bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Coordination of two enhancers drives expression of olfactory trace amine-associated receptors

Aimei Fei[1,8], Wanqing Wu[1,8], Longzhi Tan[3,8], Cheng Tang[4,8], Zhengrong Xu[1], Xiaona Huo[4], Hongqiang Bao[1], Mark Johnson[5], Griffin Hartmann[5], Mustafa Talay[5], Cheng Yang[1], Clemens Riegler[6], Kristian Joseph[6], Florian Engert[6], X. Sunney Xie[3], Gilad Barnea[5], Stephen D. Liberles[7], Hui Yang[4], and Qian Li[1,2,*]

[1]Center for Brain Science, Shanghai Children's Medical Center, Department of Anatomy and Physiology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China;

[2]Shanghai Research Center for Brain Science and Brain-Inspired Intelligence, Shanghai 201210, China;

[3]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA;

[4]Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Research Center for Brain Science and Brian-Inspired Intelligence, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China;

[5]Department of Neuroscience, Division of Biology and Medicine, Brown University, Providence, RI 02912, USA;

[6]Department of Molecular and Cellular Biology and Center for Brain Science, Harvard University, Cambridge, MA 02138, USA;

[7]Howard Hughes Medical Institute, Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA;

[8]These authors contributed equally to this work.

*Correspondence to liqian@shsmu.edu.cn, phone: +86-21-63846590 ext. 776985

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Summary

Olfactory sensory neurons (OSNs) are functionally defined by their expression of a unique odorant receptor (OR). Mechanisms underlying singular OR expression are well studied, and involve a massive cross-chromosomal enhancer interaction network. Trace amine-associated receptors (TAARs) form a distinct family of olfactory receptors, and here we find that mechanisms regulating *Taar* gene choice display many unique features. The epigenetic signature of *Taar* genes in TAAR OSNs is different from that in OR OSNs. We further identify that two TAAR enhancers conserved across placental mammals are absolutely required for expression of the entire *Taar* gene repertoire. Deletion of either enhancer dramatically decreases the expression probabilities of different *Taar* genes, while deletion of both enhancers completely eliminates the TAAR OSN populations. In addition, both of the enhancers are sufficient to drive transgene expression in the partially overlapped TAAR OSNs. We also show that the TAAR enhancers operate in *cis* to regulate *Taar* gene expression. Our findings reveal a coordinated control of *Taar* gene choice in OSNs by two remote enhancers, and provide an excellent model to study molecular mechanisms underlying formation of an olfactory subsystem.

Keywords: olfactory receptor, trace amine-associated receptor, main olfactory epithelium, olfactory subsystem, enhancer

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

In the mammalian olfactory systems, the ability to detect and discriminate a multitude of odorants relies on expression of a wide range of receptor genes in olfactory sensory neurons (OSNs) [1]. In the main olfactory epithelium (MOE), olfactory receptor genes are mainly composed of two families of seven-transmembrane G protein-coupled receptors (GPCRs): odorant receptors (ORs) [1] and trace amine-associated receptors (TAARs) [2]. Some OSNs express other types of olfactory receptors that are not GPCRs: membrane-spanning 4-pass A receptors (MS4A) [3] and guanylyl cyclase D (GC-D) [4, 5]. OSNs expressing different olfactory receptor gene families constitute distinct olfactory subsystems that detect specific categories of odorants. In addition to odorant recognition, ORs also play an instructive role in targeting the axons of OSNs into specific locations in the olfactory bulb [6-8]. Thus, correct expression of olfactory receptor genes is critical for precise translation of external odor information into the brain.

In mice, the *OR* gene family consists of ~1,100 functional genes and form the largest gene family in the genome. *OR* transcription is initiated by epigenetic switch from the repressive H3K9me3/H4K20me3 (histone H3 trimethyl lysine 9 and histone H4 trimethyl lysine 20) state to the active H3K4me3 (histone H3 trimethyl lysine 4) state on a stochastically chosen *OR* allele [9]. This process involves an enzymatic complex with histone demethylases including LSD1 (lysine-specific demethylase 1) [10]. Once a functional OR protein is expressed, a feedback signal is triggered to prevent the de-silencing of other *OR* genes by inhibiting the histone demethylase complex, thereby stabilizing the *OR* gene choice throughout the lifetime of each OSN [11-14]. In addition, *OR* gene choice is facilitated by multiple intergenic OR enhancers (63 in total, also known as “Greek Islands”) that interact with each other to form an interchromosomal enhancer hub [15-18]. The OR enhancer hub is hypothesized to insulate a chosen active *OR* allele from the surrounding repressive heterochromatin compartment. Thus, deletion of a single OR enhancer typically affects 7-10 *OR* genes within the same *OR*

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

clusters adjacent to the enhancer (within 200 kb). For instance, knockout of the H, P, or Lipsi enhancers decreases the expression of 7, 10, or 8 *OR* genes, respectively [15, 19, 20]. One exception is the J element (or called element A [21]) that exerts its function on 75 *OR* genes over ~ 3 megabase genomic distance [22]. Nevertheless, while the OR enhancers interact in *trans* to form the multi-enhancer hub, their restricted effects on proximal OR gene expression suggest that they operate in *cis*. This may be due to redundancy among the 63 enhancers in hub formation. As a result, deletion of a single or a few OR enhancers did not result in overall changes in *OR* gene expression [17].

On the other side, TAARs form a distinct olfactory receptor subfamily and are evolutionarily conserved in jawed vertebrates, including humans [23]. TAARs are distantly related to biogenic amine receptors, such as dopamine and serotonin receptors, and not related to ORs. The TAAR family is much smaller than the OR family, with 15 functional members in mouse, 17 in rat, and 6 in human [24]. In mouse, *Taar* genes are arranged in a single cluster in chromosome 10 and are numbered based on their chromosomal order, from *Taar1* to *Taar9*, with five intact *Taar7* genes (*Taar7a*, *Taar7b*, *Taar7d*, *Taar7e* and *Taar7f*; *Taar7c* is the only *Taar* pseudogene) and three intact *Taar8* genes (*Taar8a*, *Taar8b* and *Taar8c*). *Taar1* is mainly expressed in the brain, while the other *Taar* genes are mainly expressed in the dorsal zone of MOE except that *Taar6* and two members of *Taar7* (*Taar7a* and *Taar7b*) are expressed more ventrally [2, 25, 26]. Several TAARs respond to volatile amines some of which are ethologically relevant odors that serve as predator signal or social cues and evoke innate behaviors [27-35]. *Taar* genes do not co-express with *OR* genes, suggesting that TAARs constitute a distinct olfactory subsystem [2]. Consistent with this notion, genetic evidences suggested distinct mechanisms of receptor choice for TAARs and ORs [25, 26]. However, the nature of the regulatory mechanisms of *Taar* gene choice remains elusive.

Here, we found that the *Taar* gene cluster is decorated by different heterochromatic marks in TAAR and OR OSNs. We further searched for *Taar* specific enhancers by

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

performing ATAC-seq (assay for transposase-accessible chromatin using sequencing) on purified TAAR OSNs. We identified two TAAR enhancers in the 200 kb *Taar* gene cluster, with TAAR enhancer 1 located between *Taar1* and *Taar2* and TAAR enhancer 2 between *Taar6* and *Taar7a*. Both enhancers are evolutionarily conserved in placental mammals, including humans. Deletion of either one leads to specific abolishment or dramatic decrease of distinct TAAR OSN populations, suggesting that the two enhancers act in a non-redundant manner. Furthermore, the entire TAAR OSN populations are eliminated when both of the TAAR enhancers are deleted. In transgenic animals bearing the TAAR enhancers, each enhancer is capable of specifically driving reporter expression in subsets of TAAR OSNs. We next provide genetic evidence that the TAAR enhancers act in *cis*. Taken together, our study reveals two enhancers located within a single gene cluster that coordinately control expression of an olfactory receptor gene subfamily.

Results

Enrichment of TAAR OSNs. In order to gain genetic access to TAAR-expressing OSNs, we generated *Taar5-ires-Cre* and *Taar6-ires-Cre* knockin mouse lines (Figure 1A), in which Cre recombinase is co-transcribed with *Taar5* and *Taar6* gene, respectively. Following transcription, Cre is independently translated from an internal ribosome entry site (IRES) sequence [36]. Each Cre line was crossed with Cre-dependent reporter lines, including *lox-L10-GFP* [37] and *lox-ZsGreen* [37], to fluorescently label the specific population of TAAR OSNs. The GFP- or ZsGreen-positive cells were then sorted by fluorescence-activated cell sorting (FACS) and enrichment of TAAR OSNs was verified by RNA-seq (Supplementary Figure 1A). The GFP- or ZsGreen-negative cells were also sorted to serve as control cells, approximately 70-80% of which are composed of OR-expressing mature OSNs. We indeed observed a dramatic increase in *Taar* gene expression and decrease in *OR* gene expression in sorted positive cells compared to control cells (Figure 1B).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Next, we examined the expression of individual *Taar* genes. There are 15 functional *Taar* genes and 1 pseudogene (*Taar7c*) clustered on mouse chromosome 10 forming the single *Taar* cluster without any other annotated genes. Consistent with the previous observation that *Taar1* is mainly expressed in the brain [38], we did not detect *Taar1* expression in the sorted cells (Figure 1C). Surprisingly, all of the other 14 functional *Taar* genes were detected at various expression levels in reporter-positive cells (Figure 1C). We expected to obtain pure *Taar5* or *Taar6* receptor expression as TAAR OSNs obey the "one-neuron-one-receptor" rule [2]. A parsimonious interpretation of this observation is that TAAR OSNs undergo receptor switching at much higher frequencies than OR OSNs [14]. Nevertheless, we have successfully enriched TAAR OSNs that express a mixture of different TAAR family members. In addition, we analyzed the canonical OSN signaling proteins and chaperones, including *Gnal*, *Adcy3*, *Cnga2*, *Ano2*, and *Rtp1*/*2*, in enriched TAAR OSNs. We observed comparable expression levels to control cells (Figure 1D, Supplementary Figure 1B), suggesting that TAAR OSNs use the same signaling pathways as OR OSNs.

The *Taar* cluster is covered by heterochromatic marks in TAAR OSNs but not in OR OSNs. Previous studies have shown that unlike the *OR* clusters, the *Taar* cluster is not decorated by heterochromatic silencing marks of H3K9me3, indicating a different regulatory mechanism on receptor gene choice [26]. However, the experiments were performed on the whole MOE tissue, which mostly consists of OR OSNs and contains less than 1% TAAR OSNs. Therefore, it is possible that H3K9me3 is excluded from the *Taar* cluster in OR OSNs and not in TAAR OSNs, but it is diluted below the level of detection in the whole MOE preparation. To examine this, we carried out H3K9me3 native ChIP-seq analysis on purified TAAR OSNs from *Taar5-ires-Cre*; *lox-L10-GFP* mice and found that the *Taar* cluster, as well as the *OR* clusters are actually marked with high levels of H3K9me3 modification (Figure 1E). By contrast, in the sorted GFP-negative cells, the *Taar* cluster is devoid of H3K9me3 repressive marks, whereas H3K9me3 marks are enriched in the *OR* clusters (Figure 1E), in agreement with the previous study [26]. This observation suggests that *Taar* gene expression

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

undergoes epigenetic regulation in TAAR OSNs analogous to the regulation of *OR* genes in OR OSNs, and that it may also require histone demethylases including LSD1. However, the mechanisms of *Taar* gene silencing in OR OSNs might be different from *OR* gene silencing in TAAR OSNs.

Identification of two putative TAAR enhancers. To further dissect the regulatory elements of *Taar* genes, we performed ATAC-seq to assay regions of open chromatin in TAAR OSNs sorted from *Taar5-ires-Cre*; *lox-L10-GFP*, *Taar5-ires-Cre*; *lox-ZsGreen* or *Taar6-ires-Cre*; *lox-ZsGreen* mice [39]. As a control, all mature OSNs that are mainly composed of OR OSNs were sorted from *Omp-ires-GFP* mice (Figure 2A). In total, we obtained 6 replicates of TAAR OSN samples and 3 replicates of OMP-positive mature OSN samples for ATAC-seq. We first examined a number of genes encoding olfactory signaling molecules, such as *Gnal*, *Adcy3*, *Cnga2*, *Ano2*, and *Rtp1*/*2*. Consistent with the high levels of these genes in RNA-seq data, we observed strong ATAC-seq peaks with comparable intensities in both TAAR and OMP-positive OSNs (Supplementary Figure 2). The peaks are mostly located near the promoter regions, which also allows us to determine the primary isoforms expressed in the MOE (Supplementary Figure 2).

Next, we screened for neural population-specific peaks using DiffBind package. By quantitatively comparing ATAC-seq peaks in TAAR and all mature OSNs, we were able to identify 6093 differential peaks with 3290 peaks enriched in OMP-positive OSNs and 2803 peaks enriched in TAAR neurons (Figure 2B). We then focused on two population-specific peaks in the genomic regions surrounding the *Taar* cluster. We found two peaks of about 900 bp that are highly enriched in TAAR OSNs, suggesting that they may function as putative TAAR enhancers to regulate *Taar* gene expression (Figure 2C). We termed these sequences TAAR enhancer1 and TAAR enhancer 2. TAAR enhancer 1 is positioned between *Taar1* and *Taar2*, while TAAR enhancer 2 is located between *Taar6* and *Taar7a*. Quantitative analysis of peak intensity after normalizing to the median values in each sample revealed a 6-fold and 4-fold increase

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

for TAAR enhancer 1 and TAAR enhancer 2 in TAAR OSNs compared to OMP-positive OSNs, respectively (Figure 2D). Thus far, 63 OR enhancers - known as "Greek islands" - have been identified and shown to contribute to *OR* gene choice in OR OSNs [15, 16]. We therefore tested if TAAR OSNs have limited chromatin accessibility at OR enhancers. Unexpectedly, we observed similar chromatin accessibility in TAAR OSNs to the whole mature OSN population (Figure 2E). The normalized ATAC-seq peak intensities at all of the OR enhancers showed positive linear correlation with the Pearson correlation coefficient of 0.89 ($p < 0.0001$, Figure 2F). The co-existence of putative TAAR enhancers and OR enhancers in TAAR OSNs indicates that they may form an enhancer hub to facilitate *Taar* gene choice. Consistent with this notion, we observed dramatic reduction of *Taar* gene expression after deletion of Lhx2 or Ldb1, which were shown to facilitate the formation of the OR enhancer hub formation [17] (Supplementary figure 3).

Evolutionary conservation of the putative TAAR enhancers. The *Taar* gene family is hypothesized to emerge after the segregation of jawed from jawless fish [23]. To identify the evolutionary origin of the two putative TAAR enhancers, we compared the *Taar* cluster sequences from 8 Glires, 21 Euarchontoglires, 40 placental mammals, and 60 vertebrates (Supplementary Figure 4A). The two putative TAAR enhancers are among the most conserved regions in addition to *Taar* genes (Figure 3A). We then focused on the two TAAR enhancers and searched for the publicly available genome databases with the mouse sequences of the TAAR enhancers. We retrieved homologous sequences from Eutheria or placental mammals, but failed to find any homologous sequences from closely related Metatheria or marsupial mammals. We next selected one representative species for each Eutheria order, including human, chimpanzee, tarsier, rat, rabbit, hedgehog, pig, cow, sperm whale, horse, cat, big brown bat, elephant, and armadillo. We selected koala and opossum as representative species of Metatheria, and platypus as a representative species of Prototheria (Figure 3B). The nucleotide sequences of TAAR enhancers as well as neighboring *Taar* genes from the various species were extracted and aligned with homologous mouse

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sequences. The VISTA plot revealed highly conserved synteny of the *Taar* genes in all the species selected (Figure 3C). In contrast, the two TAAR enhancers are only conserved in placental mammals but not in marsupial mammals (Figure 3C). Interestingly, TAAR enhancer 1 is present in all placental mammals except very few species, such as dolphin and gibbon (Supplementary Figure 4B), while TAAR enhancer 2 is less conserved and more variable in different species of placental mammals. We observed loss of TAAR enhancer 2 in dog, sea lion, seal of the Carnivora order, and in whale, dolphin of the Cetartiodactyla order, and also in gibbon. In addition, we found duplications of TAAR enhancer 2 in cow, sheep of the Cetartiodactyla order, and also in horse of the Perissodactyla order, which may be accompanied by genome duplications of *Taar6* and *Taar7* family members (Supplementary Figure 4C). In conclusion, the sequences of the two TAAR enhancers are conserved in placental mammals, supporting the notion that they play an essential role in *Taar* gene regulation.

Deletion of TAAR enhancer 1 results in specific reduction of TAAR OSNs. To examine the function of TAAR enhancer 1, we generated TAAR enhancer 1 knockout mice using CRISPR-Cas9 genome editing system (Figure 4A, B). We then performed RNA-seq on the MOE dissected from homozygous, heterozygous, and wild type littermates to obtain transcriptomic profiles for each genotype. We first compared the gene expression levels of TAAR enhancer 1 homozygous knockout mice and wild type mice. Using the criteria of $q < 0.05$ (q value is false discovery rate-corrected p value) and fold change > 1.5-fold, we detected 14 differentially expressed genes (DEGs), 9 of which are the *Taar* genes (Supplementary Figure 5A). Thus, we carefully inspected the expression changes of all *Taar* genes and genes close to the *Taar* cluster (Figure 4C, Supplementary Figure 5B). Of the 14 functional olfactory *Taar* genes (*Taar1* is not expressed in the MOE and *Taar7c* is a pseudogene), the mRNA levels of 8 receptors (*Taar2, Taar3, Taar4, Taar5, Taar6, Taar7a, Taar7b, and Taar9*) were significantly decreased in TAAR enhancer 1 knockout mice (fold change < -1.5, $q < 0.05$), whereas the mRNA level of a single *Taar* gene, *Taar7e*, was increased (fold change > 1.5, q <

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

0.05). In addition, the mRNA levels of 3 *Taar* genes (*Taar8a, 8b, 8c*) tended to decrease (fold change < -1.5, q > 0.05), and those of 2 *Taar* genes (*Taar 7d, 7f*) were unaltered (-1.5 < fold change <1.5, q > 0.05). It is intriguing that deletion of TAAR enhancer 1 had dramatic effects on the expression of *Taar* genes that are located in both proximal and distal positions within the *Taar* cluster, but had minimal effects on expression of *Taar* genes in the center of the cluster. This observation suggests that TAAR enhancer 1 might form a loop to affect *Taar* gene expression over around 180 kb genomic region. The effect of TAAR enhancer 1 knockout is specific to *Taar* genes, as expression levels of other genes located next to the *Taar* cluster were not changed (Figure 4C). We also checked the expression patterns of *OR* genes and found no significant changes for all of the *OR* genes (q > 0.05), further suggesting the specific effect of TAAR enhancer 1 on the *Taar* genes (Figure 4D, Supplementary Figure 5C).

TAAR enhancer 1 regulates the probability of *Taar* gene choice. The observed changes in *Taar* mRNA expression may be due to altered probability of *Taar* gene choice or altered *Taar* transcript levels. To distinguish between these two possibilities, we performed *in situ* hybridization experiments and quantified cells with positive mRNA expression signals (Figure 4E, F). Consistent with our RNA-seq results, cells expressing *Taar2*, *Taar3*, and *Taar5* were totally abolished in homozygous TAAR enhancer 1 knockout mice, while the numbers of cells expressing *Taar4*, *Taar6*, *Taar8s*, and *Taar9* were significantly decreased. No obvious differences in the numbers of cells expressing *Taar7s* were detected. However, our probes targeting *Taar* gene coding regions cannot differentiate among the various members of *Taar7* and *Taar8* family genes. We therefore designed specific probes that distinguish among the various *Taar7* members based on their unique 3'-UTR sequences deduced from RNA-seq results. In TAAR enhancer 1 knockout mice, the numbers of cells expressing *Taar7a*, *Taar7b*, and *Taar7d* were significantly decreased, whereas, although not significant, it tended to decrease for *Taar7f* and to increase for *Taar7e* (Figure 4F). Furthermore, we found that the changes in the numbers of TAAR OSNs were in positive linear correlation with changes in mRNA expression from the RNA-seq data

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Pearson correlation coefficient r = 0.94, p = 4.9×10^{-6}, Figure 4G). Thus, like the H, P, and J elements, TAAR enhancer 1 regulates the probability of receptor gene choice rather than the transcript levels per cell [19, 22]. In agreement with the RNA-seq data, we did not detect any significant differences in the number of cells expressing two *OR* genes (dorsally expressed *Olfr578* and ventrally expressed *Olfr1507*) between wild type and TAAR enhancer 1 knockout mice (Figure 4E, F).

To further validate these findings, we performed immunohistochemistry analysis on the MOE using specific antibodies against TAAR5 and TAAR6 proteins. Again, we observed complete abolishment of TAAR5-positive OSNs (Figure 4H) and decreased numbers of TAAR6-positive OSNs in TAAR enhancer 1 knockout mice (Supplementary Figure 5D), whereas the number of OSNs expressing Olfr552 and Olfr1507 were not significantly changed (Figure 4H, Supplementary Figure 5D).

The dramatic reduction of TAAR OSNs in TAAR enhancer 1 knockout mice may be due to neuronal cell death as a result of the inability to express functional TAARs. We counted the number of apoptotic cells in the MOE by caspase-3 staining and did not find significant difference between TAAR enhancer 1 knockout and wild type mice (Supplementary Figure 5E), suggesting that deletion of TAAR enhancer 1 does not induce cell death. It is possible that OSNs originally choosing TAARs may arrest neuronal differentiation after TAAR enhancer 1 deletion.

Deletion of TAAR enhancer 2 causes decrease of TAAR OSNs in a pattern that is slightly different from deletion of TAAR enhancer 1. Next, to examine the function of TAAR enhancer 2, we generated TAAR enhancer 2 knockout mice by CRISPR-Cas9 genome editing (Figure 5A, B). RNA-seq on the MOE dissected from homozygous, heterozygous, and wild type littermates were performed to obtain transcriptomic profiles for each genotype. Using the criteria of $q < 0.05$ and fold change > 1.5-fold, we detected 6 DEGs, 5 of which were the *Taar* genes (Supplementary Figure 6A). Of the 14 functional *Taar* genes, the mRNA levels of 5

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

receptors (*Taar3, Taar5, Taar6, Taar7a,* and *Taar9*) were significantly decreased in homozygous TAAR enhancer 2 knockout mice compared to their wild type littermates (fold change < -1.5, q < 0.05). In addition, the mRNA levels of 5 *Taar* genes (*Taar4, Taar7f, Taar8a, Taar8b,* and *Taar8c*) tended to decrease (fold change < -1.5, q > 0.05), and those of 4 *Taar* genes (*Taar2, Taar7b, Taar7d,* and *Taar7e*) were unaltered (-1.5 < fold change <1.5, q > 0.05) (Figure 5C, Supplementary Figure 6A, 6B). Similar to TAAR enhancer 1, deletion of TAAR enhancer 2 did not change the expression levels of neighbor genes in the *Taar* gene cluster or any of the *OR* genes (Figure 5C, 5D, Supplementary Figure 6C). We then performed *in situ* hybridization experiments to quantify positive cells expressing olfactory receptor mRNAs. In accordance with our RNA-seq results, the numbers of TAAR OSNs expressing *Taar3*, *Taar4*, *Taar5*, *Taar6*, *Taar7a*, *Taar7f*, *Taar8s*, and *Taar9*, were dramatically decreased in TAAR enhancer 2 knockout mice (Figure 5E, 5F). The changes in the numbers of TAAR OSNs are in positive linear correlation with changes in mRNA expression from RNA-seq data (Pearson correlation coefficient r = 0.88, p = 0.0001), suggesting that TAAR enhancer 2 also regulates the probability of receptor gene choice (Supplementary Figure 6D). Further analysis showed that cells expressing *Taar3* and *Taar5* were completely abolished in TAAR enhancer 2 knockout mice, which was also observed in TAAR enhancer 1 knockout mice. However, we also observed subpopulations of TAAR OSNs in which the effects in TAAR enhancer 2 knockout mice were different from those in TAAR enhancer 1 knockout mice. For instance, the cells expressing *Taar2* were totally eliminated in TAAR enhancer 1 knockout mice (Figure 4E, F), but their numbers were unchanged in TAAR enhancer 2 knockout mice (Figure 5E, F). And the numbers of cells expressing *Taar7b* and *Taar7d* were significantly decreased in TAAR enhancer 1 knockout mice (Figure 4E, F), but were unaltered in TAAR enhancer 2 knockout mice (Figure 5E, F). On the contrary, the numbers of cells expressing *Taar7f* were largely decreased in TAAR enhancer 2 knockout mice (Figure 5E, F), but were not changed in TAAR enhancer 1 knockout mice (Figure 4E, F). As a control, we did not detect significant changes of cells expressing *Olfr578* and *Olfr1507* between wild type and TAAR enhancer 2 knockout mice (Figure 5E, F). These results suggest that

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

TAAR enhancers 1 and 2 may function jointly, redundantly, or separately to regulate expression of different *Taar* genes.

We next verified the above observation by immunostaining using TAAR5 or Olfr552 antibodies. We observed complete abolishment of TAAR5-positive OSNs, whereas OSNs expressing Olfr552 were not significantly changed (Supplementary Figure 6E). We also counted the numbers of apoptotic cells in the MOE by caspase-3 staining and did not find significant difference between TAAR enhancer 2 knockout and wild type mice (Supplementary Figure 6F). This suggests that deletion of TAAR enhancer 2 may lead to neuronal differentiation arrest instead of neuronal apoptosis, analogous to the effects of TAAR enhancer 1 deletion.

Deletion of the two TAAR enhancers results in complete elimination of TAAR OSNs. Our results show that deletion of either one of the two enhancers decreases the expression of distinct groups of *Taar* genes. To investigate if the whole *Taar* gene repertoires are fully controlled by the two enhancers, we generated a mouse line (TAAR enhancer 1 & 2 double knockout) with both enhancers deleted (Figure 6A, B). We then counted the numbers of TAAR OSNs in the double knockout mice and their littermate controls. Strikingly, all of the TAAR OSNs were completely eliminated after both of the enhancers were deleted (Figure 6C, D). Again we did not detect significant changes of cells expressing *Olfr578* and *Olfr1507* between wild type and TAAR enhancer 1 & 2 double knockout mice (Figure 6E, F). Taken together, the loss-of-function experiments suggest that the two TAAR enhancers we identified are specifically and absolutely required for *Taar* gene expression.

TAAR enhancers are sufficient to drive reporter expression in the TAAR OSNs. To provide further evidence for the function of the two TAAR enhancers, we then tested if they are sufficient to drive expression of neighboring genes in OSNs. We firstly detected activities of the two TAAR enhancers in zebrafish as the majority of the OR enhancers can drive reporter expression in zebrafish OSNs [15]. Indeed, the two

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

TAAR enhancers induced GFP reporter expression in the MOE of zebrafish, similar to the OR enhancers (Supplementary Figure 7A-C).

To further examine the functional roles of the TAAR enhancers, we constructed PiggyBac transgenic plasmids with one of the two TAAR enhancers placed upstream of the minimal promoter sequence from mouse Hsp68 (heat shock protein 68kDa) and followed by the *GFP* or *tdTomato* reporter gene (Figure 7A). We then generated transgenic mice by injecting PiggyBac plasmids into the pronucleus of fertilized eggs. For TAAR enhancer 1-GFP transgenic mice, we obtained 12 founder lines, 9 of which exhibited robust GFP expression in OSNs (Figure 7A). We kept one TAAR enhancer 1 transgenic line for breeding and further analysis. The GFP-positive OSNs were located in both dorsal and ventral domain of the MOE (data not shown), where the *Taar* genes are normally expressed [2, 25, 26]. To examine whether TAAR enhancer 1 is indeed specific for TAAR OSNs, we analyzed if GFP-positive OSNs overlapped with TAAR or OR OSNs using mixed *Taar* or *OR* probes. About 84.3% of GFP-positive OSNs were co-labeled with *Taar* genes, and the overlap dropped to 4.1% for *OR* genes ($p < 0.0001$) when the same number of mixed receptor probes were used. We also observed a decreased overlap of 0.3% for *OR* degenerate probe ($p < 0.0001$) (Figure 7B, C). By contrast, around 14.4% of TAAR OSNs expressed GFP reporter, and the overlap dropped to 1.9% for mixed *OR* probes ($p < 0.0001$) and 0.5% for *OR* degenerate probe ($p < 0.0001$) (Figure 7B, C). For TAAR enhancer 2-tdTomato transgenic mice, we obtained 5 founder lines, 3 of which exhibited robust tdTomato expression in OSNs (Figure 7A). We kept one TAAR enhancer 2 transgenic line and examined if TAAR enhancer 2 is indeed specific for TAAR OSNs. We found that about 73.2% of tdTomato-positive OSNs were co-labeled with *Taar* genes, and the overlap dropped to 1% for *OR* genes ($p < 0.0001$) when the same number of mixed receptor probes were used. The overlap percentage is around 0.3% for *OR* degenerate probe ($p < 0.0001$) (Figure 7D, E). On the other hand, around 13.9% of TAAR OSNs expressed tdTomato reporter, and the overlap dropped to 0.2% for mixed *OR* probes ($p < 0.0001$) and 0.05% for *OR* degenerate probe ($p < 0.0001$) (Figure 7D, E).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Together, these results validate the specific activity of the two TAAR enhancers in TAAR OSNs.

The aforementioned loss-of-function experiments strongly suggest that the two TAAR enhancers work coordinately to regulate *Taar* gene expression. The transgenic mice provide another good system for us to examine the coordination between the two TAAR enhancers. We crossed the TAAR enhancer 1-GFP transgenic mice with the TAAR enhancer 2-tdTomato transgenic mice, and visualized the overlap between GFP- and tdTomato-positive cells. About 14.3% of OSNs co-expressed GFP and tdTomato driven by TAAR enhancer 1 and TAAR enhancer 2, respectively. On the other hand, 67.5% and 18.2% of OSNs expressed each reporter alone (Figure 7F). We further analyzed the projection pattern of OSNs into the olfactory bulb. The GFP-postive and tdTomato-positive glomeruli were largely overlapped in the dorsal region of olfactory bulb, where TAAR glomeruli cluster (called DIII domain [25]). We also observed a few glomeruli that were positive for GFP or tdTomato alone (Figure 7G). Thus, these data again indicate that the two TAAR enhancers function coordinately to regulate the development of TAAR OSNs.

TAAR enhancer 1 functions in *cis*. Each OSN in the MOE expresses an olfactory receptor gene in a monogenic and monoallelic fashion. However, we observed similar expression levels of *Taar* genes in heterozygous TAAR enhancer 1, TAAR enhancer 2, or TAAR enhancer 1 & 2 double knockout mice compared to their wild type littermates (Figure 4E, 4F, 5E, 5F, 6C, 6D, Supplementary Figure 5B, 6B), seemingly violating the monoallelic expression rule. A possible explanation for this observation is that the intact enhancers in heterozygous mice could act both in *cis* and in *trans* to regulate *Taar* gene expression in both alleles. Another explanation is that TAAR OSNs are genetically programmed to express *Taar* genes. As a result, although TAAR enhancers may just operate in *cis* like OR enhancers, TAAR OSNs can only express *Taar* genes from the allele with the functional enhancers, thereby maintaining the number of TAAR OSNs. To differentiate between these two possibilities, we crossed

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

TAAR enhancer 1 knockout mice with *Taar2-9* cluster knockout mice, in which all of the olfactory *Taar* genes from *Taar2* to *Taar9* are deleted [40]. The successful deletion of the olfactory *Taar* genes was verified by the absence of the mRNA and protein expression of TAARs in the MOE by RNA-seq, RNA *in situ* hybridization, and immunohistochemistry experiments (Supplementary Figure 8A-C). Although TAAR enhancer 2 is deleted in the *Taar2-9* cluster knockout mouse line, TAAR enhancer 1 is preserved. By crossing heterozygous TAAR enhancer 1 knockout mice (designated as ΔTAAR-enhancer1/TAAR-enhancer1 here for clarity) with heterozygous *Taar2-9* cluster knockout mice (Δ*Taar2-9*/*Taar2-9*), we acquired four genotypes of mice for further analysis: (1) wild type littermates (genotype 1) with both alleles intact; (2) ΔTAAR-enhancer1; *Taar2-9*/TAAR-enhancer1; *Taar2-9* (genotype 2) with one allele lacking TAAR enhancer 1 and the other allele intact; (3) TAAR-enhancer1; *Taar2-9*/TAAR-enhancer1; Δ*Taar2-9* (genotype 3) with one allele lacking *Taar2-9* cluster and the other allele intact; and (4) ΔTAAR-enhancer1; *Taar2-9*/TAAR-enhancer1; Δ*Taar2-9* (genotype 4) with one allele lacking TAAR enhancer 1 and the other allele lacking *Taar2-9* cluster (Figure 8A). Next, we performed RNA *in situ* hybridization experiments to examine the expression of *Taar* genes that were entirely eliminated (*Taar2*, *Taar3*, and *Taar5*) or significantly reduced (*Taar6*) in homozygous ΔTAAR-enhancer1 mice (Figure 4E, F, H). In ΔTAAR-enhancer1; *Taar2-9*/TAAR-enhancer1; Δ*Taar2-9* (genotype 4) mice, cells expressing these *Taar* genes showed similar reduction pattern to that in homozygous ΔTAAR-enhancer1 mice (Figure 8B, C). These results suggest that TAAR enhancer 1 operates in *cis* to regulate *Taar* gene choice (Figure 8A), in analogy to the OR enhancers [19-22].

Discussion

In the present study, we identified two enhancers that specifically regulate probability of *Taar* gene choice and thus stabilize the cell fate decision of the TAAR subsystem. Deletion of TAAR enhancer 1 in both alleles resulted in complete disappearance of 3 *Taar* genes (*Taar2*, *Taar3*, and *Taar5*), decrease of 8 *Taar* genes (*Taar4*, *Taar6*,

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Taar7a, *Taar7b*, *3 Taar8s*, and *Taar9*), slight increase of 1 *Taar* gene (*Taar7e*), and no significant change of 2 *Taar* genes (*Taar7d*, *Taar7f*); while homozygous knockout of TAAR enhancer 2 resulted in abolishment of 2 *Taar* genes (*Taar3* and *Taar5*), decrease of 8 *Taar* genes (*Taar4*, *Taar6*, *Taar7a*, *Taar7f*, 3 *Taar8s*, and *Taar9*), and 4 unchanged *Taar* genes (*Taar2*, *Taar7b*, *Taar7d*, and *Taar7e*). It seems that various *Taar* genes are differentially regulated by the two enhancers. For example, expression of some *Taar* genes (*Taar3* and *Taar5*) requires presence of both enhancers and deletion of either one completely abolishes their expression. And expression of some *Taar* genes (*Taar2*) is solely dependent on one TAAR enhancer. While expression of the majority of *Taar* genes requires either TAAR enhancer, meaning that the two enhancers have partially overlapped function. Thus, the two TAAR enhancers work coordinately to achieve expression of the whole *Taar* gene repertoires. This conclusion is further supported by analyses of the double knockout animals lacking both of the TAAR enhancers. In the double knockout mice, all of the TAAR OSNs are completely eliminated, suggesting that the two enhancers are fully responsible for *Taar* gene expression and hence TAAR OSN development. This organization resembles the regulation of protocadherin-α family genes by the combined activity of two enhancers [41], and thus provides another great model to study gene regulation by cooperative enhancers.

Our data revealed that TAAR enhancer 1, and possibly TAAR enhancer 2, operate in *cis*, which is similar to previously identified OR enhancers, including the H, P, and J elements [19-22]. However, the phenotypes of heterozygous enhancer knockout mice are very different. In mutant mice lacking the H or P elements in one allele, the number of cells expressing the specific class II OR genes regulated by those two elements is decreased to half of that in wild type mice [19]. In contrast, when the J element or the two TAAR enhancers are deleted in one allele, the number of cells expressing class I OR genes or *Taar* genes is nearly the same as that in wild type mice. This is consistent with the hypothesis that OSNs have dedicated cell fates to express the various olfactory receptor gene families (class I *OR* genes, class II *OR* genes, and *Taar* genes)

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

prior to the first receptor choice [22, 25, 26]. When the OR enhancers that regulate expression of class II *OR* genes (*e.g.* the H and P elements) are deleted in one allele, OSNs can express class II *OR* genes from the remaining intact enhancer in the other allele and more than 60 functional OR enhancers in both alleles. As a result, the expression probability of distinct class II *OR* genes regulated by the intact enhancer is largely diluted, leading to the decrease of OSNs expressing those class II *OR* genes by half. By contrast, when the J element or either one or both of the two TAAR enhancers are deleted in one allele, receptor selection is restricted to the intact enhancer from the other allele. Thus, the remaining intact enhancer still plays a major role in receptor selection and fill in the OSN population. As a result, the numbers of OSNs expressing class I *OR* genes or *Taar* genes are similar in heterozygous and wild type animals.

The next intriguing question is how the cell fate of OSNs is determined prior to the first receptor choice. OSNs in the MOE are continuously renewed from basal stem cells. It is conceivable that certain cell-specific transcription factors, nucleosome remodeling complexes, and epigenetic regulators predefine the cell fate of OSNs during stem cell differentiation. Identification of such factors will greatly advance our understanding of the development of the distinct olfactory subsystems. Besides, TAAR enhancers emerge after the separation of placental mammals from marsupial mammals and are not found in zebrafish. However, our transient reporter assay showed that the two mouse TAAR enhancers are capable of driving reporter expression in the nose of zebrafish larvae (Supplementary Figure 7A-C). This result has two implications: (1) the teleost species may utilize other uncharacterized teleost-specific enhancers to regulate expression of the largely expanded *Taar* genes (e.g. 112 members in zebrafish); (2) although TAAR enhancers are not conserved in zebrafish, the regulatory factors proposed above might be conserved to fulfill the TAAR enhancer activities.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The *Taar* gene cluster and OR gene clusters are sequestered in different nuclear compartments. In OR OSNs, silent OR gene clusters aggregate to form ~ 5 heterochromatic foci in the center of cell nuclei, while the *Taar* gene cluster is localized to the thin rim at the nuclear periphery [18, 42, 43]. However, the nuclear locations of the *Taar* gene cluster and the OR gene clusters in TAAR OSNs have not been carefully examined. Here, we found that the heterochromatic histone modifications of the *Taar* gene cluster differ between the two olfactory subsystems. Although OR gene clusters are covered by H3K9me3 heterochromatin modifications in both TAAR and OR OSNs, the same heterochromatin marks of *Taar* gene cluster are only present in TAAR OSNs. This observation may indicate that the conformation of the *Taar* gene cluster might be different in the nucleus of TAAR and OR OSNs. On the other hand, activation of receptor genes is often accompanied by a shift in their nuclear localization. In OR OSNs, the active *OR* allele escapes from the central heterochromatin foci into euchromatic territory with the help of OR enhancer hub [17, 42]. In TAAR OSNs, the active *Taar* allele is thought to transition from the peripheral nuclear lamina to a more permissive interior euchromatin center [43]. The movement of the *Taar* gene cluster away from the repressive heterochromatin environment may need assistance of the two TAAR enhancers. In addition, the two TAAR enhancers are likely to form a TAAR enhancer hub with OR enhancers based on the finding that OR enhancers are also open in TAAR OSNs. The recent development and application of single-cell chromatin conformation capture method [18, 44] could help to reveal the three-dimensional structure of the TAAR enhancer hub in TAAR OSNs and the potential interaction between the two TAAR enhancers with distinct *Taar* genes at higher resolution. This may explain the differential regulation of *Taar* genes by the two TAAR enhancers. Therefore, revealing the spatial organization of the *Taar* gene cluster and the two TAAR enhancers in TAAR vs. OR OSNs would further advance our understanding of how the entire *Taar* gene repertoire is regulated.

Methods

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Mice. All mouse experiments were approved by the Animal Ethics Committee of Shanghai Jiao Tong University School of Medicine and the Institutional Animal Care and Use Committee (Department of Laboratory Animal Science, Shanghai Jiao Tong University School of Medicine, animal protocol number A-2016-049). Mice were housed in standard conditions with a 12-hour light/12-hour dark cycle and have access to rodent chow and water ad libitum. Both male and female mice were used for experiments. Generation of genetically manipulated mouse lines including *Omp-ires-GFP* and *Taar2-9* cluster knockout mice have been described previously [14, 40].

We used gene targeting in ES cells to generate two strains of mice, *Taar5-ires-cre* and *Taar6-ires-cre*, that express Cre recombinase in TAAR5- and TAAR6-expressing OSNs. To generate the *Taar5-ires-cre* allele, the DNA sequence encoding Cre was introduced into the 3' untranslated sequence in the *Taar5* gene immediately following the stop codon with an internal ribosome entry site (IRES) sequence preceding the Cre sequence to allow bicistronic expression of TAAR5 and Cre [45]. To select positive ES clones, the target construct included a cassette of the neomycin gene flanked by FRT sites. The neomycin gene was removed by crossing the mice to a germline Flp recombinase line. The same strategy was used to generate the *Taar6-ires-cre* allele.

TAAR enhancer 1 or TAAR enhancer 2 was deleted using the CRISPR-Cas9 genome editing technique with co-injection of spCas9 mRNA and sgRNAs. Two groups of sgRNAs were designed at upstream and downstream of the target sequences. The targeted sequences for TAAR enhancer 1 were 5'-TGGTTGTGAGTTGCTTGTGG-3' (sgRNA1 sense), 5'-TCAGCCTGTTAATTACCTGA-3' (sgRNA1 antisense), 5'-AGAACTTTCAGAGAGTTCCC-3' (sgRNA2 sense), 5'-GAACCCAGAACTGACTTTTG-3' (sgRNA2 antisense), 5'-TATTCTAGAAATACAGATGT-3' (sgRNA3 sense), and 5'-AGCATCCTGGAGGTGAAATG-3' (sgRNA3 antisense). The targeted sequences for TAAR enhancer 2 were 5'-GTAAATAAAAACTTTCCCTC-3' (sgRNA1 sense), 5'-CTCCATCGTCACAAAGCCTG-3' (sgRNA1 antisense), 5'-

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

CCCTCAAAAAGTTTGTTTTT-3' (sgRNA2 sense), and 5'-CAGGTCTTTTTTAGTGGACT-3' (sgRNA2 antisense). The RNA mixtures (50 ng spCas9 mRNA per μl and 100ng sgRNA mixture per μl) were injected into zygotes of C57BL/6J mice. The two-cell stage embryos were then transferred into surrogate mothers. We obtained 11 candidate founders for TAAR enhancer 1 knockout mice and 8 candidate founders for TAAR enhancer 2 knockout mice. We eventually established one TAAR enhancer 1 knockout line with a 1372 bp deletion (mm10, chr10: 23,929,882-23,931,253) and one TAAR enhancer 2 knockout line with a 1433 bp deletion (mm10, chr10: 23,987,372-23,988,804) that were used for RNA-seq, *in situ* hybridization, and immunohistochemistry experiments.

TAAR enhancer 1 and 2 double knockout mice were generated by injecting spCas9 mRNA and the above sgRNAs targeting TAAR enhancer 1 into zygotes of TAAR enhancer 2 homozygous or heterozygous mice. We obtained 3 candidate founders and eventually established one double knockout line with a 1205 bp deletion (mm10, chr10: 23,930,050-23,931,254) for TAAR enhancer 1 and a 1433 bp deletion (mm10, chr10: 23,987,372-23,988,804) for TAAR enhancer 2.

The TAAR enhancer 1 or TAAR enhancer 2 transgenic mice were generated using the PiggyBac Transposon system. Briefly, TAAR enhancer 1 (mm10, chr10: 23,930,200-23,931,190, 991 bp) or TAAR enhancer 2 (mm10, chr10: 23,987,501-23,988,362, 862 bp) was cloned into the modified PiggyBac transposon vector, in which enhancers were followed by Hsp68 minimal promoter and EGFP (for TAAR enhancer 1) or tdTomato (for TAAR enhancer 2) sequences. The validated plasmid (100 ng per μl) as well as PiggyBac transposase mRNA (25 ng per μl) were co-injected into fertilized eggs of C57BL/6J mice. The two-cell stage embryos were then transferred into surrogate mothers.

Fluorescence-activated cell sorting. Mice were sacrificed with CO_2 followed by cervical dislocation. The MOE tissue was dissected and cells were dissociated using

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Papain Dissociation System (Worthington Biochemical) following manufacturer's instructions with minor modifications. Briefly, dissociation reaction was incubated at 37 °C for 15 minutes. The tissue was triturated for 10-15 times with a cut P1000 pipette tip. Cells were then filtered by 40 μm strainer (Falcon) and centrifuged at 400 g for 2 minutes. Cell pellets were resuspended in DMEM (Gibco) and kept on ice for sorting. OSNs were sorted on a FASCJazz Cell Sorter (BD) with a 488-nm laser. A representative set of density plots (with gates) is shown in Supplementary Figure 1. Sorted cells were subsequently proceeded to ChIP-seq, ATAC-seq, and RNA-seq experiments.

Ultra-low-input native ChIP-seq (ULI-ChIP-seq). ULI-ChIP-seq of H3K9me3 was performed as previously described [46] except for the library preparation portion. For each reaction, 1,000–10,000 cells and 0.25 μg of antibody (ABclonal, A2360) were used. Libraries were prepared with an NEBNext Ultra kit (NEB) and 12–15 PCR cycles. Reads were mapped to the mouse reference genome (mm10) with Bowtie 2 (version 2.2.9) using the default parameters.

ATAC-seq. ATAC-seq was performed as previously described with minor modifications [47]. During qPCR-based library quantification, a length of 300 bp was used in concentration calculation. Reads were mapped to the mouse reference genome (mm10) using Bowtie 2 (version 2.2.9) with the following parameters: --local –X 2000. Significant ATAC-seq peaks were identified using MACS2 (version 2.1.1.20160309) with the "--nomodel --shift -100 --extsize 200 --keep-dup all" parameters and the default threshold of q value (adjusted p value) at 0.05. Differential peak analysis between TAAR OSNs and OMP-positive mature OSNs was performed using DiffBind (version 2.8.0) in R (version 3.5.1) with standard parameters.

RNA-seq. Total RNAs of FACS enriched cells or the whole MOE tissues were extracted with a RNeasy Mini kit (Qiagen) with DNase treatment. Libraries of FACS enriched cells were prepared with the Smart-seq2 procedure [48], similar to our previous

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

work [49]. Libraries of the whole MOE tissues were prepared with an NEBNext Ultra Directional RNA kit (NEB) with polyA bead selection. Reads were mapped to the mouse reference genome (mm10) with Hisat2 (version 2.1.0) and quantified with Cufflinks (version 2.2.1). Raw read counts mapped to genes were calculated using featureCounts from Subread package (version 1.6.2). DEG analysis was performed using DESeq2 1.20.0 package with the default settings. Significantly changed genes were identified with q value cutoff of 0.05.

Whole-mount imaging. Unfixed whole-mount olfactory bulbs from mice at P14-P21 were exposed and fluorescent reporters were directly visualized and imaged by a Nikon Ti2-E&CSU-W1 confocal microscope.

***In situ* hybridization.** *In situ* hybridization was performed as previously described [2]. Full-length, anti-sense cRNA riboprobes labeled with digoxigenin were prepared to detect mRNA expression of *Taars*, *Olfr1507*, and *Olfr578* in single color *in situ* hybridization experiments. To specifically detect the various *Taar7* family genes, we used cRNA riboprobes labeled with digoxigenin as follows: *Taar7a* (656 bp sequence amplified by primers 5'-GATTCTTGGTTTAGTTGGGG-3' and 5'-TGTAAATCTTGAATGGGTC-3'), *Taar7b* (1049 bp sequence amplified by primers 5'-GATTCTTGGTTTAGTTGGGGA-3' and 5'-GCACACCTTTGAAAACTTC-3'), *Taar7d* (1044 bp sequence amplified by primers 5'-GATTCTTGGTTTAGTCGGGG-3' and 5'-CTACTGAGCTACCTCTTCAG-3'), *Taar7e* (704 bp sequence amplified by primers 5'-TTTCTCTTGCTCCCAGGTTCTC-3' and 5'-GTGCATATCTTTGAACACTTC-3'), and *Taar7f* (500 bp sequence amplified by primers 5'-ACTCTGTCAACTGAGGCTCAG-3' and 5'-CAAGAGCAAATTATACAGGG-3').

For two-color *in situ* hybridization experiments, mixed riboprobes labeled with digoxigenin against all of the *Taar* genes or 8 *OR* genes (*Olfr103*, *Olfr145, Olfr180, Olfr277, Olfr578, Olfr644, Olfr1034, Olfr1507*) and riboprobes labeled with fluorescein against GFP (720 bp sequence amplified by primers 5'-ATGGTGAGCAAGGGCGA-3'

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

AND 5'-TTACTTGTACAGCTCGTCCATGC-3') were used. Degenerate probes labeled with digoxigenin used to detect *OR* expression were generated as previously described [26]. Images were taken using a Leica TCS SP8 confocal microscope. To quantify the number of cells expressing *Taars* or *ORs*, every 25th coronal section (14 μm thickness) throughout the MOE of mice at P14 was collected. At least five sections at similar anatomical positions in the MOE from mice with different genotypes were used to count positive cell numbers.

Immunohistochemistry. Coronal MOE sections (14 μm thickness) from P14 mice were fixed with 4% paraformaldehyde in PBS for 10 minutes at room temperature. The sections were washed with PBS three times (5 minutes each) and incubated with permeable buffer (0.3% Triton X-100 in PBS) containing 5% donkey serum for 30 minutes. Primary antibodies against TAAR4, TAAR5, TAAR6 (homemade), caspase-3 (Cell Signaling, 9661), GFP (Abcam, ab13970), and tdTomato (Takara, 632496) were used at 1:5,000, 1:5,000, 1:1,000, 1:500, 1:1,000, and 1:500 dilution in incubation buffer (1% BSA, 0.01% sodium azide, 0.3% Triton X-100 in PBS). Primary antibody incubations were performed at 4 °C for two overnights. The sections were then rinsed three times (5 minutes each) in PBS, and incubated with different fluorophore-conjugated secondary antibodies (Jackson ImmunoResearch) at 37 °C for 30 minutes. Slides were rinsed three times (5 minutes each) and coverslipped using mounting medium containing DAPI (SouthernBiotech). Images were taken using a Leica TCS SP8 confocal microscope.

Phylogenetic tree building. The selected species for phylogenetic tree building are mammals that have the predicted *Taar* gene sequences and at least one of the two potential TAAR enhancers. The phylogenetic tree was built using TimeTree (http://www.timetree.org/) [50].

Nucleotide percent identity plot. Using NCBI Genome Data Viewer (https://www.ncbi.nlm.nih.gov/genome/gdv/), we compared mouse genome sequence

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ranging from *Taar1* to *Taar7b* with that in 8 Glires, 21 EuarchontoGlires, 40 Placental mammals and 60 Vertebrates by phyloP. The display style was chosen as line graph with smooth curve. The selected species and phylogenetic relationships were shown in Supplementary Figure 3A. Next, the 991 bp TAAR enhancer 1 (mm10, chr10: 23,930,200-23,931,190) and 862 bp TAAR enhancer 2 (mm10, chr10: 23,987,501-23,988,362) sequences were separately blasted against the latest genome assembly of each species from NCBI, with the BLAST algorithm optimized for somewhat similar sequences (blastn) for more hits. The aligned sequences with the adjacent *Taar* gene sequences were downloaded for further analysis. The sequences for global alignment were extracted from the entire region of the *Taar* cluster. This approximately 295 kb mouse genomic region was compared with the same region in other species by using the web-based VISTA program (http://genome.lbl.gov/vista/mvista/submit.shtml) [51]. The local alignment of TAAR enhancer 1 and TAAR enhancer 2 sequences from different species were performed using the same program with the default settings of Shuffle-LAGAN alignment program and VISTA parameters. The setting of conservation identity was changed to 70% and that of calculation window and minimal conserved width was changed to 100 bp. The information of compared regions and selected genome databases was listed in Supplementary Table 1.

Reporter transgenic zebrafish. Sequences of the two TAAR enhancers and the Sifnos OR enhancer were amplified from mouse genomic DNA using the following primers containing the restriction enzyme sites: 5'-CAGATGGGCCCTCGAGAATGCACCAGTGCTCGTTGTG-3' (TAAR enhancer 1 forward), 5'-TAGAGTCGAGAGATCTTGTATGTAGCTGATGTCAGTATCTAGC-3' (TAAR enhancer 1 reverse), 5'-TAGAGTCGAGAGATCTTGTATGTAGCTGATGTCAGTATCTAGC-3' (TAAR enhancer 2 forward), 5'-TAGAGTCGAGAGATCTGACCAGCAGATGAAGAAAG-3' (TAAR enhancer 2 reverse), 5'-CAGATGGGCCCTCGAGCACCCCCAAGGGATTCAATG-3' (Sifnos enhancer forward), and 5'-TAGAGTCGAGAGATCTATAACTTGCTTCAAGACATGTG-3'

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Sifnos enhancer reverse). PCR products were cloned into the E1B-GFP-tol2 vector via XhoI and BglII restriction sites [52]. The cloned plasmids were purified by Qiagen purification kit and were then injected into one-cell-stage zebrafish oocytes at 40 ng/μl together with 30 ng/μl *Tol2* transposase mRNA. GFP expression was inspected at 24-48 hpf. Only embryos having at least one GFP-positive OSN were counted as positive embryos.

Statistics. Statistical analysis was performed using R (version 3.5.1) and GraphPad Prism (version 7.0a). We calculated p values by the Wald test in DiffBind package for differential peak analysis in ATAC-seq data (Figure 2B, 2D), by one-way ANOVA and post-hoc Tukey's test for multiple group comparison (Figure 4E, 4F, 4H, 5E, 5F, 6D, 8C), by the LRT and Wald test in DESeq2 package for differentially expressed gene analysis in RNA-seq data (Figure 4C, 4D, 5C, 5D, Supplementary Figure 2, 5A, 5B, 5F, 6A, 6B), and by Fisher's exact test (Figure 7C, 7E). In all figures, p values or q values are denoted as * < 0.05, ** < 0.01, *** < 0.001.

Data availability. The data supporting the findings of this study are included within the article and its Supplemental files. Reagents are available from the corresponding author upon reasonable request.

Acknowledgements

We thank the Bob Datta lab at Harvard Medical School for kindly providing *Omp-ires-GFP* mouse line, the Marius Hoener lab at Roche Innovation Center Basel for sharing *Taar2-9* cluster knockout mouse line, the Stavros Lomvardas lab for sharing RNA-seq data of Ldb1 and Lhx2 knockout mice. We would like to thank Wanxin Zeng from the Xiajing Tong lab for help in imaging, Jason Buenrostro for advice on ATAC-seq, Matthew Lorincz and Julie Brind'Amour for advice on ULI-ChIP-seq, and Stavros Lomvardas for helpful discussion. We also want to thank members of the Li lab for critical reading of the manuscript. Computational analyses were performed on the Orchestra and O2 High-performance Computer Cluster at Harvard Medical School.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Images were taken in Core Facility of Basic Medical Sciences, Shanghai Jiao Tong University School of Medicine and the Molecular Imaging Core Facility (MICF) at School of Life Science and Technology, ShanghaiTech University. This work was supported by National Natural Science Foundation of China (31771154 and 31970933 to Q.L.), Shanghai Brain-Intelligence Project from the Science and Technology Commission of Shanghai Municipality (18JC1420302), Shanghai Pujiang Program (17PJ1405400 to Q.L.), Program for Young Scholars of Special Appointment at Shanghai Institutions of Higher Learning (QD2018017 to Q.L.), Innovative research team of high-level local universities in Shanghai, Fundamental Research Funds for the Central Universities (Shanghai Jiao Tong University, 17X100040037 to Q.L.), and an NIH Director's Pioneer Award (DP1 CA186693 to X.S.X.). L.T. was supported by an HHMI International Student Research Fellowship. Work at the Barnea lab was supported by grants number 1R01DC013561 and P30GM103410.

Author contributions

A.F., W.W., L.T., and Q.L. designed experiments, analyzed data, and wrote the manuscript; A.F., W.W., L.T., H.B., and C.Y. performed morphological, histological, RNA-seq, ATAC-seq, and animal experiments; C.T., X.H., and H.Y. generated enhancer knockout and transgenic mouse lines; Z.X. performed phylogenetic analysis of TAAR enhancers; C.R., K.J., and F.E. performed zebrafish embryo injection and imaging experiments; M.J., G.H., M.T., and G.B. made the antibodies against TAAR4, TAAR5, and TAAR6 prior to publication [26] and generated the *Taar5-ires-cre* and *Taar6-ires-cre* mice.

Competing interests

The authors declare no competing interests.

Figure legends

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1. Enrichment of TAAR OSNs. (A) Schematic illustration of the design strategy for *ires-Cre* knockin mice. DNA sequences encoding the *ires-Cre* alleles were inserted after the endogenous stop codon of *Taar5* or *Taar6* genes, along with the gene encoding neomycin (neo) resistance. The neo cassette flanked by flippase recognition target (FRT) sites was further excised by crossing to mice that express germline Flp recombinase. *Taar5-ires-Cre* or *Taar6-ires-Cre* mice were then crossed with Cre-dependent reporter lines to allow for fluorescent labelling of TAAR OSNs. (B) FACS-sorted reporter-positive (*Taar5*+ or *Taar6*+) and reporter-negative (control) cells were collected for RNA-seq experiments. The values of transcripts per million (TPM) extracted for the total *Taar* and *OR* genes were plotted. (C) The TPM values of all the 15 *Taar* genes were plotted in reporter-positive (*Taar5*+ or *Taar6*+) and reporter-negative (control) cells, showing that reporter-positive cells were composed of mixed TAAR OSN populations. (D) The canonical olfactory signaling molecules were expressed at similar level in reporter-positive (*Taar5*+ or *Taar6*+) and reporter-negative (control) cells, including *Gnal*, *Adcy3*, *Cnga2*, and *Ano2*. (E) H3K9me3 ChIP-seq results showed that both the *Taar* cluster and the *OR* cluster were decorated by H3K9me3 in *Taar5*+ OSNs. In contrast, the *Taar* cluster was devoid of H3K9me3 decoration in control cells that are mainly OR OSNs.

Figure 2. Identification of two putative TAAR enhancers. (A) Schematic illustration of the strategy to identify putative TAAR enhancers using ATAC-seq. (B) 6093 differential peaks with 3290 peaks enriched in OMP-positive OSNs and 2803 peaks enriched in TAAR OSNs were identified with criterion of $q < 0.05$. (C) ATAC-seq signals across the *Taar* cluster showed two putative TAAR enhancers that were enriched in TAAR OSNs. ATAC-seq signals were normalized to median peak values of each sample. Below the signal tracks, exons of the *Taar* genes were depicted. Arrows inside exons indicate direction of sense strand. Shaded regions indicate ATAC-seq peaks called by MACS2. (D) ATAC-seq peak intensities of two TAAR enhancers in OMP-positive OSNs and TAAR OSNs were normalized to median peak values and plotted. *** $q < 0.001$, by the Wald test. (E) ATAC-seq signals of Rhodes,

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

an OR enhancer, showed similar chromatin accessibility in OMP-positive OSNs and TAAR OSNs. ATAC-seq signals were normalized to median peak values of each sample. Below the signal tracks, exons of two surrounding *OR* genes were depicted. Arrows inside exons indicate direction of sense strand. Shaded regions indicate ATAC-seq peaks called by MACS2. (F) Correlation between normalized ATAC-seq peak intensities of all 63 OR enhancers in TAAR OSNs (Y-axis) and those in OMP-positive OSNs (X-axis). Pearson correlation coefficient $r = 0.89$, $p < 0.0001$.

Figure 3. Evolutionary conservation of the two TAAR enhancers. (A) Conservation of the *Taar* cluster ranging from *Taar1* to *Taar7b* in mouse with that in 8 Glires, 21 Euarchontoglires, 40 placental mammals, and 60 vertebrates using phyloP methods. Sites predicted to be conserved were assigned positive scores and shown in blue. Peaks represent the conserved regions. (B) Phylogenetic tree of representative species selected from Eutheria, Metatheria, and Prototheria order. Each branch length was calculated from divergence times obtained from TimeTree. MYA, million years ago. (C) Nucleotide percent identity plots of TAAR enhancer 1 (left) and TAAR enhancer 2 (right) sequences from different mammals compared to corresponding mouse sequences using VISTA. The arrows indicate the transcriptional orientations of *Taar* genes in mouse genome. Conservation between 50% and 100% are shown as peaks. Highly conserved regions (> 100 bp width and > 70% identity) in exons, intergenic regions, and TAAR enhancers are shown in blue, red, and cyan, respectively.

Figure 4. TAAR enhancer 1 deletion results in massive reduction of *Taar* gene expression. (A) Schematic illustration of TAAR enhancer 1 deletion using the CRISPR-Cas9 genome editing system. Arrows indicate primers used for genotype determination. (B) Representative PCR genotyping results to distinguish between wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 1 knockout mice using primers indicated in A. (C) RNA-seq analysis showed largely decreased expression of the *Taar* genes in homozygous TAAR enhancer 1 knockout mice. The

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

log2-fold change values for the *Taar* genes and 3 surrounding genes were plotted in the *Taar* cluster region. Differentially expressed genes with criteria of q < 0.05 were represented by filled circles (n = 3). (D) The log2-fold change values for the *Taar* genes and all of the functional *OR* genes were plotted according to their relative positions along the chromosome (*Taar* genes, blue; *OR* genes, red). Differentially expressed genes with criteria of q < 0.05 were represented by filled circles (n = 3). (E) Left, representative images of *Taar5*, *Taar6*, *Taar7s*, and *Olfr1507* expression in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 1 knockout mice using single color *in situ* hybridization. Scale bar = 25 μm. Right, the numbers of OSNs expressing *Taar* genes were quantified and percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted. The error bars represented mean ± s.e. ($n \geq 5$). (F) All of the *Taar7* family genes were detected using specific probes by *in situ* hybridization. Percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted. The error bars represented mean ± s.e. ($n \geq 5$). (G) Correlation between log2-fold change values of positive cell numbers by *in situ* hybridization (Y-axis) and gene expression by RNA-seq (X-axis) for the *Taar* genes. Pearson correlation coefficient $r = 0.94$, $p = 4.9 \times 10^{-6}$. (H) Top, representative confocal images of immunohistochemistry staining for TAAR5 and Olfr552 in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 1 knockout mice. Blue fluorescence represented DAPI counterstaining. Scale bar = 25 μm. Bottom, percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted. The error bars represented mean ± s.e. ($n \geq 5$). In E, F, and H, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, by one-way ANOVA and post-hoc Tukey's test.

Figure 5. TAAR enhancer 2 deletion also results in dramatic reduction of *Taar* gene expression. (A) Schematic illustration of TAAR enhancer 2 deletion using the CRISPR-Cas9 genome editing system. Arrows indicate primers used for genotype determination. (B) Representative PCR genotyping results to distinguish between wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 2 knockout mice

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

using primers indicated in A. (C) RNA-seq data showed largely decreased *Taar* gene expression in homozygous TAAR enhancer 2 knockout mice. The log2-fold change values for the *Taar* genes and 3 surrounding genes were plotted in the *Taar* cluster region. Differentially expressed genes with criteria of $q < 0.05$ were represented by filled circles ($n = 3$). (D) The log2-fold change values for the *Taar* genes and all of the functional *OR* genes were plotted according to their relative positions along the chromosome (*Taar* genes, blue; *OR* genes, red). Differentially expressed genes with criteria of $q < 0.05$ were represented by filled circles ($n = 3$). (E) Left, representative images of *Taar2*, *Taar5*, *Taar6*, and *Olfr1507* expression in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 2 knockout mice using single color *in situ* hybridization. Scale bar = 25 μm. Right, the numbers of OSNs expressing *Taar* genes were quantified and percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted. The error bars represented mean ± s.e. ($n \geq 5$). (F) All of the *Taar7* family genes were detected using specific probes by *in situ* hybridization. Percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted. The error bars represented mean ± s.e. ($n \geq 5$). In E and F, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, by one-way ANOVA and post-hoc Tukey's test.

Figure 6. Deletion of the two TAAR enhancers causes complete elimination of TAAR OSNs. (A) Schematic illustration of TAAR enhancer 1 & 2 double knockout strategy using the CRISPR-Cas9 genome editing system. Arrows indicate primers used for genotype determination. (B) Representative PCR genotyping results to distinguish between wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 1 & 2 double knockout mice using primers indicated in A. (C) Representative images of *Taar2*, *Taar5*, *Taar6*, *Taar7*, and *Olfr1507* expression in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 1 & 2 double knockout mice using single color *in situ* hybridization. Scale bar = 25 μm. (D) The numbers of OSNs expressing *Taar* genes were quantified and percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The error bars represented mean ± s.e. ($n \geq 5$). In D, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, by one-way ANOVA and post-hoc Tukey's test.

Figure 7. TAAR enhancer 1 possesses TAAR OSN-specific enhancer activity. (A) Schematic illustration of the design strategy for TAAR enhancer 1-GFP (left) and TAAR enhancer 2-tdTomato (right) transgene constructs. The boxes in dark blue represent Hsp68 minimal promoter. The numbers of GFP-positive and tdTomato-positive independent founders among the total analyzed lines were also indicated. (B) Top, confocal images of two-color *in situ* hybridization using probes for *GFP* (green) and all of the *Taar* genes (red) in a TAAR enhancer 1-GFP transgenic mouse. Bottom, confocal images of *in situ* hybridization of mixed probes of *OR* genes (red) and the following immunohistochemistry by the GFP antibody (green) in a TAAR enhancer 1-GFP transgenic mouse. Scale bar = 25 μm. (C) Top, bar plots showing the percentages of GFP-positive cells that were co-labeled with mixed *Taar* probes, mixed *OR* probes, or *OR* degenerate probe (1083 out of 1284 cells for mixed *Taar* probe, 25 out of 616 cells for mixed *OR* probe, 1 out of 351 cells for *OR* degenerate probe). Bottom, bar plots showing the percentages of TAAR or OR OSNs that co-expressed GFP using mixed *Taar* probes, mixed *OR* probes, or *OR* degenerate probe (564 out of 3920 cells for mixed *Taar* probe, 17 out of 898 cells for mixed *OR* probe, 1 out of 203 cells for *OR* degenerate probe). *** $p < 0.001$ by Fisher's exact test. (D) Confocal images of *in situ* hybridization of mixed probes of *Taar* genes or *OR* genes (green) and the following immunohistochemistry by the tdTomato antibody (red) in a TAAR enhancer 2-tdTomato transgenic mouse. Scale bar = 25 μm. (E) Top, bar plots showing the percentages of tdTomato-positive cells that were co-labeled with mixed *Taar* probes, mixed *OR* probes, or *OR* degenerate probe (817 out of 1116 cells for mixed *Taar* probe, 9 out of 896 cells for mixed *OR* probe, 2 out of 657 cells for *OR* degenerate probe). Bottom, bar plots showing the percentages of TAAR or OR OSNs that co-expressed tdTomato using mixed *Taar* probes, mixed *OR* probes, or *OR* degenerate probe (817 out of 5863 cells for mixed *Taar* probe, 9 out of 4726 cells for mixed *OR* probe, 2 out of 3713 cells for *OR* degenerate probe). *** $p < 0.001$ by

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fisher's exact test. (F) Representative confocal images showing cells with GFP and tdTomato signals in TAAR enhancer 1-GFP; TAAR enhancer 2-tdTomato transgenic mice. Right, the numbers of cells expressing each reporter alone (GFP, tdTomato) or together were counted. Scale bar = 25 μm. (G) Whole-mount fluorescent images of the dorsal olfactory bulb in TAAR enhancer 1-GFP; TAAR enhancer 2-tdTomato transgenic mice. The area in the white box is shown on the right. Scale bar = 200 μm.

Figure 8. TAAR enhancer 1 operates in *cis*. (A) Schematic illustration of four different mouse genotypes by crossing heterozygous TAAR enhancer 1 knockout mice with heterozygous *Taar2-9* cluster knockout mice. According to the assumption that TAAR enhancer 1 functions in *cis*, in *trans* or in *cis* plus *trans*, the hypothetical percentage of positive cell numbers in different genotypes were indicated by normalizing to wild type (genotype 1) mice (set as 100%). (B) Representative images of *Taar2*, *Taar3*, *Taar5*, and *Taar6* expression in four different genotypes indicated in A using single color *in situ* hybridization. Scale bar = 25 μm. (C) The numbers of OSNs expressing *Taar2*, *Taar3*, *Taar5*, and *Taar6* were quantified and percentage of positive cell numbers normalized to wild type was plotted. The error bars represented mean ± s.e. ($n \geq 5$). ** $p < 0.01$, *** $p < 0.001$, by one-way ANOVA and post-hoc Tukey's test.

Supplementary Figure 1. Representative FACS density plots. (A) Representative plots with gates showing sorted ZsGreen-positive and ZsGreen-negative cells in *Taar5-ires-Cre*; *lox-ZsGreen* mice. (B) The classical olfactory chaperon molecules (*Rtp1* and *Rtp2*) were expressed at similar level in reporter-positive (*Taar5*+ or *Taar6*+) and reporter-negative (control) cells.

Supplementary Figure 2. Similar chromatin accessibility of canonical signaling molecules between TAAR OSNs and OMP-positive OSNs. ATAC-seq signals plotted across different genomic regions containing various olfactory signaling molecules. ATAC-seq signals were normalized to median peak values of each sample. Below the signal tracks, exons of genes were depicted. Arrows inside exons indicate

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

direction of sense strand. The NM_010307 isoform of *Gnal* gene is possibly the major variant expressed in the MOE based on the fact that ATAC-seq peaks are found in its promoter region but not the NM_177137 isoform. This is consistent with our RNA-seq data.

Supplementary Figure 3. The *Taar* genes were downregulated in mature OSNs when Lhx2 or Ldb1 is deleted. Published RNA-seq data of 5 mature OSN samples from control mice, 3 mature OSN samples from Lhx2 knockout mice, and 4 mature OSN samples from Ldb1 knockout mice were retrieved from the GEO (GSE112153) [17] and analyzed to examine changes of the *Taar* gene expression. Normalized counts of the 14 olfactory *Taar* genes were depicted in control mice, Lhx2 knockout mice, and Ldb1 knockout mice.

Supplementary Figure 4. Conservation analysis of the two TAAR enhancers. (A) Summary of 8 Glires, 21 Euarchontoglires, 40 placental mammals, and 60 vertebrates included in the conservation analysis in Figure 3A. The subset information and evolutionary relationships were obtained from UCSC. (B) VISTA nucleotide percent identity plots of the mouse *Taar* cluster ranging from *Taar1* to *Taar7a* compared with that in orangutan, gibbon, dolphin, and sperm whale. (C) VISTA nucleotide percent identity plots of mouse TAAR enhancer 2 and two adjacent *Taar* genes compared with that in in cow, sheep, and horse. The arrows indicate the transcriptional orientations of *Taar* genes in mouse genome. Conservation between 50% and 100% are shown as peaks. Highly conserved regions (> 100 bp width and > 70% identity) in exons, intergenic regions, and TAAR enhancers are shown in blue, red, and cyan, respectively.

Supplementary Figure 5. Downregulation of specific *Taar* genes by TAAR enhancer 1 deletion. (A) The 14 DEGs were identified using the criteria of $q < 0.05$ and fold change > 1.5-fold in RNA-seq data of homozygous TAAR enhancer 1 knockout mice and wild type mice. The 9 *Taar* genes that showed significant changes

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

are indicated by blue dots. The 5 other DEGs are indicated by red dots. And other *Taar* genes are indicated by blue circles. (B) Normalized counts of the 14 olfactory *Taar* genes from RNA-seq data were depicted in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 1 knockout mice. (C) Normalized counts of the *OR* genes from RNA-seq data were depicted in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 1 knockout mice. (D) Top, representative confocal images of immunohistochemistry staining for TAAR6 and Olfr1507 in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 1 knockout mice. Blue fluorescence represented DAPI counterstaining. Scale bar = 25 μm. Bottom, percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted. The error bars represented mean ± s.e. ($n \geq 5$). (E) Left, representative confocal images of immunohistochemistry staining for Caspase-3 in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 1 knockout mice. Blue fluorescence represented DAPI counterstaining. Scale bar = 25 μm. Right, percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted. The error bars represented mean ± s.e. ($n \geq 5$).

Supplementary Figure 6. Downregulation of specific *Taar* genes by TAAR enhancer 2 deletion. (A) The 6 DEGs were identified using the criteria of $q < 0.05$ and fold change > 1.5-fold in RNA-seq data of homozygous TAAR enhancer 2 knockout mice and wild type mice. The 5 *Taar* genes that showed significant changes are indicated by blue dots. Another DEG is indicated by red dots. All of the other *Taar* genes are indicated by blue circles. (B) Normalized counts of the 14 olfactory *Taar* genes from RNA-seq data were depicted in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 2 knockout mice. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, by the LRT and Wald test in DESeq2 package. (C) Normalized counts of the *OR* genes from RNA-seq data were depicted in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 2 knockout mice. (D) Correlation between log2-fold change values of positive cell numbers by *in situ* hybridization (Y-axis) and gene

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expression by RNA-seq (X-axis) for the *Taar* genes. Pearson correlation coefficient r = 0.88, p = 0.0001. (E) Left, representative images of immunohistochemistry staining for TAAR5 and Olfr552 in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 2 knockout mice. Blue fluorescence represented DAPI counterstaining. Scale bar = 25 μm. Right, percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted. The error bars represented mean ± s.e. ($n \geq 5$). (F) Left, representative images of immunohistochemistry staining for Caspase-3 in wild type (+/+), heterozygous (+/-), and homozygous (-/-) TAAR enhancer 2 knockout mice. Blue fluorescence represented DAPI counterstaining. Scale bar = 25 μm. Right, percentage of positive cell numbers in heterozygous (+/-) or homozygous (-/-) mice compared to wild type (+/+) was plotted. The error bars represented mean ± s.e. ($n \geq 5$).

Supplementary Figure 7. Functional analysis of TAAR enhancers in zebrafish. (A) Schematic depiction of the E1b reporter expression construct used in zebrafish embryo injection. Various enhancer candidates were amplified from mouse genomic DNA and fused to the E1b minimal promoter followed by the GFP sequence. (B) Representative bright field (left) and fluorescent images of injected zebrafish larvae head at 4 dpf (days post-fertilization). (C) Percentage of injected zebrafish embryos with GFP-positive OSNs at 24-48 hpf (hours post-fertilization) for the TAAR enhancers or the Sifnos OR enhancer that served as a positive control.

Supplementary Figure 8. Validation of *Taar2-9* cluster knockout mice. (A) Normalized counts of the 14 olfactory *Taar* genes and the *OR* genes from RNA-seq data were depicted in wild type (+/+), heterozygous (+/-), and homozygous (-/-) *Taar2-9* cluster knockout mice. Expression of the olfactory *Taar* genes was abolished, while expression of the *OR* genes was not significantly changed. (B) Representative images of *Taar4*, *Taar6*, and *Olfr77* expression in wild type (+/+), heterozygous (+/-), and homozygous (-/-) *Taar2-9* cluster knockout mice using single color *in situ* hybridization. Scale bar = 25 μm. (C) Representative confocal images of immunohistochemistry

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

staining for TAAR4 in wild type (+/+), heterozygous (+/-), and homozygous (-/-) *Taar2-9* cluster knockout mice. Scale bar = 25 μm.

Supplementary Table 1. The information of extracted genome sequences from various mammalian species was listed.

References

1. Buck, L. & Axel, R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**, 175-187 (1991).
2. Liberles, S.D. & Buck, L.B. A second class of chemosensory receptors in the olfactory epithelium. *Nature* **442**, 645-650 (2006).
3. Greer, P.L. *et al.* A Family of non-GPCR Chemosensors Defines an Alternative Logic for Mammalian Olfaction. *Cell* **165**, 1734-1748 (2016).
4. Sun, L. *et al.* Guanylyl cyclase-D in the olfactory CO2 neurons is activated by bicarbonate. *Proc Natl Acad Sci U S A* **106**, 2041-2046 (2009).
5. Leinders-Zufall, T. *et al.* Contribution of the receptor guanylyl cyclase GC-D to chemosensory function in the olfactory epithelium. *Proc Natl Acad Sci U S A* **104**, 14507-14512 (2007).
6. Wang, F., Nemes, A., Mendelsohn, M. & Axel, R. Odorant receptors govern the formation of a precise topographic map. *Cell* **93**, 47-60 (1998).
7. Nakashima, A. *et al.* Agonist-Independent GPCR Activity Regulates Anterior-Posterior Targeting of Olfactory Sensory Neurons. *Cell* **154**, 1314-1325 (2013).
8. Barnea, G. *et al.* Odorant receptors on axon termini in the brain. *Science* **304**, 1468 (2004).
9. Magklara, A. *et al.* An epigenetic signature for monoallelic olfactory receptor expression. *Cell* **145**, 555-570 (2011).
10. Lyons, D.B. *et al.* An epigenetic trap stabilizes singular olfactory receptor expression. *Cell* **154**, 325-336 (2013).
11. Dalton, Ryan P., Lyons, David B. & Lomvardas, S. Co-Opting the Unfolded Protein Response to Elicit Olfactory Receptor Feedback. *Cell* **155**, 321-332 (2013).
12. Serizawa, S. *et al.* Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. *Science* **302**, 2088-2094 (2003).
13. Lewcock, J.W. & Reed, R.R. A feedback mechanism regulates monoallelic odorant receptor expression. *Proc Natl Acad Sci U S A* **101**, 1069-1074 (2004).
14. Shykind, B.M. *et al.* Gene switching and the stability of odorant receptor gene choice. *Cell* **117**, 801-815 (2004).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

15. Markenscoff-Papadimitriou, E. *et al.* Enhancer interaction networks as a means for singular olfactory receptor expression. *Cell* **159**, 543-557 (2014).
16. Monahan, K. *et al.* Cooperative interactions enable singular olfactory receptor expression in mouse olfactory neurons. *eLife* **6** (2017).
17. Monahan, K., Horta, A. & Lomvardas, S. LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. *Nature* **565**, 448-453 (2019).
18. Tan, L., Xing, D., Daley, N. & Xie, X.S. Three-dimensional genome structures of single sensory neurons in mouse visual and olfactory systems. *Nat Struct Mol Biol* **26**, 297-307 (2019).
19. Khan, M., Vaes, E. & Mombaerts, P. Regulation of the probability of mouse odorant receptor gene choice. *Cell* **147**, 907-921 (2011).
20. Nishizumi, H., Kumasaka, K., Inoue, N., Nakashima, A. & Sakano, H. Deletion of the core-H region in mice abolishes the expression of three proximal odorant receptor genes in cis. *Proc Natl Acad Sci U S A* **104**, 20067-20072 (2007).
21. Cichy, A., Shah, A., Dewan, A., Kaye, S. & Bozza, T. Genetic Depletion of Class I Odorant Receptors Impacts Perception of Carboxylic Acids. *Current Biology* (2019).
22. Iwata, T. *et al.* A long-range cis-regulatory element for class I odorant receptor genes. *Nat Commun* **8**, 885 (2017).
23. Hussain, A., Saraiva, L.R. & Korsching, S.I. Positive Darwinian selection and the birth of an olfactory receptor clade in teleosts. *Proc Natl Acad Sci U S A* **106**, 4313-4318 (2009).
24. Lindemann, L. *et al.* Trace amine-associated receptors form structurally and functionally distinct subfamilies of novel G protein-coupled receptors. *Genomics* **85**, 372-385 (2005).
25. Pacifico, R., Dewan, A., Cawley, D., Guo, C. & Bozza, T. An Olfactory Subsystem that Mediates High-Sensitivity Detection of Volatile Amines. *Cell Rep* **2**, 76-88 (2012).
26. Johnson, M.A. *et al.* Neurons expressing trace amine-associated receptors project to discrete glomeruli and constitute an olfactory subsystem. *Proc Natl Acad Sci U S A* **109**, 13410-13415 (2012).
27. Hussain, A. *et al.* High-affinity olfactory receptor for the death-associated odor cadaverine. *Proc Natl Acad Sci U S A* **110**, 19579-19584 (2013).
28. Li, Q. *et al.* Synchronous evolution of an odor biosynthesis pathway and behavioral response. *Curr Biol* **23**, 11-20 (2013).
29. Li, Q. & Liberles, S.D. Aversion and attraction through olfaction. *Curr Biol* **25**, R120-129 (2015).
30. Ferrero, D.M. *et al.* Detection and avoidance of a carnivore odor by prey. *Proc Natl Acad Sci U S A* (2011).
31. Dewan, A., Pacifico, R., Zhan, R., Rinberg, D. & Bozza, T. Non-redundant coding of aversive odours in the main olfactory pathway. *Nature* (2013).
32. Li, Q. *et al.* Non-classical amine recognition evolved in a large clade of olfactory receptors. *eLife* **4** (2015).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

33. Saraiva, L.R. *et al.* Combinatorial effects of odorants on mouse behavior. *Proc Natl Acad Sci U S A* (2016).
34. Liberles, S.D. Trace amine-associated receptors: ligands, neural circuits, and behaviors. *Curr Opin Neurobiol* **34C**, 1-7 (2015).
35. Xu, Z. & Li, Q. TAAR Agonists. *Cell Mol Neurobiol* **40**, 257-272 (2020).
36. Kim, D.G., Kang, H.M., Jang, S.K. & Shin, H.S. Construction of a bifunctional mRNA in the mouse by using the internal ribosomal entry site of the encephalomyocarditis virus. *Mol Cell Biol* **12**, 3636-3643 (1992).
37. Krashes, M.J. *et al.* An excitatory paraventricular nucleus to AgRP neuron circuit that drives hunger. *Nature* **507**, 238-242 (2014).
38. Borowsky, B. *et al.* Trace amines: identification of a family of mammalian G protein-coupled receptors. *Proc Natl Acad Sci U S A* **98**, 8966-8971 (2001).
39. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. & Greenleaf, W.J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213-1218 (2013).
40. Harmeier, A. *et al.* How Female Mice Attract Males: A Urinary Volatile Amine Activates a Trace Amine-Associated Receptor That Induces Male Sexual Interest. *Front Pharmacol* **9**, 924 (2018).
41. Lomvardas, S. & Maniatis, T. Histone and DNA Modifications as Regulators of Neuronal Development and Function. *Cold Spring Harb Perspect Biol* **8** (2016).
42. Clowney, E.J. *et al.* Nuclear Aggregation of Olfactory Receptor Genes Governs Their Monogenic Expression. *Cell* **151**, 724-737 (2012).
43. Yoon, K.H. *et al.* Olfactory receptor genes expressed in distinct lineages are sequestered in different nuclear compartments. *Proc Natl Acad Sci U S A* (2015).
44. Tan, L., Xing, D., Chang, C.H., Li, H. & Xie, X.S. Three-dimensional genome structures of single diploid human cells. *Science* **361**, 924-928 (2018).
45. Gogos, J.A., Osborne, J., Nemes, A., Mendelsohn, M. & Axel, R. Genetic ablation and restoration of the olfactory topographic map. *Cell* **103**, 609-620 (2000).
46. Brind'Amour, J. *et al.* An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat Commun* **6**, 6033 (2015).
47. Buenrostro, J.D., Wu, B., Chang, H.Y. & Greenleaf, W.J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21 29 21-29 (2015).
48. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**, 171-181 (2014).
49. Tan, L., Li, Q. & Xie, X.S. Olfactory sensory neurons transiently express multiple olfactory receptors during development. *Mol Syst Biol* **11**, 844 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

50. Kumar, S., Stecher, G., Suleski, M. & Hedges, S.B. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol* **34**, 1812-1819 (2017).
51. Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M. & Dubchak, I. VISTA: computational tools for comparative genomics. *Nucleic Acids Res* **32**, W273-279 (2004).
52. Li, Q. *et al.* A systematic approach to identify functional motifs within vertebrate developmental enhancers. *Dev Biol* **337**, 484-495 (2010).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

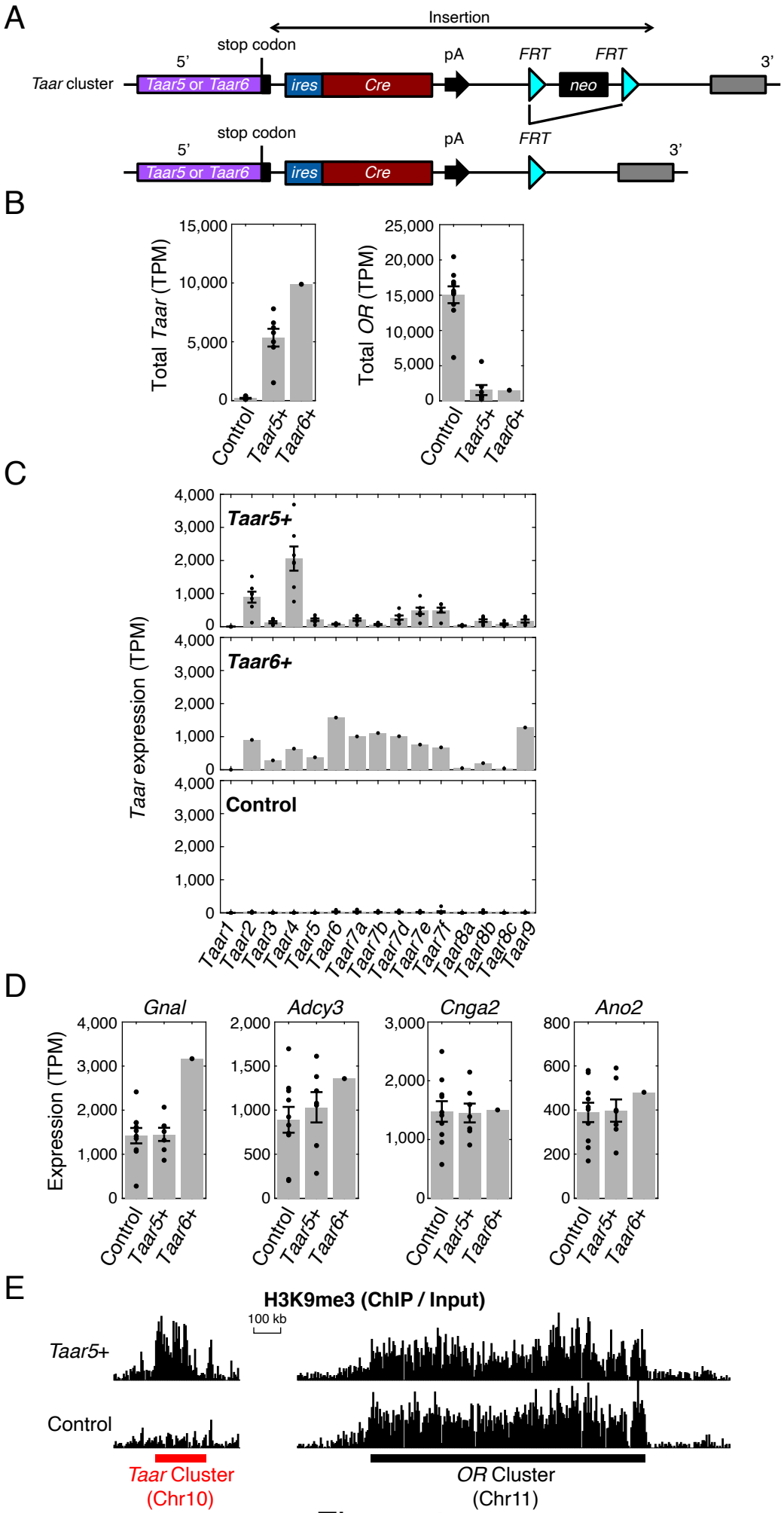


Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

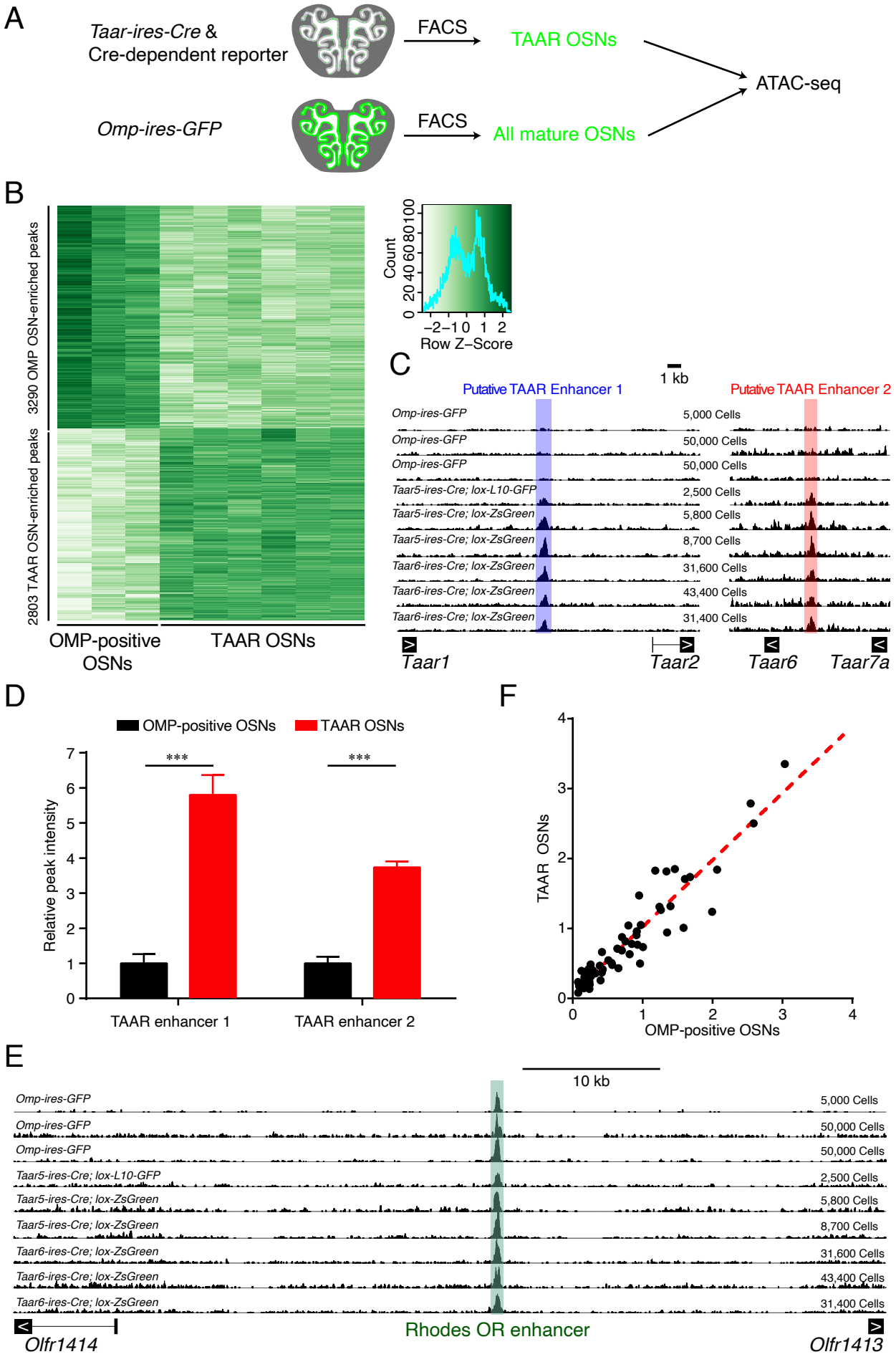


Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

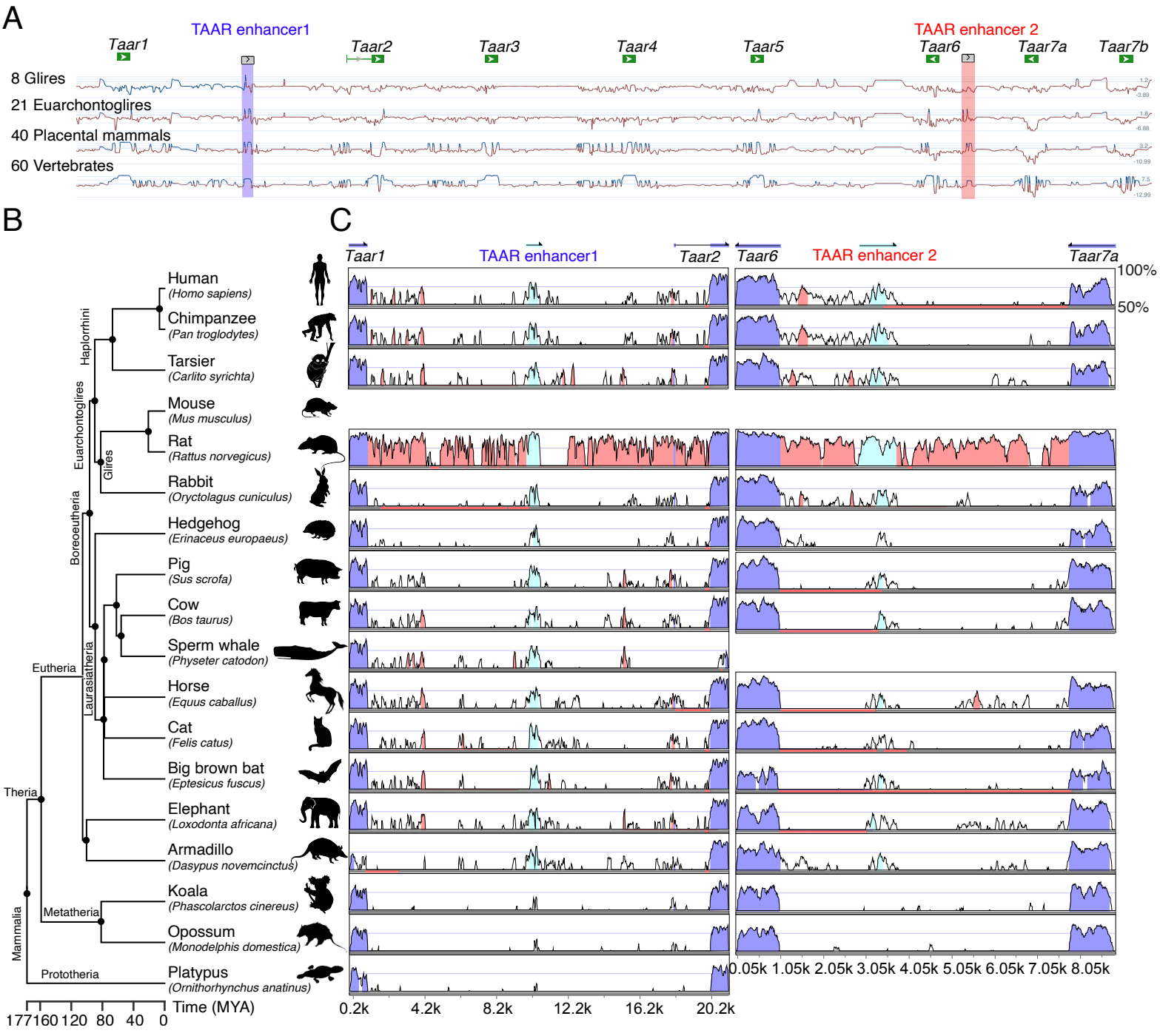


Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

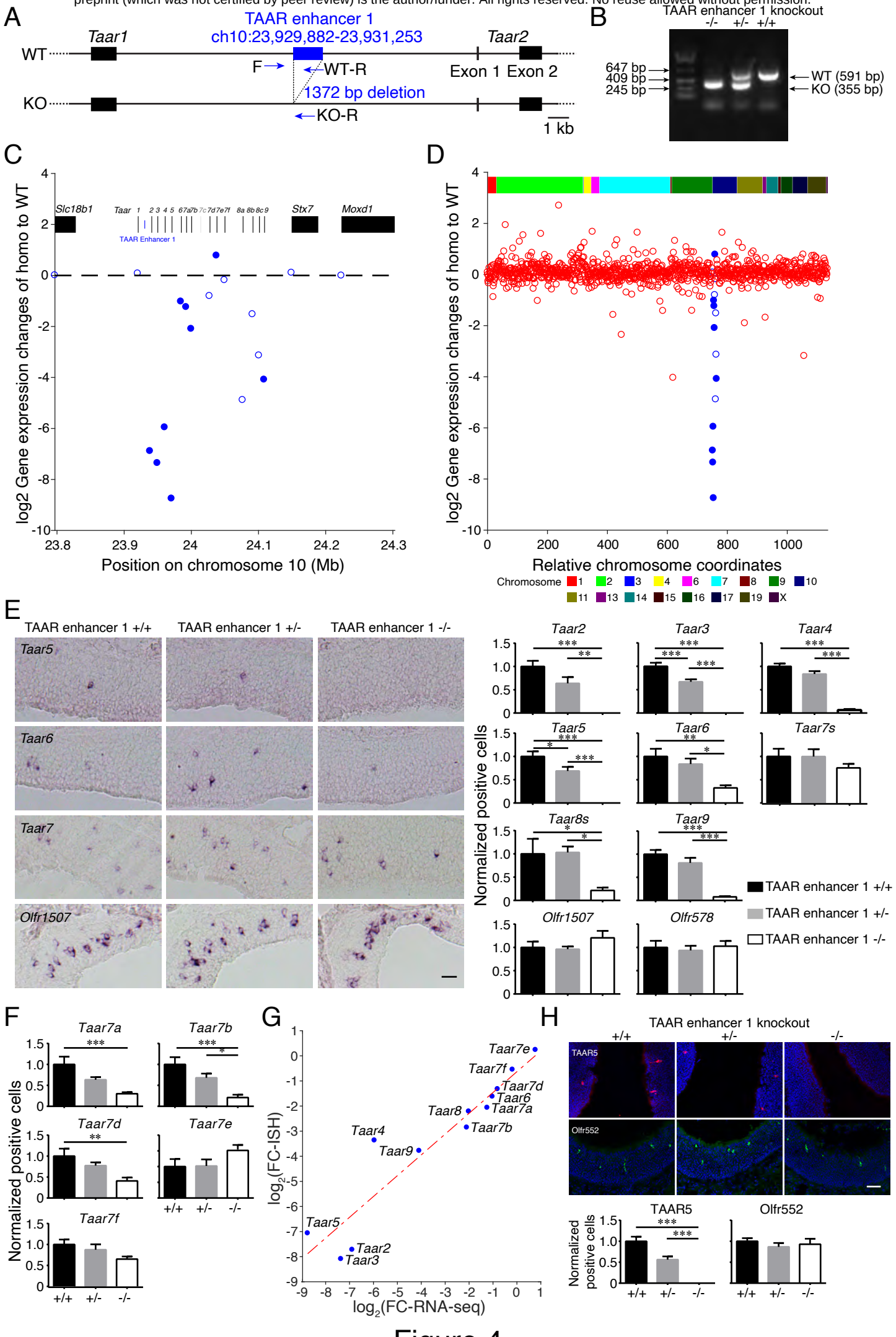
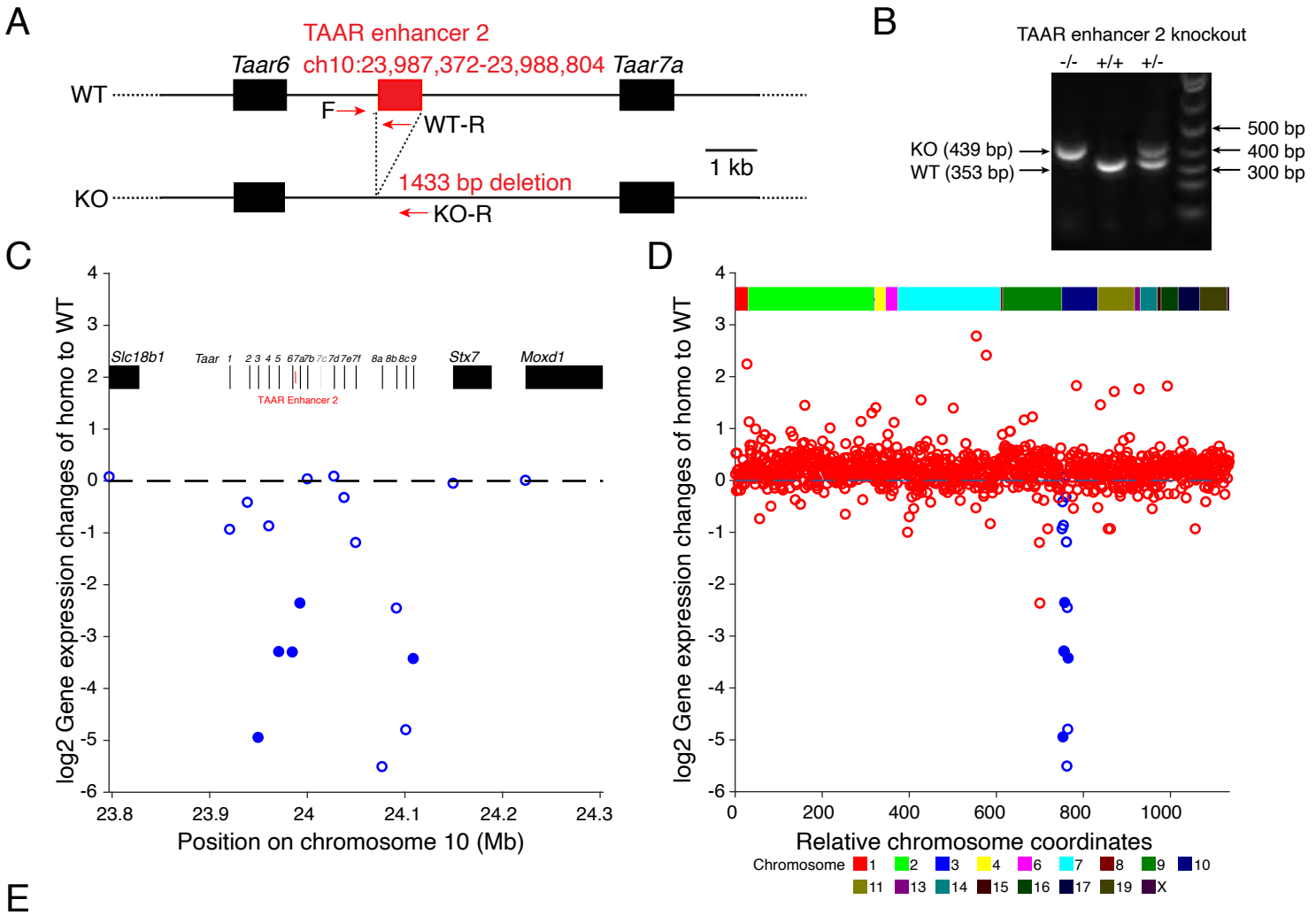
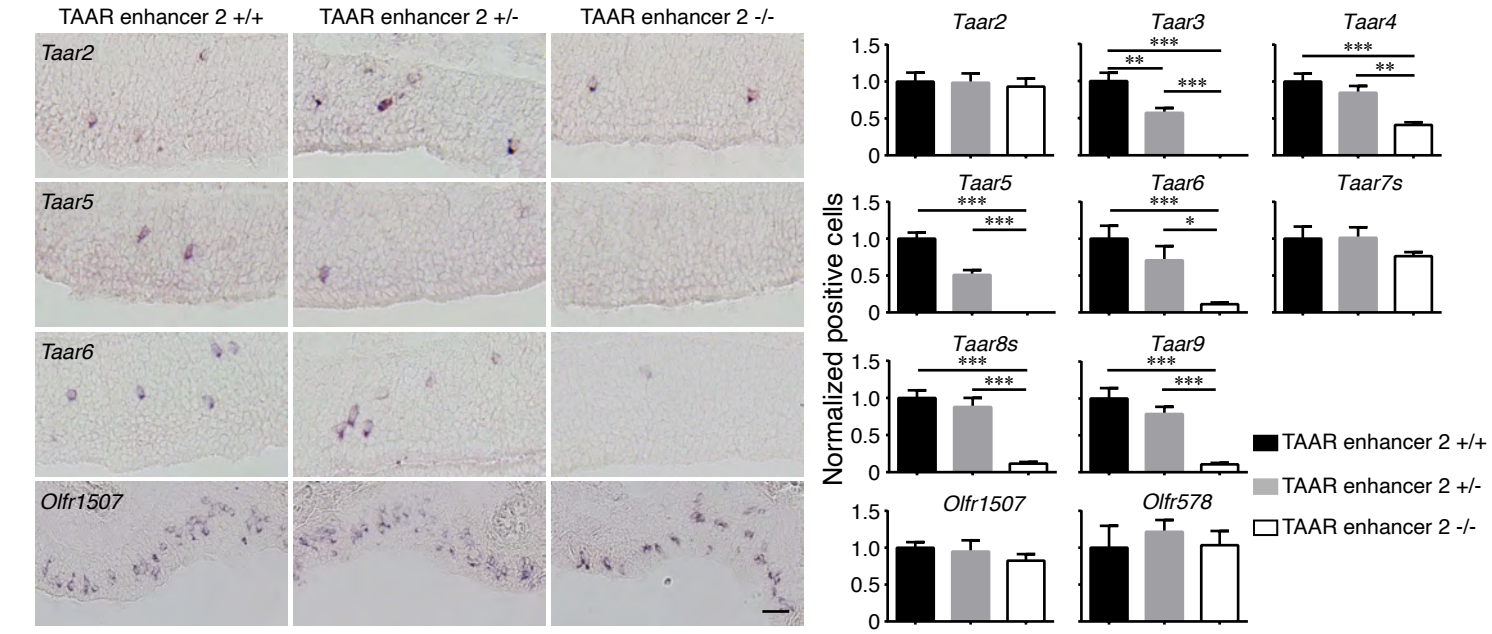


Figure 4

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.





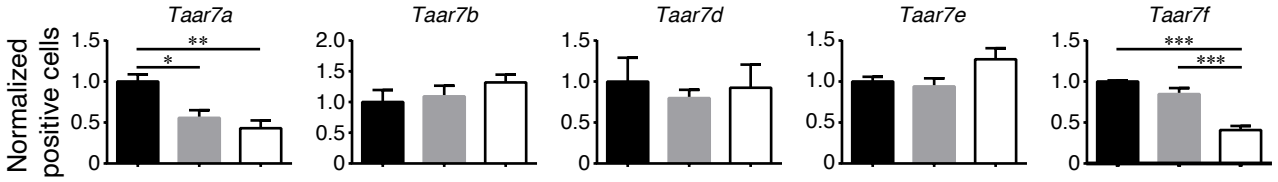


Figure 5

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

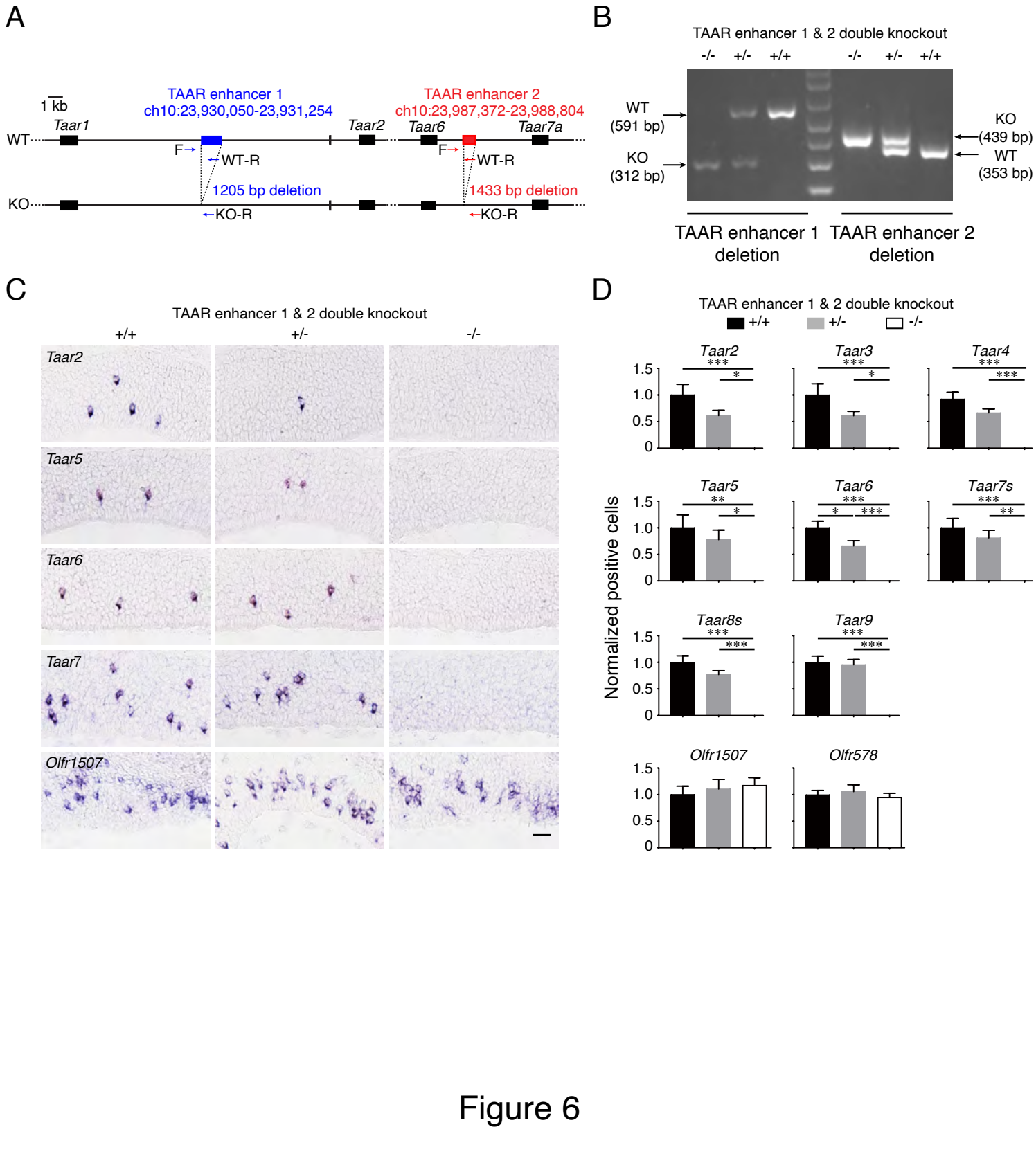


Figure 6

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

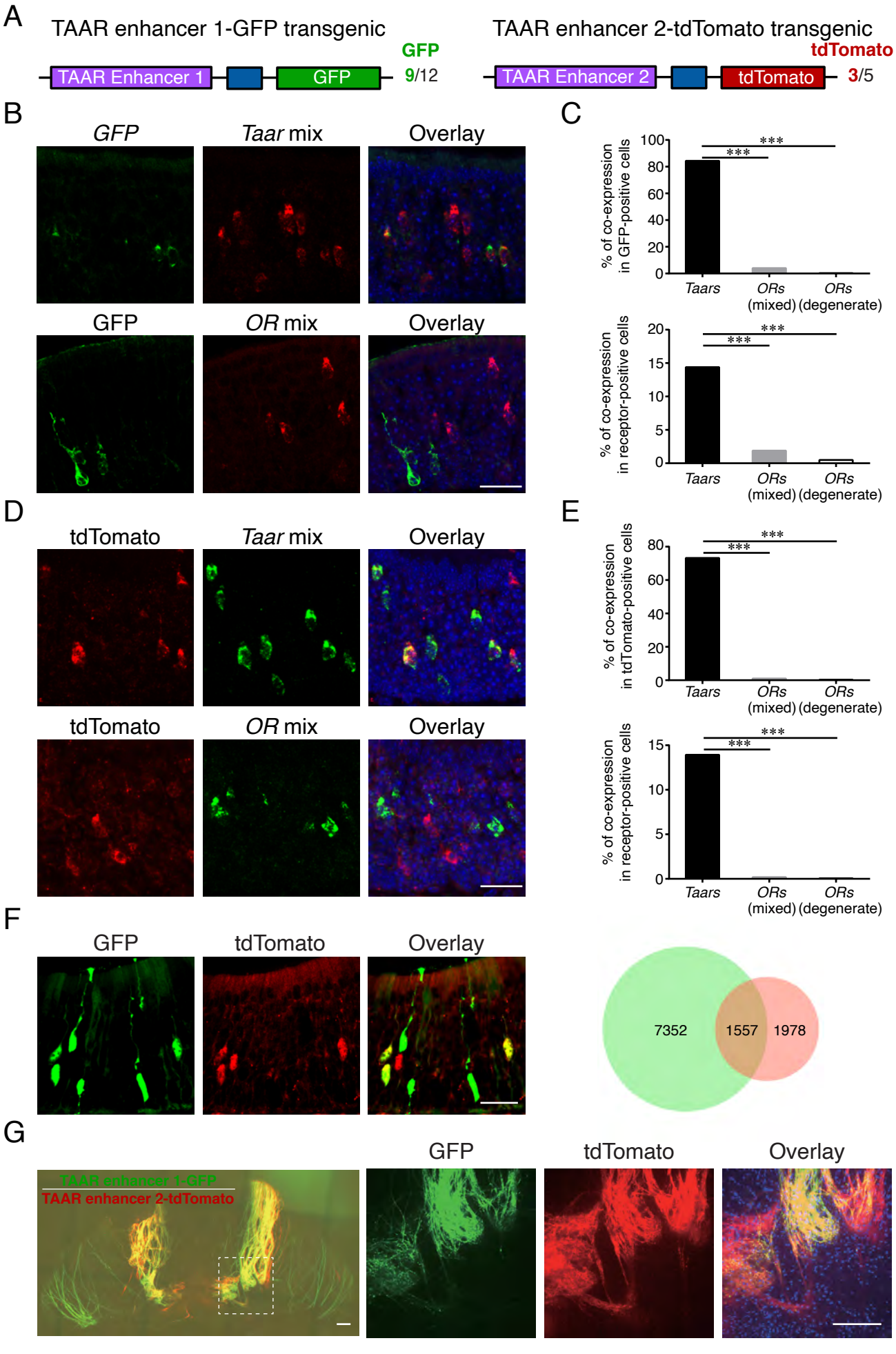


Figure 7

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

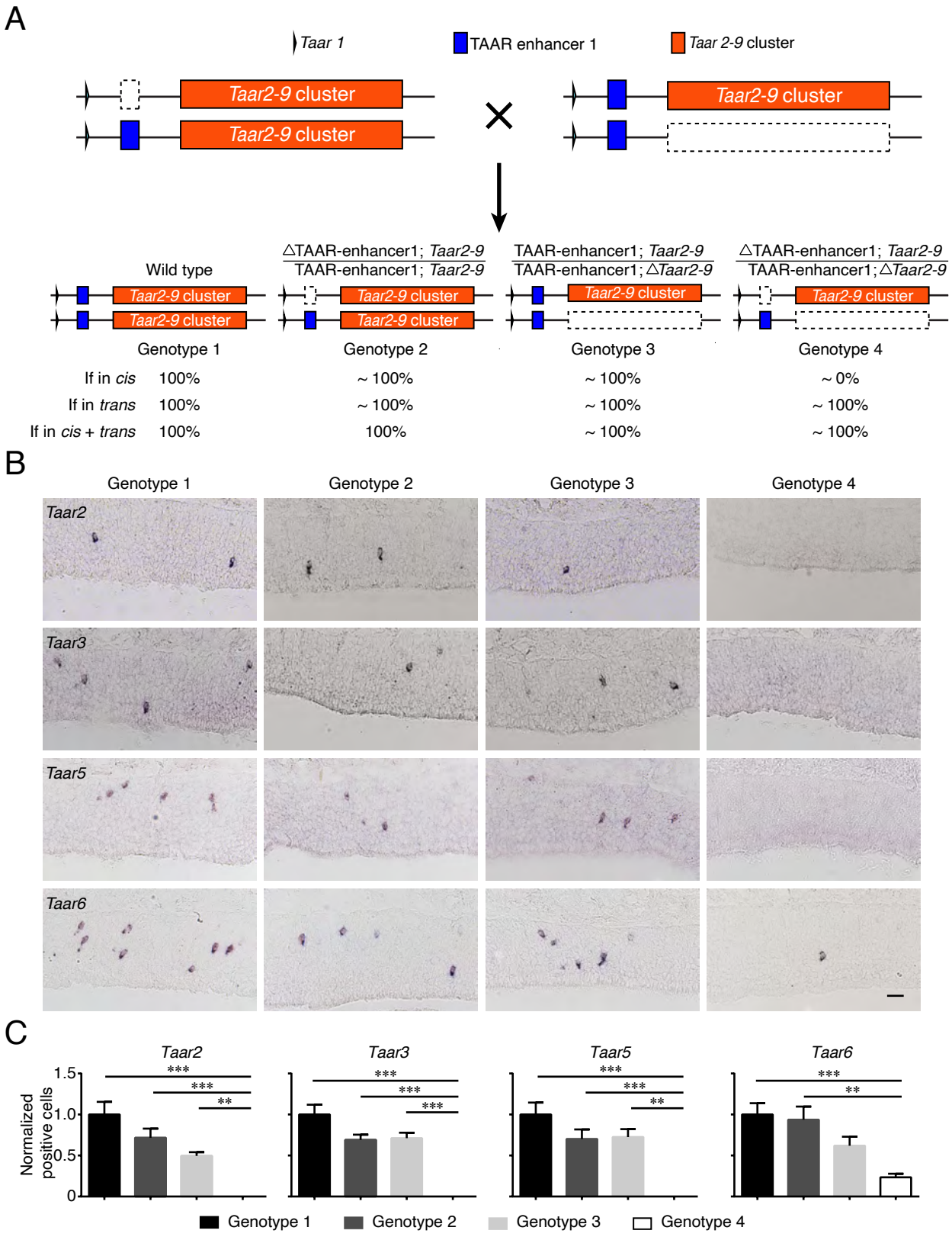
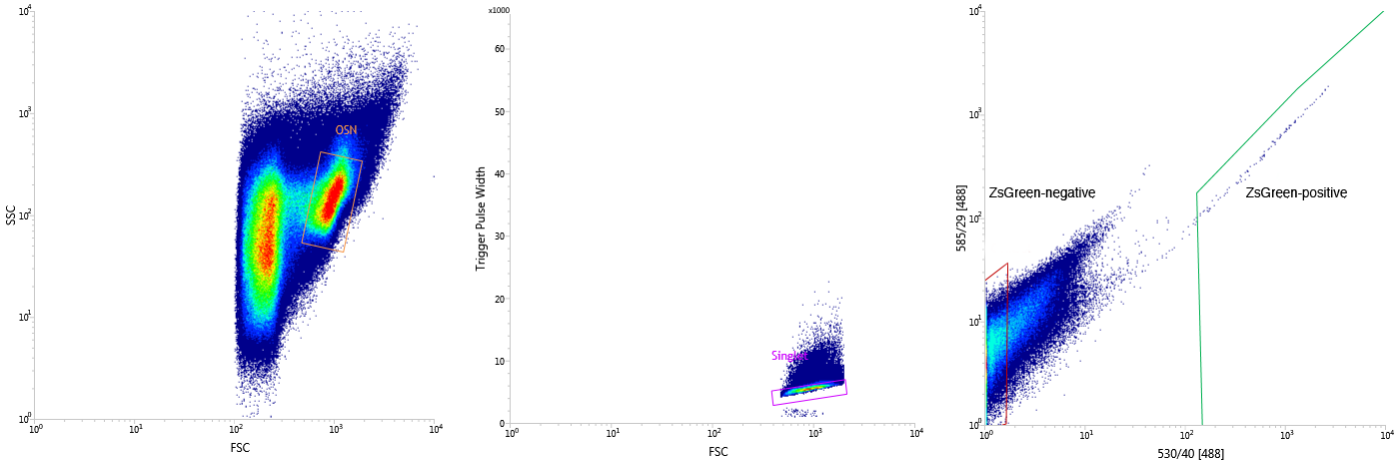


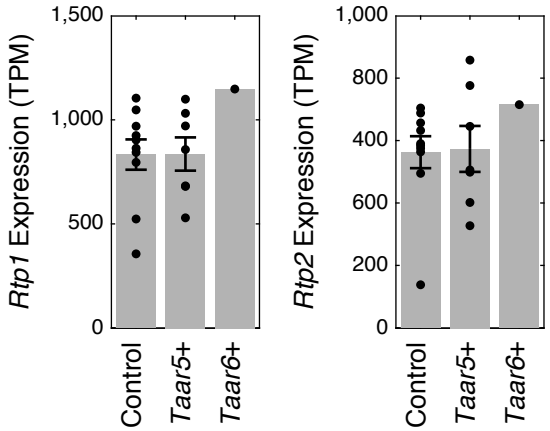
Figure 8

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A

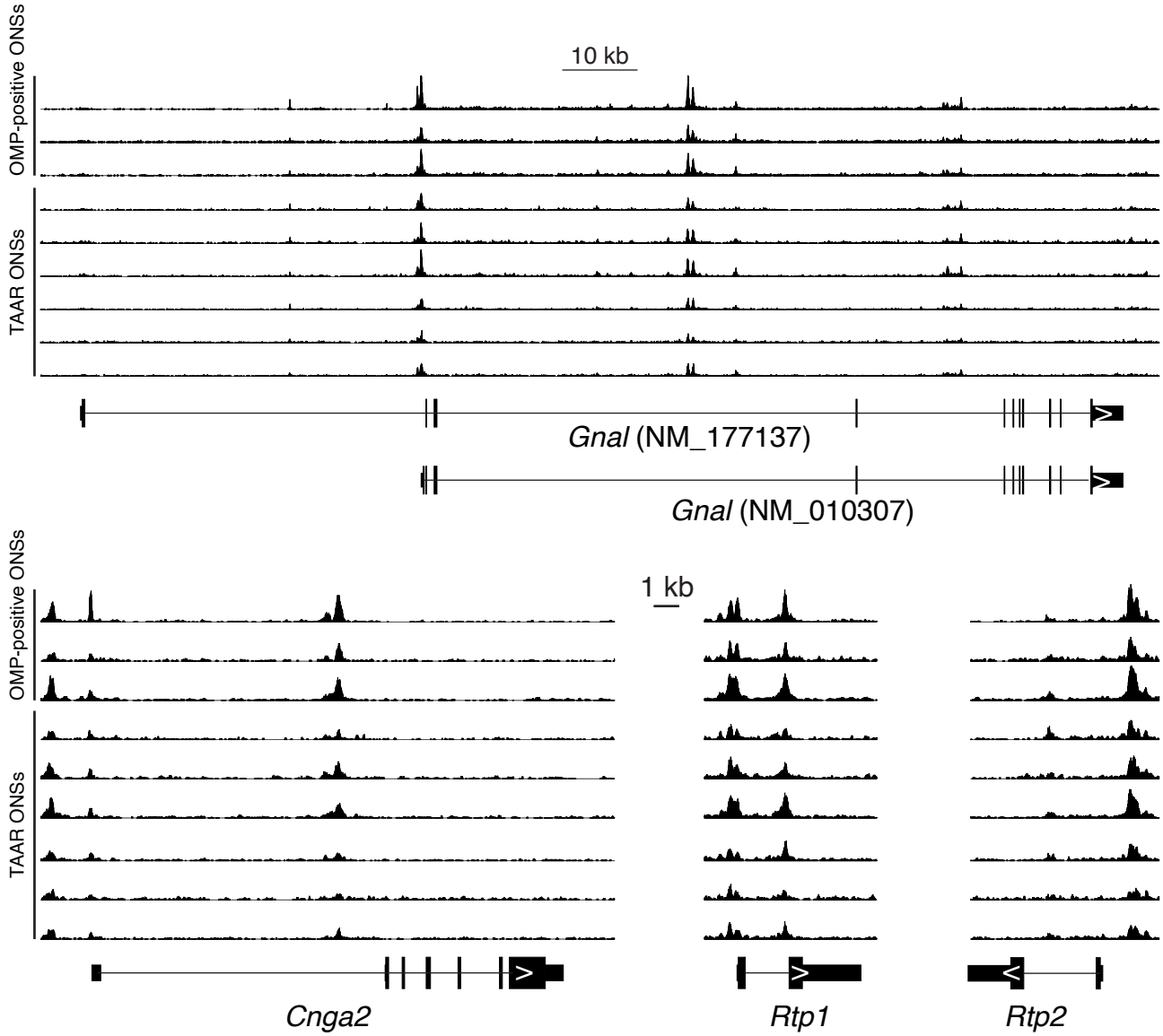


B



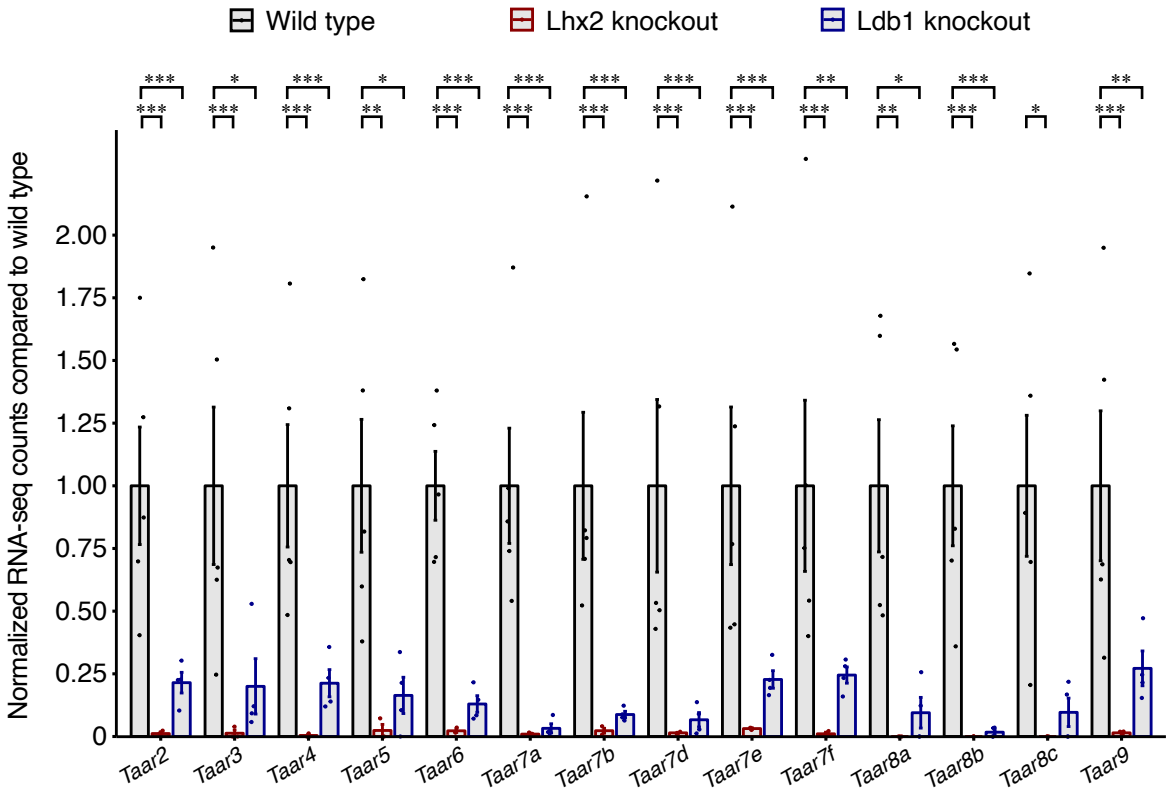
Supplementary figure 1

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



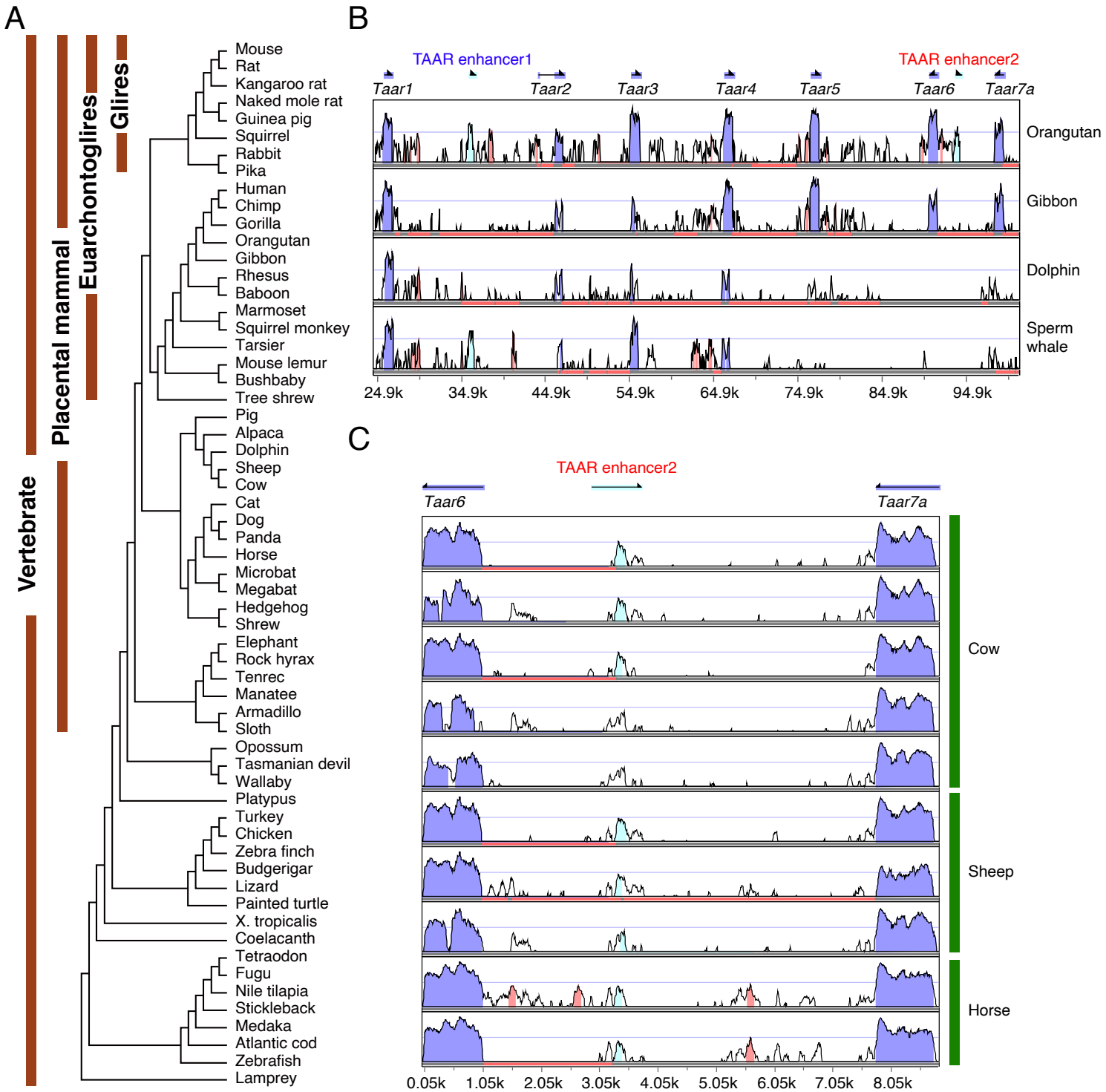
Supplementary figure 2

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



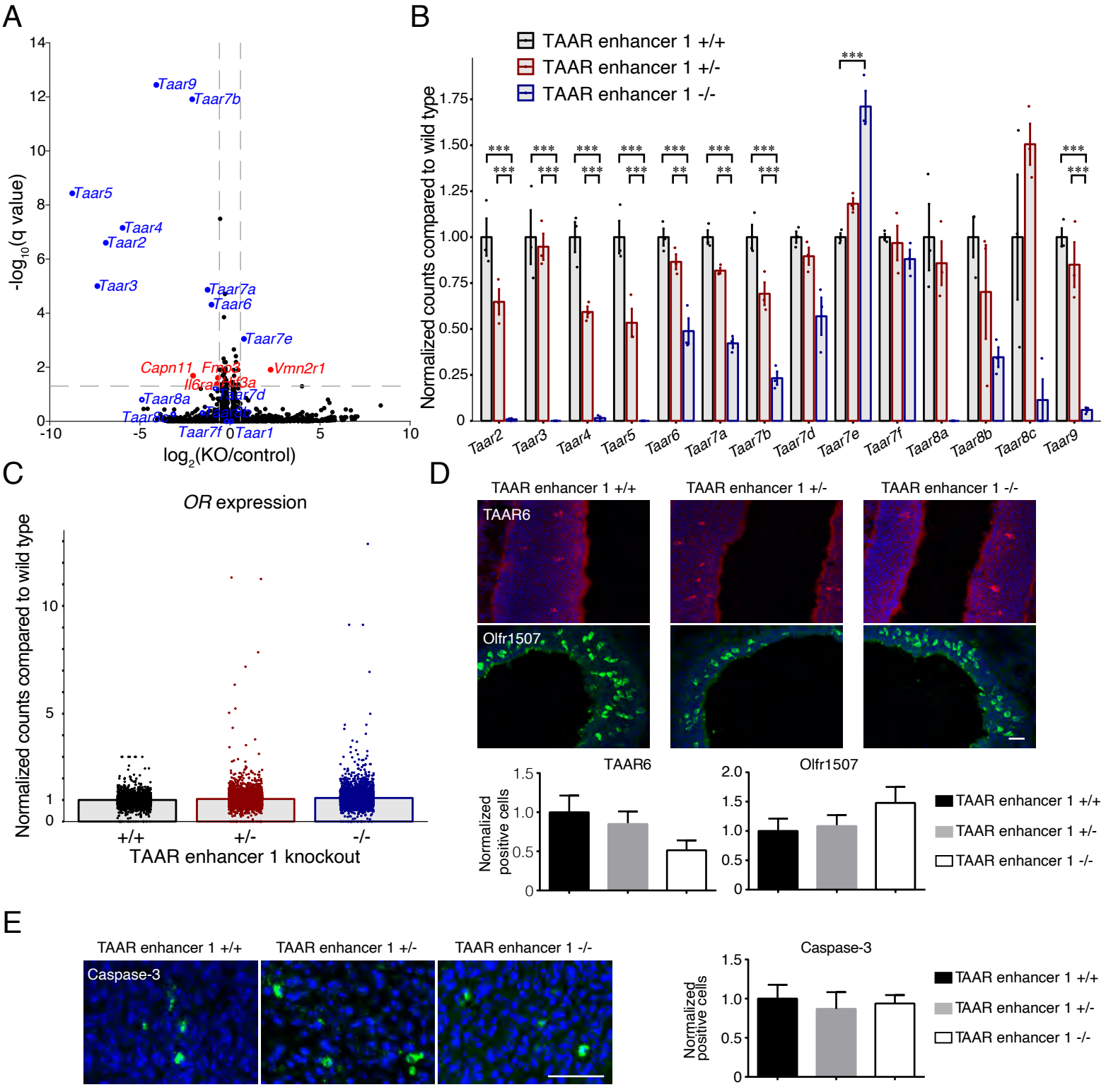
Supplementary figure 3

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



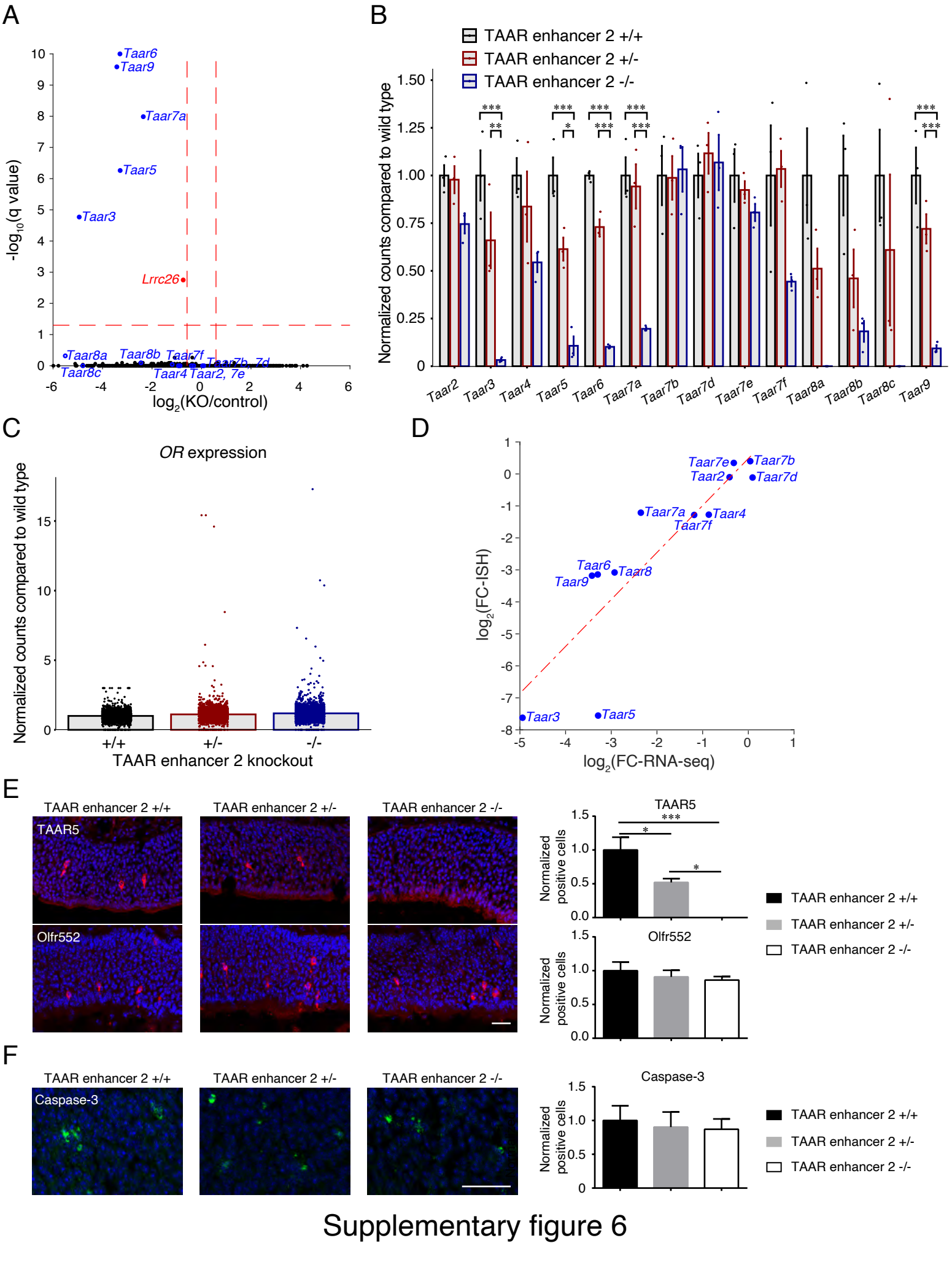
Supplementary figure 4

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



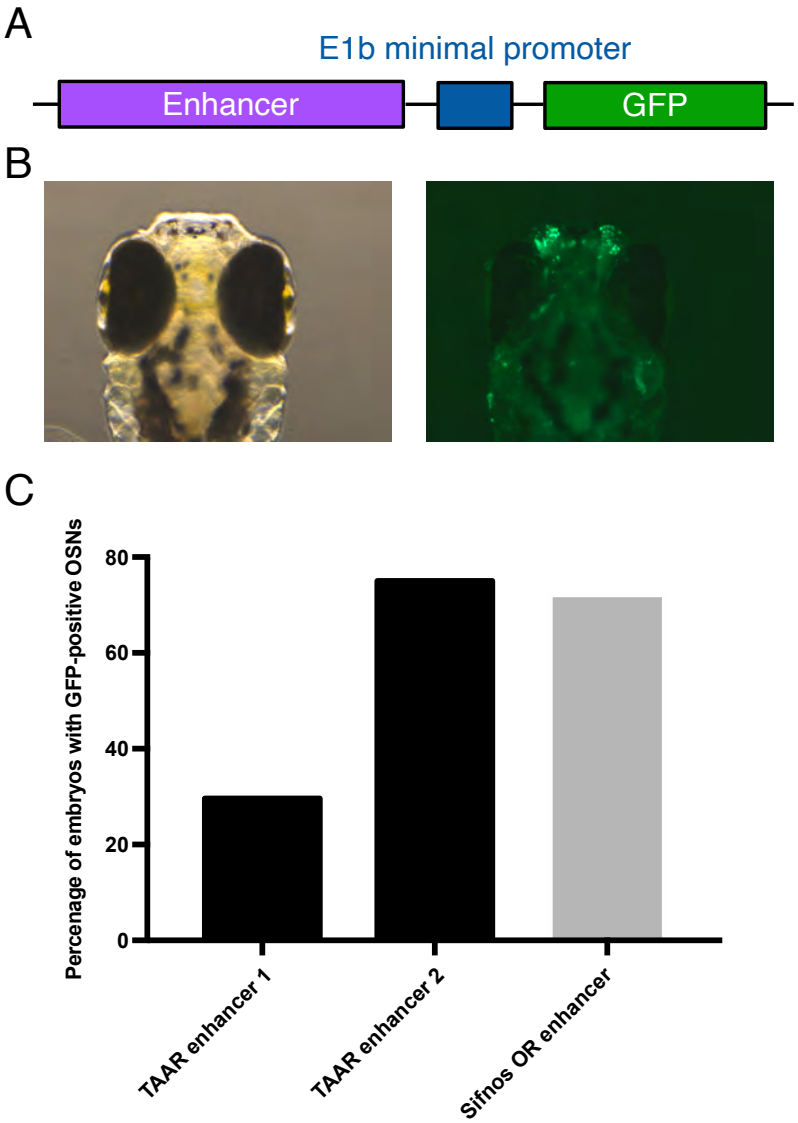
Supplementary figure 5

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



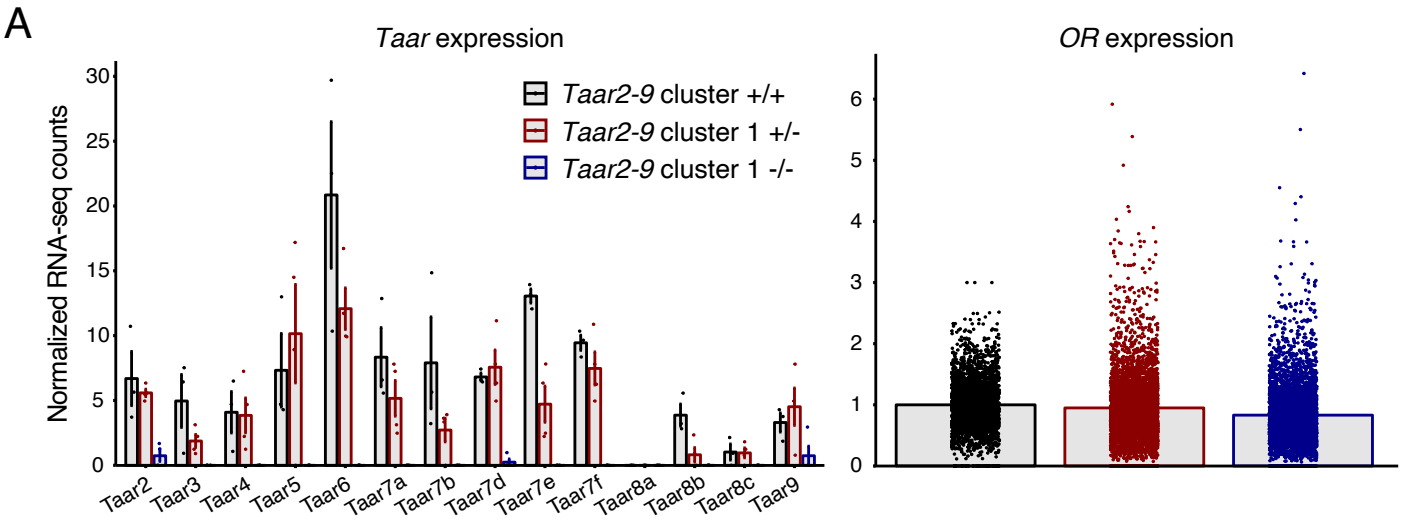
Supplementary figure 6

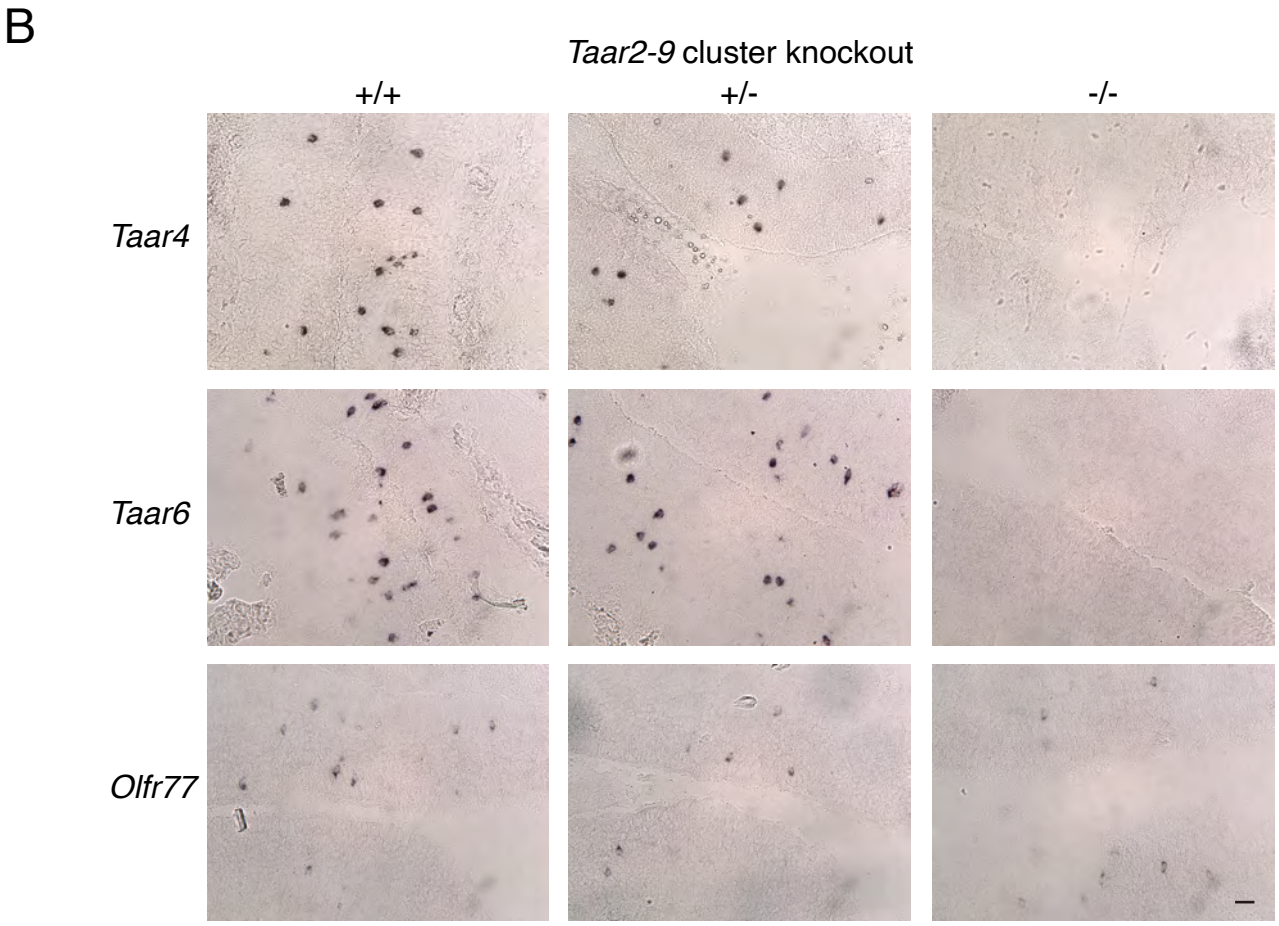
bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

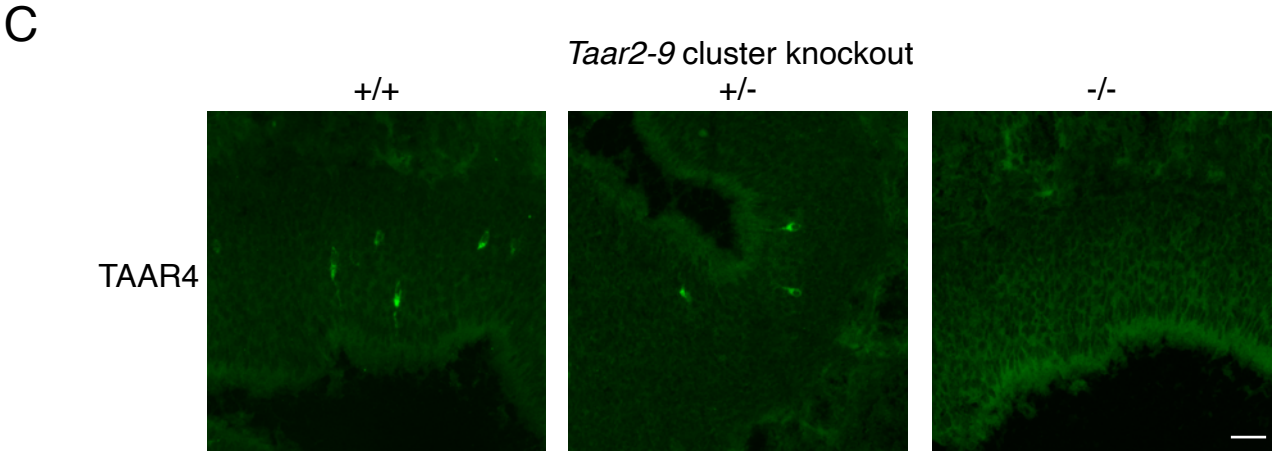


Supplementary figure 7

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.10.288951; this version posted September 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.







Supplementary figure 8

Supplementary Table 1. The information of extracted genome sequences from various mammalian species				
Species	Species	Genome assembly version	Genome coordinates	Genome regions
Mouse	Mus musculus	GRCm38.p4	chr10:23894688-24188961	Vnn1-Taar cluster-Stx7
			chr10:23920406-23941583	Taar1-Taar2
			chr10:23984609-23993481	Taar6-Taar7a
Human	Homo sapiens	GRCh38.p12	chr6:c132646026-132617150	Taar1-Taar2
			chr6:c132571359-132559024	Taar6-Taar7P
Chimpanzee	Pan troglodytes	Clint_PTRv2	chr6:c130003330-129974149	Taar1-Taar2P
			chr6:c129928126-129909449	Taar6-Taar8P
Orangutan	Pongo abelii	Susie_PABv2	chr6:c131193663-130936926	Vnn1-Taar cluster-Stx7
Gibbon	Nomascus leucogenys	Nleu_3.0	chr3:c119980447-119808147	Vnn1-Taar cluster-Stx7
Tarsier	Carlito syrichta	Tarsius_syrichta-2.0.1	NW_007247600.1:c380786-349157	Taar1-Taar2
			NW_007247600.1:c287550-257458	Taar6-Taar8
Rat	Rattus norvegicus	Rnor_6.0	chr1:c22596475-22578025	Taar1-Taar2
			chr1:c22537236-22528659	Taar6-Taar7a
Rabbit	Oryctolagus cuniculus	OryCun2.0	chr12:c122918612-122891431	Taar1-Taar2
			chr12:c122850827-122831128	Taar6-Taar6
Hedgehog	Erinaceus europaeus	EriEur2.0	NW_006804730.1:c662714-641243	Taar1-Taar2
			NW_006804730.1:c550178-516925	Taar6-Taar6
Pig	Sus scrofa	Sscrofa11.1	chr1:31037826-31074574	Taar1-Taar2
			chr1:31124661-31135694	Taar6-Taar7a
Cow	Bos taurus	ARS-UCD1.2	chr9:c70827654-70784406	Taar1-Taar2
			chr9:70591888-70609286	Taar6-Taar7aP
			chr9:c70554482-70533382	Taar8P-Taar7a
			chr9:c70711073-70693872	Taar6-Taar7a
			chr9:70618129-70633589	Taar8P-Taar7a
			chr9:c70534425-70518751	Taar7a-Taar7aP
Sheep	Ovis aries	Oar_rambouillet_v1.0	chr8:c64047965-64031195	Taar6-Taar7a
			chr8:c66398793-66380107	Taar6-Taar6P
			chr8:c64021497-64002370	Taar8P-Taar7a
Dolphin (Yangtze River do	Lipotes vexillifer	Lipotes_vexillifer_v1	NW_006784126.1:187583-324150	Vnn1-Taar cluster-Stx7
Sperm whale	Physeter catodon	ASM283717v2	chr10:30946785-31107969	Vnn1-Taar cluster-Stx7
			chr10:31012694-31027703	Taar1-Taar3P
Horse	Equus caballus	EquCab3.0	chr10:c80023263-80001724	Taar1-Taar2

			chr10:c79907708-79891692	Taar6-Taar7a
			chr10:79906594-79921740	Taar6-Taar7a
Cat	Felis catus	Felis_catus_9.0	chrB2:c122585612-122538799	Taar1-Taar2
			chrB2:122466294-122482960	Taar6-Taar7a
Big brown bat	Eptesicus fuscus	EptFus1.0	NW_007370651.1:55488823-55512018	Taar1-Taar2
			NW_007370651.1:55537933-55555994	Taar5-Taar9
Elephant	Loxodonta africana	Loxafr3.0	NW_003573420.1:27517488-27557889	Taar1-Taar2
			NW_003573420.1:27626820-27662685	Taar4-Taar7a
Armadillo	Dasypus novemcinctus	Dasnov3.0	NW_004498224.1:c223592-203709	begin-Taar2
			NW_004498224.1:c125945-91587	Taar6P-Taar7aP
Koala	Phascolarctos cinereus	phaCin_unsw_v4.1	NW_018343979.1:c15437684-15405259	Taar1-Taar2
			NW_018343979.1:c15183169-15157886	Taar6-Taar7a
Opossum	Monodelphis domestica	MonDom5	chr2:c407125000-407094705	Taar1-Taar2
			chr2:c406934871-406909558	Taar6-Taar7a
Platypus	Ornithorhynchus anatinus	Ornithorhynchus_anatinus-	NW_001794460.1:c1524776-1501778	Taar1-Taar2