the level of *Or* transcription, either positively or negatively, in both flies and mice (26, 27). Our results reveal another dimension of activity-induced gene regulation: exitron splicing. Although it is difficult to estimate the odor concentrations that animals experience in a natural environment, in principle regulatory mechanisms such as exitron-splicing may be beneficial in allowing insects to adapt to high odor or pheromone concentrations they may encounter in certain circumstances in nature.

A previous analysis of ~400 mouse olfactory receptor genes identified 21 in which a splice site within the protein-coding region was used in some but not all transcripts (33). Our finding that neuronal activity can modulate exitron-spliced transcript levels raises the possibility that activity-induced exitron-splicing may be conserved in mammals.

Several other intriguing questions are raised by our results. Among the 39 *Ors* expressed in the antenna, do the five *Ors* that are considered here share any common features? These five *Ors* are diverse in many ways. They are diverse in terms of sequence and chromosomal location (3). They are expressed in multiple morphological classes of

sensilla: *Or88a* is expressed in trichoid sensilla, *Or35a* in coeloconic sensilla, *Or82a* in basiconic sensilla. Within sensilla, *Or82a* is expressed in A neurons with large amplitudes (34); *Or35a* is expressed in B neurons with small amplitudes (35). Or88a is tuned to pheromones (16); Or82a is tuned narrowly to the odorant geranyl acetate (36); Or35a is broadly tuned to many odorants (36). A unifying feature of these five receptors is not immediately apparent to us. In this regard, we note the possibility that additional *Ors* may also contain exitrons that could be identified following activation of other ORNs with optogenetics or olfactory stimuli.

Another question raised by our results concerns the mechanism of exitron splicing. Are there particular splicing factors that recognize and bind to the weak splicing sites of these exitrons? Are these splicing factors induced by activation of the neurons? Alternatively, does the exitron-spliced transcript increase in stability upon neuronal activation? Such an increase in stability could be achieved via different mechanisms. Prior evidence suggests that the activation of neurons leads to changes in the levels of miRNAs and RNA binding proteins, both of which influence RNA stability (37, 38). Certain exitron-spliced transcripts containing premature termination codons are targets of nonsense-mediated decay (NMD) (12, 13). It is plausible that *Or88a* exitron-spliced transcripts are targets of NMD and that neuronal activation could increase their resistance to NMD.

In summary, we have identified an additional degree of freedom in the expression of *Or* genes: exitron splicing. We have shown that exitron splicing has been conserved for 20 My of evolution, and that the level of exitron-spliced transcripts is increased by neuronal activation. Our results add a dimension to the expression of odor receptors, which underlies the entirety of olfactory function.

Materials and Methods

***Drosophila* Stocks.** Flies were raised on corn syrup and soy flour culture medium (Archon Scientific) at 25 °C and 60% relative humidity in a 12:12-h light–dark cycle. *D. simulans* (14021-0251.001), *D. sechellia* (14021-0248.27), and *D. erecta* (14021-0224.01) were obtained from the *Drosophila* Species Stock Center. The *D. suzukii* stock was collected in Connecticut. The *Or88a-GAL4* line was obtained from Bloomington *Drosophila* Stock Center (#23137). The *Or67c* insertion line was from the Bloomington *Drosophila* Stock Center (#61738); it is designated w^{1118}*; Mi{ET1}lncRNA:CR44541*^{MB09023}, but we have determined that the annotated lncRNA is in fact part of the 3′UTR of *Or67c* in the antenna, as shown in Fig. 6*E*. The *UAS-CsChrimson* line was a gift from Vivek Jayaraman. *Or88a knock-in GAL4 (Or88a*^{GAL4}*)* was generated using the CRISPR method described in ref. 39. To generate *UAS-Or88a* lines, the *Or88a* gene was amplified from Canton-S w^{1118} genomic DNA and cloned into *pUAST-attB-QS* plasmids. Then, the NEB Q5® Site-Directed Mutagenesis Kit was used to delete the exitron, or mutate the exitron splice site, or insert HA tags. Plasmids were injected into embryos by Bestgene. All lines were backcrossed to Canton-S w^{1118} *(wCS)* for at least five generations before use.

RNA Purification. To extract RNA from *Drosophila* antennae, third antennal segments with the arista attached were hand-dissected from 5-d-old flies and immediately placed into 1.5-mL microfuge tubes containing liquid nitrogen. Each sample consisted of about 130 hand-dissected antennae. The tissues were then mechanically crushed with disposable plastic pestles in liquid nitrogen and lysed in RTL lysis buffer (Qiagen) on ice. Acid phenol was then added to the lysate at a 1:1 ratio, and the samples were then heated five times at 65 °C for 2 min, with vortexing in between heating periods. After cooling on ice for 5 min, the samples were spun at 15 k rpm for 10 min to separate the aqueous phase containing RNA from the phenol. The aqueous phase was then added to chloroform to remove any residual phenol. To remove DNA contamination, DNase-Zero (LGC, Biosearch Technologies) was used at 37 °C for 30 min.

RNA-Sequencing and Analysis. Antennal polyadenylated RNA was enriched using NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs). The library was prepared using the NEBNext Ultra II Directional RNA Library Prep Kit (New England Biolabs). Reads were aligned to the *D. melanogaster genome* (BDGP6) using TopHat (version 2.1.2). Integrative genomics viewer (version 2.8.9) was used to manually inspect reads mapped to *Odorant receptor* genes. Raw reads are accessible at the Genbank SRA database (BioProject accession number PRJNA1081896).

RT-PCR and qPCR. cDNA was synthesized from 0.5 μg of total RNA as templates using EpiScript according to the manufacturer's protocol (Lucigen), followed by PCR amplification with 30 or 35 amplification cycles. To control for genomic DNA contamination, each RNA sample underwent a parallel mock reverse transcription step (no RT control) in which the reverse transcriptase was omitted, before being added to PCR reactions. The RT-PCR products were then electrophoresed on agarose gels. PCR primers are listed in *SI Appendix*, Table S1. In some experiments, the RT-PCR products were purified with a Qiagen PCR purification kit and sequenced by Sanger sequencing.

qPCR was carried out using the iTaq Universal SYBR Green (Bio-Rad) system with about 20 ng of cDNA. qPCR primers are listed in the *SI Appendix*, Table S1. Target gene expression was normalized to the level of *Orco* transcripts. ΔΔCq values were plotted for relative log2 fold changes (Figs. 4–6). To calculate the percentage of *Or* transcripts, we first computed the absolute abundance of each transcript based on the standard curve analysis of the corresponding in vitro transcribed RNA.

The percentage of *Or88a* exitron-spliced transcript was calculated as (the amount of transcript amplified by the primers that amplify both transcripts − the amount of transcript amplified by the canonical transcript primers)/the amount of transcript amplified by the primers that amplify both transcripts (Fig. 1).

The percentage of *Or67c* isoforms was calculated as follows (Fig. 6):

Isoform 1: (the amount of transcript amplified by primer pair 1 − the amount of transcript amplified by primer pair 2)/the amount of transcript amplified by primer pair 1.

Isoform 2: (the amount of transcript amplified by primer pair 2 − the amount of transcript amplified by primer pair 3)/the amount of transcript amplified by primer pair 1.

Isoform 3: the amount of transcript amplified by primer pair 3/the amount of transcript amplified by primer pair 1.

Extracellular Single-Sensillum Recording (SSR). Flies that were 3 to 5 d old were used for single-sensillum recording (SSR). Chemicals with a purity ≥98% were purchased from Sigma-Aldrich and used for odorant delivery. To prepare odor cartridges, 50 μL of chemical was applied to a 13-mm diameter filter paper disc inside a Pasteur pipette. The cartridge was capped with a 1,000-μL pipette tip and allowed to equilibrate for at least 20 min prior to use. Each odor cartridge was used for no more than three presentations. The odorant stimuli were presented by placing the tip of the cartridge into a glass tube, which delivered a stream of humidified air (~2,000 mL/min) to the fly antenna, and administering a 500-ms pulse of air (~200 mL/min) through the cartridge. Neuronal firing rates were recorded and measured using AUTOSPIKE software and analyzed in Prism. To calculate the firing rate for a specific odorant, we first measured the neuronal firing rate following 500-ms odor delivery and then subtracted the spontaneous firing rate over a 500-ms interval 1.5 s before odor delivery. Likewise, we measured the firing rate to the paraffin oil diluent alone, from which we also subtracted the spontaneous firing rate. We then subtracted the response to paraffin oil from the response to the odorant, and the differences are what are shown in the graphs.

In Situ Hybridization and Immunohistochemistry. Using primers listed in *SI Appendix*, Table S1, *Or88a* DNA fragments were amplified from fly genomic DNA and then cloned into the pGEM-T Easy Vector System (Promega) for digoxigenin (DIG)–labeled RNA antisense probe synthesis using standard methods. The protocol for fluorescence in situ hybridization and immunohistochemistry were similar to those described in ref. 40. Anti-HA antibody was used as the primary antibody and Anti-Rat IgG AF488 was used as the secondary antibody for immunofluorescence. Microscopy was performed using a Zeiss LSM 880 Laser Scanning Confocal Microscope with the 40× lens and images were processed with ImageJ software.

Optogenetics. Flies were collected on the first day of eclosion and were given fly food containing 0.5 mM ATR. Flies were kept in the dark for 3 d and used for optogenetic experiments on the fourth day. For optogenetic experiments (41), flies contained in transparent plastic fly vials were placed in an area illuminated by a custom-built LED matrix with a wavelength of 625 nm (SMD 5050,

LEDlightninghut.com) programmed to deliver one-second on, one-second off pulses for 5 h at either an intensity of 0.03 mW/mm^2 or 0.14-mW/mm^2. Of note, 0.03 mW/mm^2 is the lowest intensity that was observed to have an effect in the optogenetic study of Moreira et al. (42); we note also that 0.06 mW/mm^2 was used in the original study of Klapoetke et al. (25). Following exposure, flies were immediately placed on ice for antenna dissection.

Ligand Exposure Experiments. To set up the experiment, 5-d-old flies were placed into empty vials that had a wet cotton ball at the bottom to provide moisture. A piece of mesh was then placed on top of each vial to prevent the flies from escaping. The fly vials were subsequently moved to a 10-cm-diameter × 15-cm-height container where a vial containing 3 mL of diluted chemicals (5% chemical in DMSO) was also included. The container was sealed and maintained at 25 °C and 60% relative humidity for 5 h. After 5 h, flies are taken from the container and anesthetized immediately on ice for dissection.

Data, Materials, and Software Availability. The RNA sequence data have been deposited in the Genbank SRA database and are accessible through BioProject accession number PRJNA1081896 (43). Previously published data were used for this work (15). All other study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank Zina Berman for support. This work was supported by NIH F32 DC019302 to G.J.S.T. and NIH grants R01 DC02174, R01 DC04729, and R01 DC011697 to J.R.C.

1. A. Couto, M. Alenius, B. J. Dickson, Molecular, anatomical, and functional organization of the Drosophila olfactory system. *Curr. Biol.* **15**, 1535–1547 (2005).
2. E. A. Hallem, J. R. Carlson, The odor coding system of Drosophila. *Trends Genet.* **20**, 453–459 (2004).
3. H. M. Robertson, C. G. Warr, J. R. Carlson, Molecular evolution of the insect chemoreceptor gene superfamily in Drosophila melanogaster. *Proc. Natl. Acad. Sci. U.S.A.* 100 (Suppl 2), 14537–14542 (2003).
4. W. van der Goes, J. R. van Naters, Carlson, Receptors and neurons for fly odors in Drosophila. *Curr. Biol.* **17**, 606–612 (2007).
5. L. Bai, A. L. Goldman, J. R. Carlson, Positive and negative regulation of odor receptor gene choice in Drosophila by Acj6. *J. Neurosci.* **29**, 12940–12947 (2009).
6. S. H. Fuss, A. Ray, Mechanisms of odorant receptor gene choice in Drosophila and vertebrates. *Mol. Cell Neurosci.* **41**, 101–112 (2009).
7. A. Ray, W. van der Goes, T. van Naters, J. R. Shiraiwa, Carlson, Mechanisms of odor receptor gene choice in Drosophila. *Neuron* **53**, 353–369 (2007).
8. E. Song, B. de Bivort, C. Dan, S. Kunes, Determinants of the Drosophila odorant receptor pattern. *Dev. Cell* **22**, 363–376 (2012).
9. S. Barish *et al.*, Combinations of DIPs and Dprs control organization of olfactory receptor neuron terminals in Drosophila. *PLOS Genet.* **14**, e1007560 (2018).
10. C. Andreassi, H. Crerar, A. Riccio, Post-transcriptional processing of mRNA in neurons: The vestiges of the RNA world drive transcriptome diversity. *Front. Mol. Neurosci.* **11**, 304 (2018).
11. R. E. Halbeisen, A. Galgano, T. Scherrer, A. P. Gerber, Post-transcriptional gene regulation: From genome-wide studies to principles. *Cell. Mol. Life Sci.* **65**, 798 (2007).
12. Y. Marquez, M. Höpfler, Z. Ayatollahi, A. Barta, M. Kalyna, Unmasking alternative splicing inside protein-coding exons defines exitrons and their role in proteome plasticity. *Genome Res.* **25**, 995–1007 (2015).
13. T.-Y. Wang *et al.*, A pan-cancer transcriptome analysis of exitron splicing identifies novel cancer driver genes and neoepitopes. *Mol. Cell* **81**, 2246–2260.e12 (2021).
14. Y. Zhang *et al.*, Landscape of exitrons in gastric cancer. *eBioMedicine* **84**, 104272 (2022).
15. G. J. S. Talross, J. R. Carlson, The rich non-coding RNA landscape of the Drosophila antenna. *Cell Rep.* **42**, 112482 (2023).
16. H. K. M. Dweck *et al.*, Pheromones mediating copulation and attraction in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E2829–E2835 (2015).
17. L. Schulz *et al.*, Direct long-read RNA sequencing identifies a subset of questionable exitrons likely arising from reverse transcription artifacts. *Genome Biol.* **22**, 190 (2021).
18. Y. J. Zhang, H. Y. Pan, S. J. Gao, Reverse transcription slippage over the mRNA secondary structure of the LIP1 gene. *Biotechniques* **31**, 1286, 1288, 1290, passim (2001).
19. M. Abou Alezz, L. Celli, G. Belotti, A. Lisa, S. Bione, GC-AG introns features in long non-coding and protein-coding genes suggest their role in gene expression regulation. *Front Genet.* **11**, 488 (2020).
20. F. Li *et al.*, Phylogenomic analyses of the genus Drosophila reveals genomic signals of climate adaptation. *Mol. Ecol. Res.* **22**, 1559–1581 (2022).
21. H. Guo, K. Kunwar, D. Smith, Odorant receptor sensitivity modulation in Drosophila. *J. Neurosci.* **37**, 9465–9473 (2017).
22. J. S. Kesner *et al.*, Noncoding translation mitigation. *Nature* **617**, 395–402 (2023).
23. G. Hermey, N. Blüthgen, D. Kuhl, Neuronal activity-regulated alternative mRNA splicing. *Int. J. Biochem. Cell Biol.* **91**, 184–193 (2017).
24. A. Thalhammer, F. Jaudon, L. A. Cingolani, Emerging roles of activity-dependent alternative splicing in homeostatic plasticity. *Front Cell Neurosci.* **14**, 104 (2020).
25. N. C. Klapoetke *et al.*, Independent optical excitation of distinct neural populations. *Nat. Methods* **11**, 338–346 (2014).
26. S. Koerte *et al.*, Evaluation of the DREAM technique for a high-throughput deorphanization of chemosensory receptors in Drosophila. *Front. Mol. Neurosci.* **11**, 366 (2018).
27. B. von der Weid *et al.*, Large-scale transcriptional profiling of chemosensory neurons identifies receptor-ligand pairs *in vivo*. *Nat. Neurosci.* **18**, 1455–1463 (2015).
28. S. Mitschka, C. Mayr, Context-specific regulation and function of mRNA alternative polyadenylation. *Nat. Rev. Mol. Cell Biol.* **23**, 779–796 (2022).
29. L. Statello, C.-J. Guo, L.-L. Chen, M. Huarte, Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **22**, 96–118 (2021).
30. A. D. Pourmorady *et al.*, RNA-mediated symmetry breaking enables singular olfactory receptor choice. *Nature* **625**, 181–188 (2024).
31. S. Dhamija, M. B. Menon, Non-coding transcript variants of protein-coding genes – what are they good for? *RNA Biol.* **15**, 1025–1031 (2018).
32. K. Sampath, A. Ephrussi, CncRNAs: RNAs with both coding and non-coding roles in development. *Development* **143**, 1234–1241 (2016).
33. J. M. Young *et al.*, Odorant receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels. *Genome Biol.* **4**, R71 (2003).
34. E. A. Hallem, M. G. Ho, J. R. Carlson, The molecular basis of odor coding in the Drosophila antenna. *Cell* **117**, 965–979 (2004).
35. C. A. Yao, R. Ignell, J. R. Carlson, Chemosensory coding by neurons in the coeloconic sensilla of the Drosophila antenna. *J. Neurosci.* **25**, 8359–8367 (2005).
36. E. A. Hallem, J. R. Carlson, Coding of odors by a receptor repertoire. *Cell* **125**, 143–160 (2006).
37. N. K. Vo, X. A. Cambronne, R. H. Goodman, MicroRNA pathways in neural development and plasticity. *Curr. Opin. Neurobiol.* **20**, 457–465 (2010).
38. J. Ule, R. B. Darnell, RNA binding proteins and the regulation of neuronal synaptic plasticity. *Curr. Opin. Neurobiol.* **16**, 102–110 (2006).
39. J. S. Chahda *et al.*, The molecular and cellular basis of olfactory response to tsetse fly attractants. *PLoS Genet.* **15**, e1008005 (2019).
40. N. K. Larter, J. S. Sun, J. R. Carlson, Organization and function of Drosophila odorant binding proteins. *eLife* **5**, e20242 (2016).
41. Z. He, Y. Luo, X. Shang, J. S. Sun, J. R. Carlson, Chemosensory sensilla of the Drosophila wing express a candidate ionotropic pheromone receptor. *PLoS Biol.* **17**, e2006619 (2019).
42. J.-M. Moreira *et al.*, optoPAD, a closed-loop optogenetics system to study the circuit basis of feeding behaviors. *eLife* **8**, e43924 (2019).
43. X. Shang, G. J. S. Talross, J. R. Carlson, Exitron splicing of odor receptor genes in *Drosophila*. NCBI BioProject. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1081896. Deposited 4 March 2024.