Evolutionary diversity and function of odorant receptors in birds

Robert J. Driver[1,2*], Mona A. Marie[2], Victoria J. Ko[2], Priyanka Meesa[2,3], Wanting Sun[2], Renee J. Li[2], Michael S. Brewer[1], Hsiu-Yi Lu[2], Kevin F. P. Bennett[3], Ichie Ojiro[2,4], Nivritti E. Mantha[2,5], Hiroaki Matsunami[2,6*], and Christopher N. Balakrishnan[1,7].

[1] Department of Biology, East Carolina University, Greenville, NC 27858, USA

[2] Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, NC 27710, USA

[3] Present address: Georgetown University School of Medicine, Washington, DC 20007, USA

[3] Department of Biology, Pennsylvania State University, University Park, PA 16802, USA

[4] Present address: Department of Neuroscience, New York University, NY 10012, USA

[5] Department of Engineering, Smith College, Northampton, MA 01063, USA

[6] Department of Neurobiology, Duke University School of Medicine, Durham, NC 27710, USA

[7] Division of Environmental Biology, National Science Foundation, Alexandria, VA 22314, USA

* - Corresponding author.

Abstract

Odorant receptors (ORs) form one of the largest gene families in vertebrates; most mammals have hundreds of intact OR genes. Although birds display diverse breeding and foraging behaviors, they were long assumed to rely minimally on olfaction. Here we show that, similar to mammals, bird genomes often encode hundreds – and in some species, thousands – of intact ORs. We further present evidence that most avian ORs within the bird-specific gamma-c subfamily

have undergone extensive gene conversion. We show that avian ORs are expressed in olfactory sensory neurons and respond to defined odorants. Notably, we identify ligands for avian ORs, revealing both unique functions within the gamma-c clade and functional conservation with a deeply divergent mammalian OR. These findings uncover unexpected parallels between mammals and birds in olfactory system organization while revealing a distinctive evolutionary feature: widespread gene conversion shaping the majority of avian ORs.

Introduction

In 1834, Charles Darwin tested if birds could smell by hiding a piece of meat in a folded white paper. When Darwin showed the folded paper to vultures, they showed no interest in the concealed meat. However, when the paper was torn open and the food was revealed, the vultures began to flap and peck excitedly. Darwin wrote that "it would have been quite impossible to have deceive a dog [in this way][1]." This observation aligned with other early studies concluding that olfactory signals were of limited use to birds[1–3]. In modern times, behavioral work in birds has shown promising roles for olfaction in foraging, locating nest sites, seed caching behavior, and species recognition, among other behaviors[4–7]. Additionally, specific bird species rely on olfaction for foraging, including the turkey vulture (*Cathartes aura*) and many seabirds (order Procellariformes)[8,9]. However, how the bird olfactory system functions at the molecular or cellular level remains largely unexplored.

In vertebrates, odors are primarily detected with odorant receptors (ORs), a gene family of G protein-coupled receptors expressed in the olfactory sensory neurons (OSNs) of the olfactory epithelium[10,11]. To accommodate the incredible variety of odorants in nature, ORs

constitute the largest gene family in vertebrates, for example, there are more than 2,000 OR genes in elephants[12,13]. Our previous work has shown that birds have many more OR genes in their genomes than previously realized[14]. We found that genomic OR repertoires of five bird species ranged from 50 to over 350 intact ORs per species[14]. Our genomic analyses, in synergy with recent behavioral studies, suggest a greater role for the sense of smell in birds than previously appreciated[7,14,15].

Despite the potential importance of olfaction in bird behavior and ecology, there is little understanding of the evolutionary diversification and expression of ORs, an no information regarding avian OR binding properties. First, using high-quality, publicly available assemblies (largely from the Vertebrate Genomes Project) we document patterns of OR diversity across the avian tree of life and the evolutionary processes underlying this diversity. We then localized OSNs in bird olfactory epithelium, and tested the extent to which OR expression was localized to the OSNs. Finally, we determined whether bird ORs and OSNs are capable of responding to odor, and identified specific odorants that trigger responses. Together, we aim to test whether bird ORs are functional in the avian olfactory system and to reevaluate the historical view that olfaction is largely irrelevant to bird life history.

Results

Characterization of bird odorant receptor repertoires

To understand OR repertoire size and diversity across bird species, we used our OR identification and classification procedure and detected complete, putatively functional OR genes in 148 bird genome assemblies from across the phylogeny (Fig. 1A). Previously, we

demonstrated that short-read genome assemblies risk underestimating OR gene counts due to high sequence similarity between many bird ORs[14]. To confidently capture all OR genes, we selected only genome assemblies using long read sequencing technology, with a minimum contigN50 size of 7 Mb. We found a wide range of OR repertoires among birds (Fig. 1B). We found an average of 251.9 ORs per species, with a median of 146.5 ORs per species. We found an incredible range of OR repertoire in birds (Fig. 1B). At the maximum, we identified 3,750 ORs in North Island brown kiwi (*Apteryx mantelli*), the largest number of ORs known from any animal, surpassing elephants (Fig. 1B). Conversely, we also found species with few ORs. The lowest OR counts of all birds were in crows, with only 7 in carrion crow (*Corvus corone*). Taken together, like in mammals, bird ORs are diverse and vary substantially between species.

Evolutionary Dynamics of Avian-specific Gamma-c ORs

ORs are broadly grouped into either Class I or Class II based on sequence similarity[16]. Out of a total of 37,286 ORs across all species in birds, 3.34% included alpha (Class I) ORs, while 96.61% were gamma or gamma-c (Class II) ORs (Fig. 1A, C). This large proportion of Class II ORs relative to Class I ORs is consistent with OR repertoire profiles across mammalian species. Within Class II, birds have a unique expansion of ORs in a specific OR subfamily[17,18]. This OR subgroup, known as gamma-c ORs, constituted 78.85% (29,413 total ORs) of the total bird ORs we identified, including 88.5% (419 out of 473 total) of all ORs in the chicken (Fig. 1C). In addition to being the most numerous bird OR group, gamma-c sequences within each species are highly similar to each other, as indicated by short branch lengths on the OR phylogeny (Fig. 1C) and an average of 83.83% identical nucleotide and 77.78% identical amino acid sequences across 419 chicken gamma-c ORs. Unlike other OR gene families, in which paralogs from different

species generally cluster together[19], sequence similarity between gamma-c ORs is higher within a species than between species from different bird families (Fig. 1D). However, this is not true for gamma-c OR genes from closely-related species within the same bird family, where we observe interdigitation of gamma-c genes among species, illustrating that fully monophyletic clades form following millions of years of evolutionary divergence (Supp. Fig. 1).

The unusual species-specific phylogenetic pattern we observe between bird families (Fig. 1D) would be consistent with either parallel duplication of a single ancestral gamma-c OR in bird families or concerted evolution, including unequal crossing over or gene conversion[20]. Gene conversion transfers short stretches of DNA between paralogs, typically during meiosis through DNA repair mechanisms[21,22]. This process leaves several characteristic signatures[23]: (i) incompatibility between gene trees and the true history of gene duplication, resulting in species-specific clades; (ii) the presence of subregions within genes that are nearly identical in nucleotide sequence, contrasted with other subregions that retain higher sequence diversity; and (iii) phylogenetic relationships reconstructed from homogenized subregions versus non-converted regions are incongruent.

Consistent with these predictions, we found that gamma-c ORs within species contain long stretches of highly similar nucleotides in the open reading frame (ORF). For example, in chicken (and six other bird species), a 292-bp region near the 5′ end of the ORF is nearly identical within the given species across paralogs (Fig. 2A, 2B). Additional shorter homogenized stretches occur in other parts of the ORF. Importantly, these stretches show similarity across all codon positions, including synonymous third positions, indicating that homogenization is not driven by amino acid–level selection (Fig. 2A). In contrast, human ORs exhibit relatively

uniform diversity across the ORF (Fig. 2A). Additionally, statistical analysis of gamma-c sequence alignments in chicken indicated thousands of recombination events, a signature of gene conversion. In pairwise comparisons of aligned gamma-c ORs, at least 2,000 segments within the homogenized regions of the open reading frame showed evidence for undergoing gene conversion events ($P < 0.0001$, Supp. Mat., File 7)[24]. Consistent with gene conversion events, gene trees reconstructed from homogenized versus variable regions of gamma-c ORFs were incongruent across six bird species (Fig. 2B). Moreover, gamma-c ORs are arranged in genomic clusters, often on avian dot chromosomes – for example, all chicken gamma-c ORs are found on dot chromosomes 16, 29, and 34[14,25]. Because gene conversion most frequently occurs among genes in close proximity[26], the persistence of chromosome-specific substitutions (Fig. 2C) further supports intrachromosomal gene conversion. Together, these findings indicate that gamma-c ORs are mosaics of homogenized and less-homogenized subregions, and suggest that their unusual phylogenetic pattern reflects extensive gene conversion during bird evolution.

Expression of bird ORs in OSNs

The results above demonstrate that birds have genomic OR repertoires comparable in size to those in mammals. We next sought to determine how bird ORs function in the olfactory system, if at all. In animals, olfactory bulb size relative to the whole brain is a long-standing measurement used to assess potential relevance to the olfactory system[18,27]. We found a positive correlation between the revised counts of intact ORs from genomes with long-read technology (see above) with the size of the olfactory bulb to brain ratio, both overall (Fig. 3A) and individually within each subclass of OR (Supp. Fig. 2). This result is consistent with the idea that all classes of ORs are functioning in the olfactory system.

Having made a potential functional link between OR count and olfactory system function, we next investigated bird OR expression. To be functional and tied to olfaction, we would expect bird OR expression in the olfactory epithelium, and, specifically, the OSNs. To test this, we conducted bulk RNA-seq from olfactory epithelium and used pectoralis muscle as a negative control in the chicken, zebra finch (*Taeniopygia guttata*), and brown-headed cowbird (*Molothrus ater*), and mapped reads to OR genes. We found that overall expression of ORs was over 70 times higher in the chicken olfactory epithelium compared to pectoralis muscle (Fig. 3B, student's paired t-test, $t = 5.83$, $P < 0.001$). OR expression varied among olfactory epithelium samples within a species, but was consistent with the expression levels of known OSN markers in each sample, suggesting heterogeneity in olfactory epithelium content in each sample (Supp. Fig. 3). We also found that most individual OR genes were expressed in the olfactory epithelium, including the gamma-c ORs (Fig. 3C). Further, we found cellular-level OR expression in the chicken olfactory epithelium through *in situ* hybridization (Fig. 4). We found that the distribution of ORs (including gamma-c ORs) in OSNs colocalized with Cnga2, an established OSN marker (Fig. 4, Supp. Fig. 4). We note that, due to their high nucleotide-level similarity, our gamma-c OR probe is expected to hybridize to most, if not all, gamma-c ORs. Our observation that far more OSNs are labeled with the gamma-c OR probe than with other OR probes is consistent with this notion. This expression confirms our expectation that bird ORs contribute to the sense of smell in birds.

Bird ORs respond to odors

Given the presence of ORs in chicken OSNs, we anticipated that bird ORs respond to odors. First, we aimed to identify active ligands for gamma-c ORs in the chicken. Given the

large number of gamma-c ORs and overall high sequence similarity, we first tested a consensus OR designed from the gamma-c OR alignment, based on our hypothesis that consensus ORs both support robust functional expression in heterologous cells and approximate the response of native ORs[28] (see STAR methods). We first tested the consensus of all chicken gamma-c ORs for a response to a panel of odors. We found that the chicken gamma-c consensus OR responded to 2-isopropyl-3-methoxypyrazine and 2-isobutyl-3-methoxypyrazine, members of the key pyrazine food odor group (Fig. 5A). Having methoxypyrazine response from a consensus OR, we next cloned native ORs directly from chicken genomic DNA (Fig. 5B see STAR methods). To test the function of the cloned ORs, in addition to N-terminal rhodopsin tag, we modified the C-terminal ends of each OR to enhance their expression[29] (see STAR methods). Several native ORs responded to 2-isopropyl-3-methoxypyrazine, 2-isobutyl-3-methoxypyrazine, and 2-sec-butyl-3-methoxypyrazine (Fig. 5C). Importantly, each OR has a unique relative response profile against a panel of pyrazines (Fig. 5C). Our results show that bird ORs, particularly those within the gamma-c subfamily, have the ability to respond to odors and each OR may have unique function.

Bird odorant receptor response to odors activating mammalian orthologs

In addition to gamma-c ORs, which are highly variable across bird species, we wanted to test odor responses from a bird OR that is stable across evolutionary time, notably, OR51E2 that belongs to the alpha (Class I) family and responds to short chain carboxylic acids. Recent studies revealed the structure and binding mechanisms of human OR51E2[30,31]. We found that chicken has an orthologous OR with human OR51E2 (Fig. 5D), and we therefore tested chicken OR51E2 response to a set of short chain carboxylic acids. We found that chicken OR51E2 responds to the same odorants, propionate and acetate, as the human ortholog, but additionally responds to

butyrate, indicating moderate functional differences between the two ORs. (Fig. 5E). We therefore show that mammalian and bird OR orthologs show consistent odor response profiles.

Bird OSNs respond to odors in vivo

After demonstrating an *in vitro* response of chicken gamma-c ORs to pyrazines, we next tested for a response to pyrazines in the live chicken to test whether OSNs expressing gamma-c ORs are activated by odor stimulation (Fig. 6A). We exposed individual chickens in a container to a cassette containing either 1% (v/v) or 10% 2-isobutyl-3-methoxypyrazine. We also exposed chickens to 1% acetophenone, commonly used in mammalian olfactory experiments, as well as to no-odor stimulation as controls[32]. We then labelled gamma-c expressing neurons with *in situ* hybridization, and performed immunohistochemistry against phospho-ribosomal protein S6 (pS6) to label active neurons (see STAR methods)[28]. In chickens receiving no odor, there was minimum pS6 labeling overlapping with gamma-c expression. Following 1% acetophenone exposure, we observed pS6-positive OSNs, but this signal did not colocalize with OSNs expressing gamma-c (Fig. 6B-C). We found that a subset of OSNs expressing gamma-c ORs were labeled with pS6 antibodies following exposure to 1% and 10% 2-isobutyl-3-methoxypyrazine (Fig. 6B-C). Quantification of pS6 intensity in OSNs expressing gamma-c ORs demonstrated that pyrazine stimulation, but not acetophenone stimulation, produced a significant increase in pS6 signals in OSNs expressing gamma-c ORs relative to no odor controls (Fig. 6C, Supp Fig. 5). Together, these results support that a subset of OSNs expressing gamma-c ORs is activated by exposure to 2-isobutyl-3-methoxypyrazine.

Discussion

In this study, we show that birds often have hundreds to thousands of ORs in their genomes, including the organism with largest known OR repertoire. We find that most bird OR genes belong to an expansion of a specific Class II subgroup, known as the gamma-c ORs, and that these receptors may uniquely evolve through gene conversion. In the chicken, we demonstrate that ORs are expressed in OSNs and that some of them respond to pyrazine odors. We also show that for OR51E2, a highly conserved OR across vertebrate evolution, chicken and human ORs respond to the same short chain carboxylic acids.

Diverse bird species have OR counts comparable to other vertebrates

We analyzed only genomes generated with long-read sequencing and a minimum contigN50 size of 7 Mb in order to avoid undercounting ORs caused by nearly identical sequences within stretches of gamma-c ORs[14]. As a result, we find that bird OR repertoires had much larger counts than previously reported[12,27]. For example, remarkably, the North Island brown kiwi OR repertoire changed from 63 based on the short-read genome[12] to 3,750 in the present study based on the long-read assembly. The kiwi forages at night and uses its bill, with nostrils located at the tip, to probe for soil invertebrates[33]. It has the largest olfactory bulb relative to brain size of any bird, consistent with our finding that olfactory bulb size correlates with OR counts[34,35]. Together, these results further support an exceptional enhancement of olfactory ability and sensitivity in North Island brown kiwi. These thousands of ORs in kiwi contrast with birds that have few genomic ORs, such as crows, which also exhibit smaller olfactory bulbs[35–37]. High variation in OR counts in birds is analogous to the pattern observed in mammals, with fewer than 20 ORs in several toothed whales but over 2,000 ORs in elephants[12,13].

Although the average OR count in birds (258 ORs) was lower than reported mean mammalian OR counts (680 ORs[12]), the average was comparable to functional genomic ORs in crocodiles (276 ORs[12]), and was greater than the average OR counts reported for bony, cartilaginous, and jawless fish[12]. As OR counts may correlate with olfactory acuity such as the ability to discriminate odor pairs[38], it is possible that many bird species have olfactory acuity that is similar to species with known reliance on olfaction.

ORs are expressed in the OSNs in birds

Our study, including bulk RNA-seq in three bird species and in situ hybridization in chicken, suggests the expression of a large majority of ORs including gamma-c OR genes in the OSNs. These results are consistent with two earlier bulk RNA-seq studies in other bird species that also showed OR expression, including gamma-c ORs, in the bird olfactory epithelium[39,40,]. Here we show consistent results that many genes in the alpha (Class I), gamma and gamma-c (Class II) OR repertoire show expression in the olfactory epithelium in three species.

We further found that in the chicken, OSNs are present in the posterior turbinate of the olfactory epithelium, and that bird ORs are expressed in the OSNs. Although previous studies have shown the location of neurons in the olfactory epithelium of embryonic chicken and kiwi[35,41,42], this is the first time that the location of the OSNs was shown using known OSN markers such as Cnga2 and ORs. Additionally, we saw multiple ORs that belong to alpha, gamma, and gamma-c ORs expressed in the OSNs. Within OSNs, more neurons were positive for gamma-c ORs than alpha or gamma ORs. Our probe used for in situ hybridization likely hybridizes to mRNA transcripts from most gamma-c genes due to the high overall identities and

nearly identical nucleotide stretches. Therefore, we are limited in our ability to test the location of individual gamma-c ORs. This high sequence similarity of the gamma-c ORs may make it difficult to determine whether each avian OSN expresses only a single OR, as occurs in mammals. In summary, we conclude that just as in mammals, bird ORs expressed in the OSNs of the olfactory epithelium.

Gene conversion as an evolutionary mechanism in sensory biology

Here, we present evidence for extensive gene conversion of gamma-c ORs, resulting in highly similar paralogs within species. Sequence homogenization through gene conversion may purge mutations in regions of the OR important for folding, trafficking, or stability. Such a mechanism is likely to be more common in birds, as gene conversion has also been implicated in the evolution of major histocompatibility complex genes (MHC) and toll-like receptor genes[22,43]. In particular, in the chicken a large tandem cluster of the avian MHC gene family is located on dot chromosome 16, the same dot chromosome that contains some gamma-c ORs[44]. Consequently, gene conversion could cause different members of gamma-c ORs within a species to acquire similar functions toward ecologically relevant odors for the species yet potentially limiting their ability to discriminate a broad range of environmental odors.

Interestingly, however, our data suggest that many putative ligand-interacting regions of the OR, such as transmembrane domains 5 and 6, remain relatively diverse at both the nucleotide and amino acid levels (Fig. 2A-B). This raises the possibility that gamma-c ORs within a species may still retain functional diversity despite extensive homogenization elsewhere. Future studies that functionally characterize a large set of ORs will be essential for answering this question.

Bird ORs respond to odors

We report the first ever response of a bird OR to an odorant by showing that chicken gamma-c ORs respond to pyrazine odors 2-methoxy-3-isobutylpyrazine and 2-sec-butyl-3-methoxypyrazine. To humans, 2-methoxy-3-isobutylpyrazine has a potent green pepper-like aroma (pers. obs.). In nature, pyrazines are an important class of key food odors and it is possible that chickens detect pyrazines when foraging[45]. Future studies in chickens can help uncover the response and possible relevance to pyrazines in chicken behavior and ecology. In one example, chickens with prolonged exposure to 2-methoxy-3-isobutylpyrazine had a larger egg size[46]. Therefore, there could be potential implications of pyrazine odor detection and bird development.

Despite high sequence similarities, our results support the idea that different gamma-c ORs respond to different odorants based on our pS6 immunostaining showing that only a subset of gamma-c expressing OSNs responds to 2-methoxy-3-isobutylpyrazine. Future studies can characterize the response of different gamma-c ORs in the chicken, as well as across other bird species. In addition to the first response of a bird OR to an odor, this is also the first response of a gamma-c OR to an odor, implicating this diverse and prolific OR subfamily in the olfactory system.

OR51E2 binding properties conserved across vertebrate evolution

In addition to finding that odors bind to the highly evolutionary dynamic gamma-c ORs, we found that bird and human OR51E2, forming a clear orthologous relationship across many vertebrates, shows conserved response to short chain carboxylic acids. Future investigations

should describe the responses of other bird mammal OR orthologs to determine if functional conservation is widely conserved across both orthologs and paralogs, as well as characterize the binding properties of these ORs across additional vertebrate classes.

Limitations of the study

In our genomic analyses, we used assemblies generated with different sequencing technologies and assembly methods, which may have affected the recovery of ORs and led to varying degrees of mapping errors. In our RNA-seq analyses, the extremely high sequence similarity among gamma-c OR genes made it challenging to assign some reads to individual ORs. Many ambiguous reads were excluded when mapping reads to specific OR genes, potentially underestimating gamma-c OR expression. For our functional evaluation of ORs in heterologous cells, some ORs did not reach the cell membrane, leading to an apparent lack of response. We relied on pS6 immunostaining to assess colocalization with individual OR expression and OSN activation; however, positive detection of a neuronal activation marker (i.e., pS6) does not necessarily equate to an electrical response of OSNs. Finally, the high sequence similarity of chicken gamma-c ORs prevented our *in situ* hybridization probe from targeting individual genes; instead, our results reflect expression of many gamma-c ORs collectively.

Acknowledgements

We would like to thank the Vertebrate Genomes Project (VGP) for providing the majority of the high-quality long read genome assemblies used in this study. Without the VGP, we would have been unable to recover gamma-c OR repertoires from many bird species, and we greatly thank

this tremendous resource. We would like to thank Dr. Kenneth Anderson, Becca Wysocky, and Christina Sigmon at North Carolina State University Department of Poultry Science for their dedication and assistance with obtaining chicken tissue and their help with chicken odor exposure experiments. We would like to thank Dr. Richard Mooney, Dr. Audrey Mercer, and Michael Booze at Duke University School of Medicine Department of Neurobiology for assistance with zebra finch tissues. We thank Dr. Marc Schmidt at University of Pennsylvania for graciously collecting brown-headed cowbird tissue. We thank Emily Xu for chicken tissue sectioning and manuscript review. This work was funded by National Science Foundation awards IOS 1655730 (to C.N.B), DBI 2208965 (to C.N.B.), and DEB 1457541 (to R.J.D.). This work was funded by National Institutes of Health R01 DC022770, R01 DC020353, and R01 DC021585 (to H.M.). This work was funded by the Society for Integrative & Comparative Biology, the American Museum of Natural History, the Society for the Study of Evolution, the American Genetics Association, the American Ornithological Society, and the Wilson Ornithological Society.

Figure captions.

Figure 1. Intact genomic OR repertoires in birds. **A.** Intact OR counts from 148 bird species assemblies using long read sequencing. Black bars show total intact genomic OR repertoires. Purple bar graph shows the proportion of each species' genomic repertoire across Class I alpha receptors (light gray), Class II gamma receptors (light purple), and Class II gamma-c receptors (dark purple). Species relationships are illustrated by the phylogeny. **B.** Histogram displaying the range of total intact genomic OR counts in birds. We binned total OR counts into sets of 15 ORs. Birds of interest are displayed in black using Phylopic. Dark blue line shows the median count

for birds. A comparison with genomic intact OR counts from mammals is shown with red arrows using orange Phylopic images. **C.** Phylogeny of 473 intact genomic ORs in the chicken. Gray shows Class I alpha receptors, light purple shows Class II gamma receptors, dark purple shows Class II gamma-c ORs. **D.** Phylogeny of gamma-c OR genomic repertoires from six bird species representing different families. Each color is a different bird species.

Figure 2. Evidence for gene conversion in bird gamma-c ORs. **A.** Nucleotide sequence identity in from aligned OR repertoires. Chicken gamma-c OR (top left) ORF shown in purple, with red highlighting a homogenized subregion and blue highlighting a comparatively diverse region. Human OR alignments show that nucleotide sequence identity is uniform throughout ORF (top right). If only third codon positions are shown, nucleotide sequence identity remains distinct across subregions in chicken gamma-c ORs (top left). Nucleotide sequence at third codon positions remains uniform across subregions in human ORs. **B.** Phylogenies from six bird species using distinct subregions of gamma-c. Purple bar shows chicken gamma-c ORF as in panel A, with red correspond to a homogenized region and blue corresponding to a diverse region (see corresponding regions in panel A). Phylogeny from homogenized region (left) forms species-specific clades largely corresponding to Figure 1D. Phylogeny from diverse region (right) forms unique phylogenetic patterns without clear species-specific clades. **C.** Chromosome-specific signatures of gamma-c ORs in chicken. Coordinates of representative chicken gamma-c ORs from three dot chromosomes are shown. Nucleotides divergent from consensus sequence are colored. Substitutions largely follow chromosome-specific patterns.

Figure 3. Evidence for role of OR function in bird olfactory system. **A.** Bird total intact genomic OR counts correlate with olfactory bulb to brain size ratio in birds. **B.** Total OR expression levels measured from olfactory epithelium and pectoralis muscle in three bird species. **C.** OR

expression levels of individual ORs in olfactory epithelium and pectoralis muscle tissue. Each species is depicted with a phylogeny of its intact genomic OR repertoire. Inner ring shows expression of each OR in pectoralis muscle, outer ring shows expression of each of in olfactory epithelium. Cyan indicates low level of expression, red high level of expression.

Figure 4. Odorant Receptors (ORs) distribution in chicken olfactory epithelium (OE). **A.** In situ hybridization (ISH) for CNGA2, Gamma-C, Or10a7, and Or52r1 (top to bottom, respectively). **B.** Proportions for Ors expression relative to the OSN marker CNGA2. Scale bar shown in 100uM.

Figure 5. Functional testing of ORs *in vitro*. **A.** Phylogeny showing the position of the chicken consensus sequences in relation to native chicken ORs. Orange is the consensus of all chicken gamma-c, red, blue, green, and pink are consensus sequences of subsets of chicken gamma-c ORs. **B.** Increasing response of chicken gamma-c consensus OR (green) with increasing concentrations of 2-sec-butyle-3-methoxypyrazine (left) and 2-isopropyl-3-methoxypyrazine (right) odors relative to empty vector (gray). **C.** Responses of chicken gamma-c subclade consensus ORs (blue, green, red, pink) with increasing concentrations of the two pyrazine odors (see 5B) relative to empty vector (black). The strongest dose response curve with increasing concentrations of pyrazines was present in subclade consensus OR 2. **D.** Increasing response of native chicken ORs with increasing concentrations of 2-isopropyl-3-methoxypyrazine and 2-methoxy-3-isobutylpyrazine relative to empty vector. **E.** Location of human OR51E2 (red) when placed on the chicken OR phylogeny. Human OR51E2 has a 1:1 orthologous relationship with chicken OR51E2 (blue). **F.** Response of human OR51E2 (left) to increasing concentrations of three short chain fatty acid odors, acetate (red), propionate (black), and butyrate (purple). Response of chicken OR51E2 to the same odors (right).

Figure 6. Pyrazine exposure increases pS6 in gamma-c-positive olfactory sensory neurons (OSNs). **A.** Scheme of pS6 experiment. We placed adult chickens in containers and exposed to odorants for 1 hour. We then extracted tissue from the olfactory epithelium within the chicken maxilla, and sectioned tissue using Leica CM 1850 Cryostat. We then performed *in situ* hybridization (ISH) and pS6 Immunoflorescence, performed image analyses and quanitification. **A.** In situ hybridization (ISH) combined with Immunoflorescence (IF) for gamma-c and pS6, respectively. **B.** Semi-quantitative analysis for pS6 color intensity comparing treatment groups No odor, Acetophenone 1% (ACE15), 2-methoxy-3-isobutylpyrazine 1% and 2-methoxy-3-isobutylpyrazine 10%. Scale bar shown in 20uM.

Supp. Fig. 1. Closely related bird species within the same bird family do not show monophyletic gamma-c OR clades. Two species in the Old World warbler family (Sylviidae) as well as tufted duck (*Aythya fuligula*) and two swan species (*Cygnus atratus* and *C. olor*) show interdigitation of gamma-c OR in the phylogeny.

Supp. Fig. 2. Olfactory bulb size is correlated with OR counts of OR subfamilies. This relationship suggests that each individual subfamily may be involved in smell, including the gamma-c ORs.

Supp. Fig. 3. Expression of OSN markers is higher in olfactory epithelium than pectoralis muscle. We summed expression of three OSN positive markers, CGNA2, OMP, and ADCY3, and observed high levels in the olfactory epithelium samples. This suggests that we have isolated tissue containing OSNs. Additionally, samples with larger OSN marker expression consistently had larger OR expression (Fig. 3B).

Supp Fig 4. Semiquantitative analysis of ISH signals for OSN marker CNGA2 and chicken ORs in the olfactory epithelium. Five randomly selected fields of view (FOVs) were analyzed for each probe staining to quantify the percentage of olfactory sensory neurons (OSNs) expressing the canonical OSN marker CNGA2, and different chicken ORs, gamma-c, Or10a7, and Or52r1 within the chicken olfactory epithelium (OE).

Supp Fig 5. pS6 intensity semi-quantitative measurement in individual Gamma-C–positive neurons. Each dot represents the pS6 fluorescence intensity measured per cell within the identified Gamma-C–positive population. Semi-quantification was performed at the single-cell level, and data are expressed as relative fluorescence intensity normalized to background signal.

STAR Methods

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Robert Driver (rjd43@duke.edu).

Materials availability

Correspondence and requests for materials should be addressed to the lead contact. Any newly generated plasmids can be obtained from the lead contact following reasonable requests.

Data and code availability

Accession numbers are listed in the key resources table.

This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.

All original code has been deposited at Dryad and is publicly available as of the date of publication. DOIs are listed in the key resources table.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental model and study participant details

In vivo animal studies:

The RNA sequencing study design looked for the presence and expression levels of ORs in bird olfactory epithelium and pectoralis muscle. Five individuals of four species were obtained in this study for gene expression analyses: *Gallus gallus*, *Taeniopygia guttata*, and *Molothrus ater*. In all three species, individuals were obtained from captive group-housed populations and all individuals were in good health. Individuals were sacrificed as part of ongoing experiments in their laboratory. Obtaining olfactory epithelium and pectoralis tissue was a byproduct of previously scheduled sacrifices for other purposes. IACUC protocols for proper care and sacrifice were followed for the associated ongoing experiment. For *Taeniopygia guttata*, and *Molothrus ater*, all individuals were wild-type adults. For *Gallus gallus*, individuals were 21-week-old Hyline W-36 white leghorns. For *Molothrus ater*, all indiviuals were male. For *Gallus gallus* and *Taeniopygia guttata*, all individuals were female.

For *in situ* hybridization, we used three female *Gallus gallus* Hyline W-36 white leghorns at 60 weeks of age.

In vitro cell lines:

All cell-based assays used Hana3A cells, a cell line derived from HEK293T cells, originally from an immortalized human embryonic kidney cell line[47]. Hana3A cells express RTP1L, RTP2, REEP1, and G_{aolf}, proteins that aid in the expression of ORs and the transport of ORs to the cell surface[47]. We grew Hana3A cells in Gibco Minimal Essential Media ThermoFisher FBS, 5mL Gibco Glutamax, Gibco streptomycin, and penicillin.

Method Details

Assembly selection

We investigated OR diversity in birds by selecting publicly available genome assemblies on GenBank (Supp Mat., https://www.ncbi.nlm.nih.gov/genbank/). Assemblies for each species implemented some form of long-read sequencing technology, including Pacific Biosciences or Oxford Nanopore methods. Genomes varied in the assembly methods used and in the size and total number of contigs and scaffolds. We selected only assemblies using long read sequencing, with a minimum contigN50 size of 7 Mb, due to the difficulty in recovering total OR counts in assemblies with shorter contigs[14]. In total, we analyzed 148 different bird assemblies, including species from the three main lineages of birds, the Palaeognathae, Galloanserae, and Neoaves. The species set represents diverse ecology, diets, and trophic levels.

OR identification and classification

To detect putatively functional ORs in the selected genomes, we created a BLAST query with a set of 2,110 OR protein sequences from 6 mammals (*Ornithorhynchus anatinus, Didelphis virginiana, Bos taurus, Canis lupus, Rattus norvegicus, Macaca mulatta*), 2 birds (*Gallus gallus, Taeniopygia guttata*), and 1 crocodilian (*Gavialis gangeticus*). We obtained this query OR set by combining previously published OR subgenomes[17,18,48]. Using this query file, we performed

TBLASTN searches against all 148 bird genomes with a threshold of $E < 1e–20$. The TBLASTN –num_alignments option was set to 200,000 to capture all genomic ORs similar to a single query sequence. To remove pseudogenized and truncated ORs, we filtered for hits > 250 amino acids long. For any single location on the genome, we filtered out hits within 100 bp of each other, and selected the lowest *E*-value associated with that location.

After obtaining unique BLAST hits, we extracted the associated nucleotide sequence from the genome as well as 300-bp regions flanking the hit both upstream and downstream. We used a modified Perl script to detect open reading frames (ORFs) within each extracted region[49,50]. We then aligned these ORFs to each other as well as to the human Olfactory Receptor Family 2 Subfamily J Member 3 (OR2J3) sequence using the E-INS-I default parameters in MAFFT[51]. Using the previously characterized transmembrane domains of OR2J3 as a guide, we removed any sequences that had five or more amino acid insertions or deletions within a transmembrane domain in the alignment[50,51]. This included ORFs with stop codons appearing prior to the end of the seventh transmembrane domain. We performed this filtration with custom python scripts (see Supp. Mat., File 6).

Using this alignment, we recorded the position of the first amino acid in the first transmembrane domain. To estimate the location of the ORF start codon, we used modified Perl scripts to find the most appropriate methionine upstream of this recorded transmembrane start position[49,50]. ORF sequences were then truncated at the 5’ ends to begin with this methionine. This set of ORFs was then aligned using the E-INSI-I parameters in MAFFT[51] to a set of *T. guttata* reference ORs[52] as well as 11 non-OR rhodopsin-like G-protein coupled receptors (non-OR GPCRs) that functioned as an outgroup[50]. We then used clustalW to generate a neighbor-joining

tree from this alignment with 1000 bootstraps, gaps removed, and Kimura's distance correction[53,54]. We then removed any ORFs that were phylogenetically more closely related to the non- OR GPCRs.

We classified all remaining ORFs as functional ORs. Using this final set, we ran a maximum likelihood tree using IQ-TREE with automatic model selection and 1000 SH-like approximate likelihood ratio test replicates[55]. Using ML support values, we collapsed all nodes <50% support into a polytomy using iTOL software, and rooted the tree using the ancestral branch leading to the 11 non-OR GPCRs[56]. We classified bird ORs into subfamilies alpha, gamma, and gamma-c based on the subfamily of the query sequence used to identify the OR and the location of the OR in one of the three distinct avian OR clades[48,52]. We then counted the final number of OR sequences as well as the number of ORs from each subfamily.

Estimation of tree topology

To analyze OR counts in a phylogenetic context, we sought to create a phylogeny of the surveyed bird species. The bird species used in this study are a unique set, with no preexisting published phylogenies containing the species in a single tree. Therefore, we used topologies from existing phylogenies in the literature. We used an established bird phylogenetic tree as the topology for our tree structure[57]. Three species in our analysis were not present in the original phylogeny- *Manacus candei*, *Pyrocephalus nanus*, and *Oenanthe melanoleuca*. These species are a result of phylogenetic splits, but the sister species were present in the original phylogeny[57]. Therefore, the species in the dataset were placed at the position of their respective sister species in tree[57]. We then used the drop.tip() function with the *ape* package[58] in R to remove all species not present in our analysis, leaving us with a tree of species with OR counts. Our resulting tree

had polytomies in the tanager group, which we then resolved manually using reference topologies[59,60].

Estimation of branch lengths

To determine the branch lengths for our literature-based topologies, we mined the genomes of our 148 species subset for ultraconserved elements (UCEs). We downloaded fasta files from GenBank (accession numbers of assemblies in Supp. Mat., File 1). We next followed the UCE discovery procedure recommended in PHYLUCE[61] (Faircloth 2015) to extract UCE loci from reference genomes. We next aligned UCE loci from all genomes using MAFFT[51] (and trimmed using GBLOCKS[62], both implemented in PHYLUCE. We subset the alignments to retain only those with no missing data and concatenated them into a single alignment. We ran IQ-TREE[55] using a subset of 15 taxa and the model finder option to select a sequence evolution model. We ran IQ-TREE on the full dataset constrained to the topology of Stiller et al.[63] and using the best model from the subset run (TVM+F+R6). Finally, with R package ape (R Core Team 2024,[58]) we converted the resulting tree to an ultrametric tree using the relevant calibration dates used by Stiller et al.[63].

Olfactory bulb size: phylogenetic generalized least squares

The olfactory bulb size relative to the brain size was available for 24 species in our dataset, published in Corfield et al.[34]. We omitted species in this analysis that were not represented in the published dataset. To control for the phylogenetic non-independence of our trait comparisons across bird species, we ran phylogenetic generalized least squares (PGLS) models. The phylogenetic trees with branch lengths generated from the UCE dataset were converted to a

correlation structure in R using the *ape* package function corBrownian to estimate a Brownian motion (BM) model of trait evolution and corMartens to estimate an Ornstein-Uhlenbeck (OU) model[64]. The OU model may better replicate actual biological processes due to an additional parameter to the "random walk" of BM in that there is a greater attraction to an initial central value the further the trait is from this value. We then used the function gls in the R *nlme* package. For each trait comparison, we compared the AIC values of each model to determine whether to select BM or the additional parameter in OU. We then ran ANOVA tests on our models followed by general linear hypotheses tests to determine significance, using the *multcomp* package in R.

Olfactory epithelium sample collection (for downstream RNA-seq)

To determine the location of the bird OE and specific OE regions (the anterior, middle, and posterior conchae), we referenced morphological descriptions and images of the maxilla[36,65]. We originally practiced dissections on bird carcasses donated by the North Carolina Museum of Natural Sciences. In this unique dissection, the maxilla was cut transversely through the nares and then from this incision the sides of the maxilla were cut proximally towards the lores. There were three cuts in the maxilla, one transverse and distal, the other two sagittal from the nares to the lores. From this, the proximal half of the maxilla can be lifted up from the nares, exposing the tissue in the maxilla. We sampled as much tissue as possible in this part of the maxilla, and tried to sample from all three regions (anterior, middle, posterior) of the conchae, and placed immediately in microcentrifuge tubes on dry ice. Following sample collection, samples were stored in -80 C freezers. We obtained pectoralis muscle tissue at the same time, following olfactory epithelium sampling.

We obtained olfactory epithelia from three bird species: *Gallus gallus* (chicken), *Taeniopygia guttata* (zebra finch), and *Molothrus ater* (cowbird). In total, we obtained three OE and pectoralis samples from each species. R.J.D. sampled the chickens immediately following a routine dispatch in the laboratory of Dr. Ken Anderson at the Prestage Department of Poultry Science at North Carolina State University. The chickens were 21-week old hyline W-36 white leghorn hens (female). R.J.D. collected the zebra finch samples from the laboratory of Dr. Richard Mooney in the Department of Neurobiology at the Duke University School of Medicine. All zebra finches were adult females from separate parents. Dr. Marc Schmidt at the Department of Biology at the University of Pennsylvania collected and dissected the cowbirds. All brown-headed cowbirds were adult males. All four species were sampled from captive populations, including the domesticated chicken and zebra finch.

RNA extractions and sequencing

To extract RNA from the olfactory epithelium and pectoralis tissue, we cut a small amount of tissue (roughly 2x2 cm) from each sample, and cut samples on dry ice. We immediately transferred tissue to 1mL RNAzol RT (Molecular Research Center, Inc., Cincinnati, OH) according to the manufacturer's brochure (March 2017), and dissolved the sample with a homogenizer. We then added 400uL water to DNA, protein, and polysaccharides, and then waited 15 minutes to precipitate. We centrifuged to remove these at 12,000 g for 15 minutes. We next added 5uL 4-bromoanisole to 1mL of supernatant for phase separation, waited 3-4 minutes, and then centrifuged at 12,000 g for 10 minutes. We performed this optional step of the protocol twice. We then precipitated the isolated RNA by adding equal volume isopropanol to the supernatant, waited 15 minutes, and then centrifuged 12,000 g for 10 minutes. We then washed

with 400uL 75% ethanol and spun at 4,000 g for 3 minutes, and repeated this step twice. We then solubilized in water. We tested RNA concentration and purity using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and RNA quality and integrity were assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) at the Brody Integrative Genomics Core in the Department of Pathology & Laboratory medicine at East Carolina University.

We examined RNA quality with the 4200 TapeStation (Agilent Technologies, Santa Clara, CA), with RNA integrity number (RIN) of samples ranged from 6 to 10. We determined RNA concentration with the Qubit Fluorometric Quantitation (Thermo Fisher, Waltham, MA), with 150 ng of RNA samples used for each NGS library preparation. We prepared stranded cDNA libraries using the TruSeq Stranded LT mRNA kit (Illumina, San Diego, CA) in accordance with the manufacturer's protocol using the poly-adenylated RNA isolation. We performed sequencing of paired-end reads (100 bp × 2) by pooling all the samples together on the NextSeq 2000 system with a P3 200 cycles reagent. We de-multiplexed and trimmed raw sequence reads for adapters the on-instrument DRAGEN GenerateFastQ pipeline (v3.7.4).

Read mapping

We mapped reads using the Spliced Transcripts Alignment to a Reference (STAR) aligner[66]. We were interested in OR expression specifically, so we generated the STAR reference genome not from the available species genome assemblies, but from our previously established genomic OR repertoires of each species. Additionally, not all ORs are annotated in previously published assemblies. We found the genomic OR repertoires for chicken, zebra finch, and cowbird from our previously described genomic scans. From our final curated OR alignments, we used custom

R scripts and bedtools to extract nucleotides from the associated genome[67]. We generated the reference genome of OR sequences without using a GTF reference annotation. We then mapped reads to the genomic OR repertoires using STAR default parameters. We also mapped all other genes back to each species reference genome to find the expression of all genes in the OE and pectoralis samples.

Counting and differential expression

We counted the number of reads in output SAM files using the *dplyr* package in R[68]. To measure gene expression, we converted raw counts to fragments per kilobase of transcript per million (FPKM). We did not filter genes with low expression due to previous reports of many bird ORs showing low expression levels[39,40]. We used a standard linear model with "tissue" (either pectoralis "PEC" or olfactory epithelium "OE") as the independent variable testing within *Gallus gallus*, *Taeniopygia guttata*, and *Molothrus ater*. We ran a student's paired t-test comparing log-transformed FPKM values between OE and pectoralis samples for chicken, zebra finch, and cowbird, as an alternative way to measure differential expression from a relatively small number of overall genes. For mapping to phylogenetic trees, we used trees created as described previously, using maximum likelihood methods in IQ-TREE[55]. We overlayed expression heatmap plots to the phylogeny using the gheatmap function in *ggtree* in R[69].

Olfactory epithelium sectioning

We used three female *Gallus gallus* Hyline W-36 white leghorns at 60 weeks of age. Following sacrifice and decapitation, we transported chicken heads on ice and dissected the chicken by cutting the maxilla transversely and collecting tissue within. We obtained chicken turbinate tissue

at the posterior or proximal end of the maxilla. We then embedded the dissected turbinate in optimal cutting temperature (OCT) compound histology mold. We stored embedded tissue in liquid nitrogen before transferring to -80 C. We then used a Leica CM 1850 Cryostat to section 18um tissue sections onto VWR superfrost frosted adhesion slides. We stored slides at -80 C.

OR probe preparation: PCR amplification of DNA template

Using chicken genomic DNA as template, we followed steps outlined previously[70]. We set up a PCR reaction (Table M1) using primers designed to amplify chicken ORs from the alpha, gamma, and gamma-c subfamilies (Supp Mat.) and ran for 25 cycles (Table M2). We then cleaned up amplified DNA template using the MinElute columns in the Qiagen MinElute PCR purification kit [Qiagen #28004], which yields a final elution in 10uL of EB buffer (10 mM Tris-HCl, pH 8.5). For purification, we first added 200uL PB buffer to the PCR reaction and mix thoroughly. Then, we transferred the mixture to MinElute columns and centrifuged at full speed for 30 seconds. We then washed columns with 750uL PE buffer, spun for 30 seconds at maximum speed in a microcentrifuge, and discarded flow-through. Finally, we centrifuged for two minutes at maximum speed to remove any residual PE buffer. We then transferred the MinElute column to a 1.5mL microcentrifuge tube. We then added 10uL of elution buffer directly to the center of the MinElute column and centrifuged for 2 minutes to elute the PCR product. We then loaded 1uL of the purified DNA template on a 1% agarose gel and ran gel electrophoresis to check for the recovery of appropriately sized PCR product. To be permissible, a strong and distinct band is necessary.

Probe synthesis, hydrolysis, and cleanup

We then set up the transcription reaction (Table 2M), and incubated for 120 minutes at 37C. We then prepared alkaline buffer (80mM NaHCO3, 120mM Na2CO3), making 150uL total, or 12uL of 1M NaHCO3, 18uL of 1 M Na2CO3, in 120uL of nuclease-free water. We then added 25uL alkaline buffer to the transcription reaction and incubated at 60C for 15 minutes. We then purified the reaction using RNase-free riboprobe purification columns in the Micro Bio-spin 30 chromatography column (Cat #732-6223, Bio-Rad Laboratories, Inc., USA). Prior to using the columns, we inverted columns multiple times to mix the Bio-gel resin and remove bubbles. We then removed the column's bottom tip and place in a collection tube. We next spun down the resin in the column at 3400rpm for 2 minutes and then place the column into a 1.5mL microcentrifuge tube. We pipetted the RNA probe solution directly onto the resin and then centrifuged at 3400rpm for 4 minutes. Following elution, we added 35uL UltraPure Distilled Formamide (Cat #15515-026, Life Technologies Corp., CA, USA) to the probe. We then ran 5uL of the eluted probe in a 1% agarose gel electrophoresis. The correct sized products appeared as a fuzzy band between 100 and 200bp. We then stored probes at -80C.

RNA fluorescent in-situ hybridization (FISH) and Immunofluorescent (IF) Staining

We first removed the slides with sections from the -80 C and placed on a clean flat plastic tray cleaned with 70% ethanol. We then took a hairdryer and blow-dried the slides until dry. We then loaded on a slide rack and dipped the slides in a staining bucket of 4% Paraformaldehyde (Cat#S898-07, Avantor Performance Materials LLC., PA, USA) for 15 minutes. Next we washed slides in a bucket of 1X PBS for 5 minutes, and repeated this step for a total of two washes. Meanwhile, we prepared a triethanolamine solution in a 1L glass beaker, gently mixing 8.2mL Triethanolamine (Cat # 9468-01, Baker Analyzed, Avantor Performance Materials, Inc., PA,

USA) with 800mL dH2O using a magnetic stir bar. We dipped the slide rack into the solution while the stir bar was spinning. We submerged the slides for 10 minutes, and during the first minute, we added 1.75mL Acetic Anhydride (Cat# A-6404, Sigma Chemical Co., MO, USA) dropwise while stirring. Following this submersion, we then wash the slide rack in a new bucket of 1X PBS for 5 minutes.

We then set up slide holders using a 150mm diameter Petri dish with blotting paper on the bottom, and two 1mL serological pipettes placed on top for holding slides. We soak the blotting paper in 5x saline sodium citrate buffer (SSC) (Cat # AB131156, American Bioanalytical, MS, USA) with 50% UltraPure Distilled Formamide (Cat #15515-026, Life Technologies Corp., CA, USA). We then dried the slides using blotting paper and placed the slides on the slide holders. We then added 500uL of prehybridization buffer and incubated for at least one hour at 58C. We then added the DIG probe to new prehybridization buffer at a concentration of 1uL per 200uL of buffer and heat at 85C for 5 minutes. We then added 200uL of the probe and prehybridization buffer mix (50% formamide (Invitrogen 15515-026), 5xSSC (American Bioanalytical AB13156), Baker's yeast RNA (Sigma R-6750, 250ug/ml), Herring or Salmon sperm DNA (Sigma D7290 or D1626, respectively, P/C 100ug/ml), 1mM DTT, heparin (Sigma H3393, 300U/ml)) to each slide in the slide holder. We covered the slides with strips of parafilm to form a coverslip and incubated overnight at 58C.

The next day, we warmed 5X Saline Sodium Citrate buffer (SSC) (Cat # AB13156, American Bioanalytical, MA, USA) and 0.2X SSC to 72°C in an incubator. We then made a 5% blocking solution stock, of 10g blocking reagent (Cat #11096176001, ROCHE Diagnostics GmbH, Mannheim, Germany) in 200mL maleic acid buffer and stored at 4C. We then diluted

this solution to 0.5% blocking solution in 1X PBS, and dilute 7.5% Bovine Serum Albumin (BSA) stock (Cat # 15260-037, Gibco, Thermo Fischer Scientific, MA, USA) to a 0.1% BSA solution in 1X PBS. We prepared a working solution of DIG-POD antibody by making a 1:1000 dilution of Anti-Digoxigenin-POD, Fab fragments (Cat# 11207733910, ROCHE Giagnostics GmbH, Mannheim, Germany), in the 0.5% blocking solution.

We filled two buckets with the 5X SSC and two with 0.2X SSC. We then took the slides and dry with blotting paper, and dip slides into the warm 5X SSC. The parafilm coverslip should fall off in the buffer, but if it did not, we removed with forceps. We then transferred the slides to a second bucket of 5X SSC and then moved to a bucket with the warm 0.2X SSC at 72C for 30 minutes. We repeated this step in a second bucket of 0.2X SSC at 72C for 30 minutes. We then moved slides to a wash step with a bucket of 1X PBS for 5 minutes at room temperature. We then moved slides from the slide rack to a slide mailer containing 0.5% blocking solution and incubate for 30 minutes. We then removed slides from the mailer and use blotting paper to removed excess solution. We then placed each slide back in the slide holder, with 300uL of antibody placed on top of each slide. We incubated the slide holders 45 minutes at room temperature. We then removed the antibody solution with blotting paper, placed slides back in mailers, and rinse twice with 1X PBS. Following rinses, we then washed with 1X PBS in the slide mailer three times, each for 10 minutes. We then placed the slides in 0.1% BSA solution and prepared TSA working solution. The recipe for TSA working solution is 0.003% H2O2 and a 1:400 dilution of FITC-tyramide stock in in 1X PBS. The FITC-tyramide is prepared following previous protocols[71,72]. Tyramide-FITC was generated using fluorescein-NHS ester (Cat # 46100, Pierce), tyramide (Cat # T-2879, Sigma), Dimethyl formamide (DMF) (Cat # T-8654), and triethyl amine (TEA) (Cat # T-0886, Sigma), prepared by by mixing 4 ml FITC NHS in DMF

and 1.37 ml tyramide solution and incubated in the dark in RT for 2 hrs, then added 4.6 ml ethanol. We diluted the H2O2 from an original 30% H2O2 stock in 1X PBS. After this, we removed slides from the slide mailer and removed excess fluid using blotting paper. We then add 300uL of the TSA working solution per slide and incubate for 10 minutes at room temperature in the dark. We then carefully TSA working solution using blotting paper and placed slides back to in the slide mailer. We then rinsed slides twice with 1X PBS and then wash with 1X PBS twice for 5 minutes each. If only FISH is required, we moved to the counterstaining step. If double labeling with phosphorylated S6 (pS6) antibody is required with continue with Immunofluorescent (IF) staining.

For double-labeling with the pS6 antibody, we blocked FISH labeled sections (0.1% Triton-X, 5% skim milk, in PBS) for 30 minutes, then incubated with anti-phospho-S6 (244/247) 1:300 (Cat # 44-923G, Thermo Fisher Scientific, MA, USA) diluted in blocking solution overnight at 4°C. After washing with PBS, we incubated sections with the donkey Cy3-conjugated anti-rabbit IgG 1:200 (Jackson ImmunoResearch Laboratories, PA, USA) for 45 min, washed again, and counterstained the nuclei.

For counterstaining, we replaced the 1X PBS with nuclei staining solution. To make the staining solution, we made a solution of 1% or 25uL Hoechst nuclear stain (bisbenzimide H33258, Cat # B2883, SIGMA-ALDRICH Co. St.Louis, USA), in 250mL 1X PBS, and incubated in the dark for 5 minutes. We then discard the nuclei staining solution and washed slides twice with 1X PBS for 5 minutes. We then have a final rinse with dH2O. We finally dried the excess liquid using blotting paper, and then we sealed the slides with Mowiol mounting media and a coverslip, and stored them in a slide box at 4°C.

Imaging

Images were acquired as z-stacks of optical sections at 1 μm intervals using a Zeiss Axio Observer Z1 inverted microscope equipped with a Zeiss Axiocam MRm camera at 200× or 400× magnification. The orthogonal projection function was used to generate composite images for visualization and quantification purposes.

Quantification

For quantification of percent of probe-positive cells, ImageJ2, version 2.9.0/1.53t was used. We quantified the total cell number per field of view (FOV) in the blue channel, and the probe-positive cells in the green channel, to calculate the percent of expression of probe positive divided by total cells counted in the FOV. Data processing and statistical analysis were performed in Python (version 3.11) using standard scientific libraries including NumPy, Pandas, and Matplotlib. The mean and standard error of the mean (SEM) were calculated for each experimental group using $SEM=SD/\sqrt{n}$, where SD is the standard deviation and n is the number of biological replicates. Graphical visualization of the quantified ISH data, including bar plots with mean ± SEM, was generated using Matplotlib. The plotting code was custom written in Python and applied uniformly to all datasets to ensure consistent normalization and scaling.
For pS6 signal quantification, the pS6 pixel intensity was quantified in individual cells labeled with the mRNA probe (OR-expressing) using ImageJ2, version 2.9.0/1.53t, and normalized values were analyzed. Two-Way ANOVA, followed by Dunnett's post-hoc test, statistical analysis and plotting was performed using GraphPad Prism 10 Version 10.4.1 (532).

Design of consensus OR sequence

To screen for many ORs at once initially, we designed consensus OR sequences representing the sequences of multiple ORs. For example, in the chicken, we took the amino acid sequences of all 419 chicken gamma-c ORs recovered from the genomic repertoire and aligned these ORs using the E-INS-I default parameters in MAFFT[51]. We then selected the most common amino acid residue at each position in the alignment, and created a "consensus OR" of all of the most common residues from the 5' to the 3' end of the OR ORF. The stop codon was assigned at the first consensus stop codon in the alignment. In the infrequent case that multiple residues were equally common at a single position in the alignment, we picked the amino acid at random. We trimmed all alignments at the 5' end prior to the first consensus methionine, and trimmed all alignments at the 3' end following the first consensus stop codon. We removed all residues corresponding to an insertion that was not present in the majority of ORs, or a deletion that was present in the majority of ORs. For subclade consensus ORs, we aligned subsets of chicken gamma-c ORs, and created a consensus sequence from this OR subgroup.

OR cloning procedure

We designed primers for OR amplification (see Supp. Mat., File 9) based on the OR 5' and 3' sequence. We designed primers to have an estimated denaturation temperature of 56 C to 58 C, or roughly 18-22 nucleotides from the 5' or 3' end of the OR. We then added a 5' forward MluI restriction enzyme linker *AAACGCGT) and a 3' reverse NotI linker (TTGCGGCCGC). We dissolved primers in Gibco water to 100uM and then made a 5uM working stock. For OR consensus sequences, we first centrifuged the tube, then added 100uL of TE buffer to reach a final concentration of 10 ng/uL. We then vortexed and incubated at 50 C for 20 minutes. We then created a PCR mix (Table M3) using the primer set, and ran for 25 cycles (Table M4). For

consensus ORs, template DNA was the hydrated synthesized OR, for native ORs, the template DNA was the genomic DNA from chicken skeletal muscle (Zyagen #GC-314). Following amplification, we ran 1uL of product on a 1.5% agarose gel at 100V for 20 minutes, to confirm the amplified product. We then performed a purifying step and added 200uL of PB buffer (Qiagen) directly to the PCR strip tube. We then mixed the PB buffer and the PCR product and transferred to a Qiagen MinElute column and collection tube (Qiagen), followed by a 30 second spin at 15,000 RPM, and discarded the flow through. We then added 750uL of PE buffer followed by a 30 second spin, and then a 2 minute dry spin, discarding the flow through each time. We then added 10uL of elution buffer directly to the column and transferred the column to a 1.5mL microcentrifuge tube. We then spun the column down for 1 minute. For restriction enzyme digestion, we used 100ng/uL concentration of vector DNA, and added reagents (Table M5), including MluI and NotI high fidelity enzymes (New England Biolabs), and digested for 20 minutes at 37C. Following restriction enzyme digestion, we then purified the digested product. We added 200uL PB buffer to the PCR strip tubes, mixed thoroughly, and then transferred the mixture to the same MinElute column as before. We then spun for 30 seconds at 15,000 RPM and discarded the flow-through. We then added 750uL or 3.66M guanidine hydrochloride aqueous solution, made from 35g of guanidine hydrochloride in 100mL Gibco water. We then spun for 30 seconds and discarded the flow-through. We then added 750uL PE buffer (Qiagen) and spun for 30 seconds. Then, a second round of 750uL PE buffer, spinning for 30 seconds, followed by a 2 minute dry spin, discarding flow-through with each spin. We then transferred the MinElute spin column to a 1.5mL microcentrifuge tube, and added 10uL of elution buffer directly to the column, and spun for 1 minute. We then used the purified digest to set up a ligation reaction (Table M6), and held tubes at room temperature for over 1 hour. We ligated

inserts into the rhodopsin tagged pCI vector or a modified Lucy-tagged, FLAG-tagged, and rhodopsin-tagged pCI vector, which contains MluI and NotI restriction enzyme sites, a Rho tag, as well as ampicillin and carbencillin resistance.

We then transformed *E. coli* cells with the plasmid with the ligated insert. We first thawed 20uL of competent cells and added 2.5uL of the ligated product. We held the incubating cells on ice for over 10 minutes, and then pipetted the cells onto an LB-ampicillin or LB-carbenicillin plate. We then let the colonies grow for 37 C overnight. To check for successful insertion of DNA into plasmid and transformation, we picked colonies and soaked in 20uL of Gibco water. We then performed colony PCR using primers designed for pCI (pCI 5'A: CTCCACAGGTGTCCACTC, pCI 3'A: CACTGCATTCTAGTTGTGG, Table M7) and ran for 25 cycles (Table M8). We then ran 1uL of PCR product with 5uL 1.2X loading dye (New England Biolabs) and checked for a band of the appropriate size on a 1% gel. We then added 10uL of *E. coli* with the appropriate insert to 4.5mL 2XYT-ampicillin or 2XYT- carbenicillin (100ug/mL) in a 17x100mm culture tube and shook bacteria for 18 hours at 37C. We then stored remaining volume from the colonies with the correct insert in 15-40% glycerol at -80C. Following incubation, we spun down the liquid culture for 30 seconds at 15,000 RPM twice in a 2mL tube and discarded the supernatant, retaining only the bacteria pellet. We then miniprepped plasmids with the ZymoPURE Plasmid Miniprep kit (Cat. # D4212, Zymogen, USA).

Addition of Givaudan sequence to OR

One challenge with studying native ORs are their low expression in heterologous cells. Comparisons in odor ligand selectivity and sensitivity between native ORs and consensus ORs are made difficult by low native OR expression that led to low native OR response. Previous

studies have utilized the addition or modification of N-terminal and C-terminal tags to improve OR trafficking from the ER to the cell surface[73]. Studies have shown that basic amino acid residues, such as arginine and lysine placed at the C-terminus is beneficial for cell surface expression[74]. Here, we engineered chicken native gamma-c OR receptors to investigate the potential improvement in the native OR receptors response.

We engineered ORs to modify their C-terminal amino acid sequences by replacing them with the basic Givaudan sequence RNKEVKKAIKRLFKRKCCRRR. To achieve this, we first obtained the nucleotide sequences of native ORs through DNA sequencing. Next, we designed primers to bridge the C-terminal region of the native ORs with the Givaudan sequence[29]. Native ORs featured a conserved landmark sequence, NPXIYXXRN, on their C-terminus. The primer was specifically designed as a reverse complement to the junction between the native OR sequence and the Givaudan sequence. In this primer, the 3′ end included the reverse complement of the nucleotide sequence corresponding to the NPXIYXXRN region of the native OR. These nucleotides were selected to achieve a denaturation temperature of 56°C or 58°C, with adenine (A) and thymine (T) contributing 2°C each and cytosine (C) and guanine (G) contributing 4°C each. Preceding these nucleotides, the primer included the sequence GATCGCTTTTTTCACTTCTTT, which encodes the reverse complement of the Givaudan amino acid sequence beyond the overlapping RN landmark. After obtaining OR specific bridging primers, we proceeded with OR cloning to achieve purified plasmid (previously described in STAR methods).

In vitro cell culture

Preparation of stock Hana3A cell line

We kept Hana3A cell stocks frozen in -80 C freezers and followed a previously described protocol for cell preparation[75]. We thawed frozen cells in 1mL tubes in a 37 C water bath. After thawing, we transferred the entire stock to 6mL Minimum Essential Medium Eagle (MEM) (Cat # 10-010CV, Corning, Mediatech. Inc., VA, USA) supplemented with 10% Fetal Bovine Serum (FBS) (M10 solution) in a 15mL conical tube. We then centrifuged at 1000 rpm at room temperature and collected Hana3A cells in a pellet. We then aspirated the M10, leaving the pellet, and resuspended cells with 10mL M10 with and 0.5% penicillin-streptomycin (Gibco) and 0.5% amphotericin B (Gibco) on a 100-mm cell culture plate. We then cultured cells in a 37 C incubator with a water bath and 5% CO2 overnight. The next day, we checked cells with a phase contrast microscope to check for cell health. Ideal health has cells spread out uniformly across the surface of the culture dish and with no signs of contamination. We then replace medium with fresh M10 with penicillin-streptomycin and amphotericin medium to maintain cell health.

Over the next few days, we checked cell density (Hana3A cell line divides roughly once per 24 hours) and cell health periodically using a phase contrast microscope. We waited until cells reached the desired confluency and then we were ready to passage cells to 96 well plates. First, we tilted the dish to aspirate all media from the cell culture dish, and then we pipette 10mL PBS onto cells to wash. We then aspirated the PBS from the edge of the dish, and then added 3mL 0.05% trypsin-EDTA to the cells to detach them from the bottom of the dish. To assist in detaching, we gently shake the dish to detach cells from the bottom. This gentle shaking lasts around 2 minutes, and can be observed both by eye and under a phase contrast microscope. Immediately after detaching, we neutralized the trypsin-EDTA by the addition of 5mL M10 to the dish. We then dissociate the cells from forming clumps by pipetting all 8mL up and down onto the plate. We then transfer the cells and media to a 15mL conical tube, and centrifuge for 5

minutes at 200g at room temperature. Following spin down, we aspirated all media (M10 and trypsin-EDTA), without disturbing the cell pellet. We then resuspend cells in 1mL M10 to both seed 96-well plates and for the maintenance of the cell line.

For cell line maintenance, we transfer 10mL M10 with penicillin-streptomycin and amphotericin to a new culture plate. We then transfer the desired amount of cells to continue the cell population onto the plate, and place in a 37C incubator with 5% CO2 for future experiments. For cells to be used for transfection, we transferred to 96-well plates. The amount of cell transferred to 96 well plates should remain consistent, with roughly 10% of cells on a 100% cell covered 35x10mm cell culture dish per each 96-well plate used. So, for example, if the surface of a culture plate was covered 100% in cells, and an assay used six 96-well plates, 600uL (60%) of the cells from the culture plate resuspended in 1mL M10 would be used for 96-well plate seeding.

We then added 5mL M10 (without penicillin-streptomycin and amphotericin) per 96-well plate (Corning) used to a reagent reservoir (for example, an experimental design with six 96-well plates would use 30mL M10) and added the appropriate quantity of cells to the reservoir as well. We additionally added 60uL poly-D-lysine to the reservoir per 96-well plate (for example, 360uL poly-D-lysine for a design with six 96-well plates). We then mixed all of cells, media, and reagents. Using a multichannel pipettor, we then pipetted 50uL of the mixture from the reservoir to the each well of the 96-well plate. We then cultured the cells in the 96-well plate(s) for 24 hours in the 37C incubator with 5% CO2.

Transfection of OR plasmid

Prior to transfection, we prepared all necessary plasmids through the ZymoPURE Plasmid Miniprep kit (Cat. # D4212, Zymogen, USA). Necessary plasmids include ORs of interest, RTP1S to assist in membrane transport[76], pGloSensor-20F (Cat # E1171, Promega Co., WI, USA) to encode for the cAMP-responsive element, and empty pCI mammalian expression vector (Cat # E1731, Promega Co., WI, USA) as a negative control. Rho-tagged ORs of interests are located in the pCI vector. We then observed the cells to be transfected for normal shape distribution, and a confluency of 30-50%. We then prepare two mixtures for each 96-well plate, first we prepared a DNA transfection mixture of 500uL MEM, 10uL pGLO, and 5uL RTP1S. After this is distributed, ORs are added to a final concentration of 100ng/uL per 96-well plate. We then add the second mixture of 500uL MEM with 20uL Lipofectamine 2000 (Cat # 11668019, Thermo Fisher Scientific Inc., MA, USA). Once the plasmids and Lipofectamine mixture are added together, we then incubate at room temperature for 15 minutes. Then for each 96-well plate, we added 5mL M10 to the mixture.

With the mixture prepared, we then removed the 96-well plates from the incubator and gently tapped the plates upside-down on sterile paper towels so that M10 from the previous day is removed from wells. Using a multichannel pipettor, we then added 50uL of the transfection mixture to each well. We then incubated cells with transfected plasmids for 24 hours in the 37C incubator with 5% CO2.

Loading glo sensor stimulation buffer

First, we observed cells under the microscope to check for roughly 50%-80% confluence and a healthy appearance. We then added 10mM HEPES and 15mM NaN3 to HBSS to later add to the stimulation buffer. For a 500mL bottle of HBSS, we added 5mL 1M HEPES and 5mL 1.5M

NaN3. We then removed GloSensor cAMP Reagent (glo green) (Cat # E1291, Promega, WI, USA) from -80C storage and warmed to room temperature. We added 75uL glo green reagent to 2.76mL HBSS to create a stimulation buffer. We then removed the transfection media mixture from the cells by gently inverting and tapping 96-well plates onto sterile paper towels in a safety cabinet. We then used a multichannel pipettor to distribute 25uL of the stimulation buffer to each well. We then placed the 96-well plate at room temperature covered in aluminum foil in a drawer for 2 hours. Glo green reagent is light sensitive and we kept plates in a dark, odorless environment.

During incubation, we prepared odorants. Initially, we diluted odorants to 100mM concentrations in 95% ethanol and stored at -20C. We then added odors to the stimulation buffer at the desired concentration for odor exposure to ORs. For example, we would serially dilute 100mM working solutions to concentrations of 300uM, 100uM, 30uM, 10uM, 3uM, 1uM, in continuum, at the time of stimulation. Each OR in each experiment also received a no odor negative control, with only stimulation medium, to assess background cAMP activity levels. For odorants with sulfur, we added 30uM CuCl2 to the odorant stimulation buffer mixture.

Following incubation, we transferred each 96-well plate to the CLARIOstar Plus multi-mode plate reader (BMG LABTECH). We then take a baseline measurement of fluorescence activity of the cells prior to the addition of odorants. Following the blank measurement, using a multichannel pipettor, we distributed 25uL of the odorant and stimulation medium mixture to each well in the 96-well plate. Immediately following odorant exposure, we placed the 96-well plate back in the CLARIOstar plate reader and began reading cAMP activity for 15 minutes, at 10 cycles, or one cycle per 90 seconds.

Data analysis

We analyzed plate reader data using custom python scripts. We normalized fluorescence activity to the blank zero time point measurement prior to odor stimulation, and subtracted this value by 10 to allow for 0 to signify a lack of response. We analyzed dose-response curves to ORs by fitting a least squares function to the data in GraphPrism 10 Version 10.4.1 (532).

OR cell surface expression

Flow cytometry was performed to assess the cell-surface expression of ORs. HEK293T cells were plated in 6 well plate (Corning) at a density of approximately 3.5×10^5 cells (2.5% confluency per well) and cultured overnight. After 18–24 hours, OR plasmids (1,000 ng), which were N-terminally tagged with the first 20 amino acids of human rhodopsin (rho-tag) in the pCI mammalian expression vector (Promega), were transfected along with 200 ng of RTP1S and 10 ng of eGFP using Lipofectamine 2000 (Cat # 11668019, Thermo Fisher Scientific, MA, USA).

At 18–24 hours post-transfection, cells were detached using Cell Stripper (Cat # 25-056-CI, Corning, Mediatech Inc., VA, USA) and resuspended in ice-cold Phosphate Buffered Saline (PBS) (Cat # SH30256.01, Cytiva, HyClone Laboratories, UT, USA) supplemented with 15 mM Sodium Azide (NaN_3) (Cat # S2002, Sigma-Aldrich, MO, USA) and 2% Fetal Bovine Serum (FBS) (Cat # SH3008802HI, Cytiva, HyClone Laboratories, UT, USA). The cell suspension was transferred to 5 ml round-bottom polystyrene tubes (BD), centrifuged at 4□, and resuspended again in PBS containing 15 mM NaN_3 and 2% FBS. Cells were then incubated with a primary antibody (1/400 dilution, mouse anti-rhodopsin clone 4D2, MABN15, Sigma-Aldrich, MO, USA) for 30 minutes, followed by washing with PBS containing 15 mM NaN_3 and 2% FBS.

After another centrifugation step, cells were stained with a secondary antibody (1/200 dilution, phycoerythrin-conjugated donkey anti-mouse F(ab′)2 fragment, 715-116-150, Jackson Immunologicals) for 30 minutes in the dark. To distinguish dead cells, 7-amino-actinomycin D (1/500 dilution, 129935, Calbiochem) was added. The samples were immediately analyzed using a BD FACSCanto II flow cytometer, with gating applied to select GFP-positive, single, spherical, and viable cells. Phycoerythrin fluorescence intensities were quantified and visualized using FlowJo v10.8.1. An empty pCI plasmid served as the negative control.

Live chicken exposure to pyrazines

We collected 64 week old white leghorn chickens and placed in boxes with airflow through an external vent. We allowed chickens to be in the box for one hour to acclimate the chickens from any outside odors they may have encountered prior to the experiment. Following one hour acclimation, we placed odors on blotting paper that was then placed inside a Sakura Tissue-Tek Uni-cassette system. Odors added to the cassette were either 1% acetophenone, 1% 2-methoxy-3-isobutylpyrazine, or 10% 2-methoxy-3-isobutylpyrazine, diluted in ethanol. No odor control chickens received blotting paper and cassette with no odor added. We then placed cassettes on the floor of the boxes with the chickens. Chickens were standing on a grated floor at the bottom of the box, and we placed cassettes below the grated floor to prevent chickens from interfering with cassettes.

Following one hour of odor exposure, we then sacrificed chickens. We then dissected chickens were to obtain posterior olfactory epithelium tissue. We embedded posterior olfactory epithelium in OCT medium, and flash froze in liquid nitrogen prior to transfer to -80C. We then cut 18uM sections of the posterior olfactory epithelium using a Leica CM 1850 Cryostat and placed on VWR superfrost slides, and stored at -80C.

Citations

1. Darwin, C. (1891). Journal of researches into the natural history and ecology of the countries visited during the voyage of H.M.S. Beagle round the world, under the command of Capt. Fitz Roy, R.N. (London□: Ward, Lock & Co .).

2. Hill, Alex. (1905). Can birds smell? Nature *71*, 318–319. https://doi.org/10.1038/071318b0.

3. Audubon, J.J. (1826). Account of the habits of the turkey buzzard, Vultur aura, particularly with the view of exploding the opinion generally entertained of its extraordinary power of smelling . Edinburgh New Phil J . *2*.

4. Molina-Morales, M., Castro, J., Albaladejo, G., and Parejo, D. (2020). Precise cache detection by olfaction in a scatter-hoarder bird. Anim. Behav. *167*, 185–191. https://doi.org/10.1016/j.anbehav.2020.07.002.

5. Van Huynh, A., and Rice, A.M. (2021). Odor preferences in hybrid chickadees: implications for reproductive isolation and asymmetric introgression. Behav. Ecol. Sociobiol. *75*, 129. https://doi.org/10.1007/s00265-021-03069-2.

6. Bonadonna, F., and Gagliardo, A. (2021). Not only pigeons: avian olfactory navigation studied by satellite telemetry. Ethol. Ecol. Evol. *33*, 273–289. https://doi.org/10.1080/03949370.2021.1871967.

7. Wikelski, M., Quetting, M., Cheng, Y., Fiedler, W., Flack, A., Gagliardo, A., Salas, R., Zannoni, N., and Williams, J. (2021). Smell of green leaf volatiles attracts white storks to freshly cut meadows. Sci. Rep. *11*, 12912. https://doi.org/10.1038/s41598-021-92073-7.

8. Owre, O.T., and Northington, P.O. (1961). Indication of the sense of smell in the turkey vulture, Cathartes aura (Linnaeus), from feeding tests. Amer. Midl. Naturalist *66*, 200–205. https://doi.org/10.2307/2422877.

9. Grubb, T.C. (1972). Smell and foraging in shearwaters and petrels. Nature *237*, 404–405. https://doi.org/10.1038/237404a0.

10. Buck, L., and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell *65*, 175–187. https://doi.org/10.1016/0092-8674(91)90418-x.

11. Strotmann, J., Wanner, I., Krieger, J., Raming, K., and Breer, H. (1992). Expression of odorant receptors in spatially restricted subsets of chemosensory neurones. Neuroreport *3*, 1053–1056. https://doi.org/10.1097/00001756-199212000-00005.

12. Policarpo, M., Baldwin, M.W., Casane, D., and Salzburger, W. (2024). Diversity and evolution of the vertebrate chemoreceptor gene repertoire. Nat. Commun. *15*, 1421. https://doi.org/10.1038/s41467-024-45500-y.

13. Christmas, M.J., Kaplow, I.M., Genereux, D.P., Dong, M.X., Hughes, G.M., Li, X., Sullivan, P.F., Hindle, A.G., Andrews, G., Armstrong, J.C., et al. (2023). Evolutionary constraint and innovation across hundreds of placental mammals. Science *380*, eabn3943. https://doi.org/10.1126/science.abn3943.

14. Driver, R.J., and Balakrishnan, C.N. (2021). Highly Contiguous Genomes Improve the Understanding of Avian Olfactory Receptor Repertoires. Integr. Comp. Biol. *61*, 1281–1290. https://doi.org/10.1093/icb/icab150.

15. Creece, D., Freire, R., and Massaro, M. (2025). Past research and future directions in understanding how birds use their sense of smell. Ibis, 1–29. https://doi.org/10.1111/ibi.13398.

16. Glusman, G., Bahar, A., Sharon, D., Pilpel, Y., White, J., and Lancet, D. (2000). The olfactory receptor gene superfamily: data mining, classification, and nomenclature. Mammalian Genome *11*, 1016–1023. https://doi.org/10.1007/s003350010196.

17. Niimura, Y., and Nei, M. (2005). Evolutionary dynamics of olfactory receptor genes in fishes and tetrapods. Proc. Natl. Acad. Sci. U.S.A. *102*, 6039–6044. https://doi.org/10.1073/pnas.0501922102.

18. Steiger, S.S., Fidler, A.E., Valcu, M., and Kempenaers, B. (2008). Avian olfactory receptor gene repertoires: evidence for a well-developed sense of smell in birds? Proc. R. Soc. B *275*, 2309–2317. https://doi.org/10.1098/rspb.2008.0607.

19. Zhang, X., Zhang, X., and Firestein, S. (2007). Comparative genomics of odorant and pheromone receptor genes in rodents. Genomics *89*, 441–450. https://doi.org/10.1016/j.ygeno.2007.01.002.

20. Liao, D. (1999). Concerted Evolution: Molecular Mechanism and Biological Implications. The American Journal of Human Genetics *64*, 24–30. https://doi.org/10.1086/302221.

21. Minias, P. (2024). Evolutionary variation in gene conversion at the avian MHC is explained by fluctuating selection, gene copy numbers and life history. Molecular Ecology *33*, e17453. https://doi.org/10.1111/mec.17453.

22. Alcaide, M., and Edwards, S.V. (2011). Molecular Evolution of the Toll-Like Receptor Multigene Family in Birds. Molecular Biology and Evolution *28*, 1703–1715. https://doi.org/10.1093/molbev/msq351.

23. Mansai, S.P., and Innan, H. (2010). The Power of the Methods for Detecting Interlocus Gene Conversion. Genetics *184*, 517–527. https://doi.org/10.1534/genetics.109.111161.

24. SAWYER, S. (1999). GENECONV□: a computer package for the statistical detection of gene conversion. http://www.math.wustl.edu/∽sawyer.

25. Huang, Z., Xu, Z., Bai, H., Huang, Y., Kang, N., Ding, X., Liu, J., Luo, H., Yang, C., Chen, W., et al. (2023). Evolutionary analysis of a complete chicken genome. Proc Natl Acad Sci U S A *120*, e2216641120. https://doi.org/10.1073/pnas.2216641120.

26. Andolfatto, P., and Nordborg, M. (1998). The Effect of Gene Conversion on Intralocus Associations. Genetics *148*, 1397–1399. https://doi.org/10.1093/genetics/148.3.1397.

27. Khan, I., Yang, Z., Maldonado, E., Li, C., Zhang, G., Gilbert, M.T.P., Jarvis, E.D., O'Brien, S.J., Johnson, W.E., and Antunes, A. (2015). Olfactory receptor subgenomes linked with broad ecological adaptations in Sauropsida. Mol Biol Evol *32*, 2832–2843. https://doi.org/10.1093/molbev/msv155.

28. Jiang, Y., Gong, N.N., Hu, X.S., Ni, M.J., Pasi, R., and Matsunami, H. (2015). Molecular profiling of activated olfactory neurons identifies odorant receptors for odors in vivo. Nat Neurosci *18*, 1446–1454. https://doi.org/10.1038/nn.4104.

29. Emter, R., Merillat, C., Buchli, F., Flachsmann, F., and Natsch, A. (2025). Decoding human olfaction by high heterologous expression of odorant receptors detecting signature odorants. Curr Biol *35*, 5252-5263.e4. https://doi.org/10.1016/j.cub.2025.09.041.

30. Billesbølle, C.B., de March, C.A., van der Velden, W.J.C., Ma, N., Tewari, J., del Torrent, C.L., Li, L., Faust, B., Vaidehi, N., Matsunami, H., et al. (2023). Structural basis of odorant recognition by a human odorant receptor. Nature *615*, 742–749. https://doi.org/10.1038/s41586-023-05798-y.

31. de March, C.A., Ma, N., Billesbølle, C.B., Tewari, J., Llinas del Torrent, C., van der Velden, W.J.C., Ojiro, I., Takayama, I., Faust, B., Li, L., et al. (2024). Engineered odorant receptors illuminate the basis of odour discrimination. Nature *635*, 499–508. https://doi.org/10.1038/s41586-024-08126-0.

32. Asakawa, M., Fukutani, Y., Savangsuksa, A., Noguchi, K., Matsunami, H., and Yohda, M. (2017). Modification of the response of olfactory receptors to acetophenone by CYP1a2. Sci Rep *7*, 10167. https://doi.org/10.1038/s41598-017-10862-5.

33. Cunningham, S.J., Castro, I., and Potter, M.A. (2009). The relative importance of olfaction and remote touch in prey detection by North Island brown kiwis. Anim. Behav. *78*, 899–905. https://doi.org/10.1016/j.anbehav.2009.07.015.

34. Corfield, J.R., Price, K., Iwaniuk, A.N., Gutierrez-Ibañez, C., Birkhead, T., and Wylie, D.R. (2015). Diversity in olfactory bulb size in birds reflects allometry, ecology, and phylogeny. Front. Neuroanat. *9*. https://doi.org/10.3389/fnana.2015.00102.

35. Corfield, J.R., Eisthen, H.L., Iwaniuk, A.N., and Parsons, S. (2014). Anatomical specializations for enhanced olfactory sensitivity in kiwi, Apteryx mantelli. Brain Behav. Evol. *84*, 214–226. https://doi.org/10.1159/000365564.

36. Yokosuka, M., Hagiwara, A., Saito, T.R., Tsukahara, N., Aoyama, M., Wakabayashi, Y., Sugita, S., and Ichikawa, M. (2009). Histological properties of the nasal cavity and olfactory

bulb of the Japanese jungle crow Corvus macrorhynchos. Chem. Senses *34*, 581–593. https://doi.org/10.1093/chemse/bjp040.

37. Kondoh, D., Nashimoto, M., Kanayama, S., Nakamuta, N., and Taniguchi, K. (2011). Ultrastructural and histochemical properties of the olfactory system in the Japanese jungle crow, Corvus macrorhynchos. J. Vet. Med. Sci. *73*, 1007–1014. https://doi.org/10.1292/jvms.10-0574.

38. Martinez, Q., Amson, E., and Laska, M. (2024). Does the number of functional olfactory receptor genes predict olfactory sensitivity and discrimination performance in mammals? J. Evol. Biol. *37*, 238–247. https://doi.org/10.1093/jeb/voae006.

39. Sin, S.Y.W., Cloutier, A., Nevitt, G., and Edwards, S.V. (2022). Olfactory receptor subgenome and expression in a highly olfactory procellariiform seabird. Genetics *220*, iyab210. https://doi.org/10.1093/genetics/iyab210.

40. Luo, H., Luo, S., Fang, W., Lin, Q., Chen, X., and Zhou, X. (2022). Genomic insight into the nocturnal adaptation of the black-crowned night heron (Nycticorax nycticorax). BMC Genomics *23*, 683. https://doi.org/10.1186/s12864-022-08904-y.

41. Leibovici, M., Lapointe, F., Aletta, P., and Lièvre, C.A.-L. (1996). Avian olfactory receptors: differentiation of olfactory neurons under normal and experimental conditions. Dev. Biol. *175*, 118–131. https://doi.org/10.1006/dbio.1996.0100.

42. Lalloué, F.L., Ayer-Le Lièvre, C.S., and Sicard, G. (2003). Analysis of the functional maturation of olfactory neurons in chicks before and after birth. Chem. Senses *28*, 729–737. https://doi.org/10.1093/chemse/bjg055.

43. Hess, C.M., and Edwards, S.V. (2002). The Evolution of the Major Histocompatibility Complex in Birds: Scaling up and taking a genomic approach to the major histo compatibilty complex (MHC) of birds reveals surprising departures from generalities found in mammals in both large-scale structure and the mechanisms shaping the evolution of the MHC. BioScience *52*, 423–431. https://doi.org/10.1641/0006-3568(2002)052%255B0423:TEOTMH%255D2.0.CO;2.

44. Delany, M.E., Robinson, C.M., Goto, R.M., and Miller, M.M. (2009). Architecture and Organization of Chicken Microchromosome 16: Order of the NOR, MHC-Y, and MHC-B Subregions. J Hered *100*, 507–514. https://doi.org/10.1093/jhered/esp044.

45. Dunkel, A., Steinhaus, M., Kotthoff, M., Nowak, B., Krautwurst, D., Schieberle, P., and Hofmann, T. (2014). Nature’s chemical signatures in human olfaction: a foodborne perspective for future biotechnology. Angew Chem Int Ed Engl *53*, 7124–7143. https://doi.org/10.1002/anie.201309508.

46. Barnea, A., and Rothschild, M. (2002). The odour of pyrazine increases the egg mass of domestic chickens (Gallus gallus domesticus). J. Zool. *257*, 411–416. https://doi.org/10.1017/S0952836902001012.

47. Saito, H., Kubota, M., Roberts, R.W., Chi, Q., and Matsunami, H. (2004). RTP Family Members Induce Functional Expression of Mammalian Odorant Receptors. Cell *119*, 679–691. https://doi.org/10.1016/j.cell.2004.11.021.

48. Vandewege, M.W., Mangum, S.F., Gabaldón, T., Castoe, T.A., Ray, D.A., and Hoffmann, F.G. (2016). Contrasting patterns of evolutionary diversification in the olfactory repertoires of reptile and bird genomes. Genome Biol. Evol. *8*, 470–480. https://doi.org/10.1093/gbe/evw013.

49. Montague, M.J., Li, G., Gandolfi, B., Khan, R., Aken, B.L., Searle, S.M.J., Minx, P., Hillier, L.W., Koboldt, D.C., Davis, B.W., et al. (2014). Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. Proc. Natl. Acad. Sci. U.S.A. *111*, 17230–17235. https://doi.org/10.1073/pnas.1410083111.

50. Beichman, A.C., Koepfli, K.-P., Li, G., Murphy, W., Dobrynin, P., Kliver, S., Tinker, M.T., Murray, M.J., Johnson, J., Lindblad-Toh, K., et al. (2019). Aquatic adaptation and depleted diversity: a deep dive into the genomes of the sea otter and giant otter. Mol Biol Evol *36*, 2631–2655. https://doi.org/10.1093/molbev/msz101.

51. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence slignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

52. Steiger, S.S., Fidler, A.E., and Kempenaers, B. (2009). Evidence for increased olfactory receptor gene repertoire size in two nocturnal bird species with well-developed olfactory ability. BMC Evol. Biol. *9*, 117. https://doi.org/10.1186/1471-2148-9-117.

53. Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. *16*, 111–120. https://doi.org/10.1007/BF01731581.

54. Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL–EBI. Nucleic Acids Res. *38*, W695–W699. https://doi.org/10.1093/nar/gkq313.

55. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

56. Letunic, I., and Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res *52*, W78–W82. https://doi.org/10.1093/nar/gkae268.

57. Feng, S., Stiller, J., Deng, Y., Armstrong, J., Fang, Q., Reeve, A.H., Xie, D., Chen, G., Guo, C., Faircloth, B.C., et al. (2020). Dense sampling of bird diversity increases power of comparative genomics. Nature *587*, 252–257. https://doi.org/10.1038/s41586-020-2873-9.

58. Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633.

59. Oliveros, C.H., Field, D.J., Ksepka, D.T., Barker, F.K., Aleixo, A., Andersen, M.J., Alström, P., Benz, B.W., Braun, E.L., Braun, M.J., et al. (2019). Earth history and the passerine superradiation. Proc. Natl. Acad. Sci. U.S.A. *116*, 7916–7925. https://doi.org/10.1073/pnas.1813206116.

60. Barker, F.K., Burns, K.J., Klicka, J., Lanyon, S.M., and Lovette, I.J. (2015). New insights into New World biogeography: An integrated view from the phylogeny of blackbirds, cardinals, sparrows, tanagers, warblers, and allies. The Auk *132*, 333–348. https://doi.org/10.1642/AUK-14-110.1.

61. Faircloth, B.C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics *32*, 786–788. https://doi.org/10.1093/bioinformatics/btv646.

62. Castresana, J. (2000). Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. Mol Biol Evol *17*, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334.

63. Stiller, J., Feng, S., Chowdhury, A.-A., Rivas-González, I., Duchêne, D.A., Fang, Q., Deng, Y., Kozlov, A., Stamatakis, A., Claramunt, S., et al. (2024). Complexity of avian evolution revealed by family-level genomes. Nature *629*, 851–860. https://doi.org/10.1038/s41586-024-07323-1.

64. Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633.

65. Danner, R.M., Gulson-Castillo, E.R., James, H.F., Dzielski, S.A., Frank, D.C., III, Sibbald, E.T., and Winkler, D.W. (2017). Habitat-specific divergence of air conditioning structures in bird bills. The Auk *134*, 65–75. https://doi.org/10.1642/AUK-16-107.1.

66. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

67. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

68. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686. https://doi.org/10.21105/joss.01686.

69. Yu, G. (2020). Using ggtree to visualize data on tree-like structures. Curr. Protoc. Bioinform. *69*, e96. https://doi.org/10.1002/cpbi.96.

70. Nagai, M.H., and Matsunami, H. (2023). Activity-Dependent Labeling of Olfactory Sensory Neurons Using RNA Fluorescence In Situ Hybridization Followed by Phospho-S6 Immunofluorescence. In The Olfactory System, B. J. Goldstein and H. Matsunami, eds. (Springer US), pp. 83–97. https://doi.org/10.1007/978-1-0716-3425-7_7.

71. Davidson, L.A., and Keller, R.E. (1999). Neural tube closure in Xenopus laevis involves medial migration, directed protrusive activity, cell intercalation and convergent extension. Development *126*, 4547–4556. https://doi.org/10.1242/dev.126.20.4547.

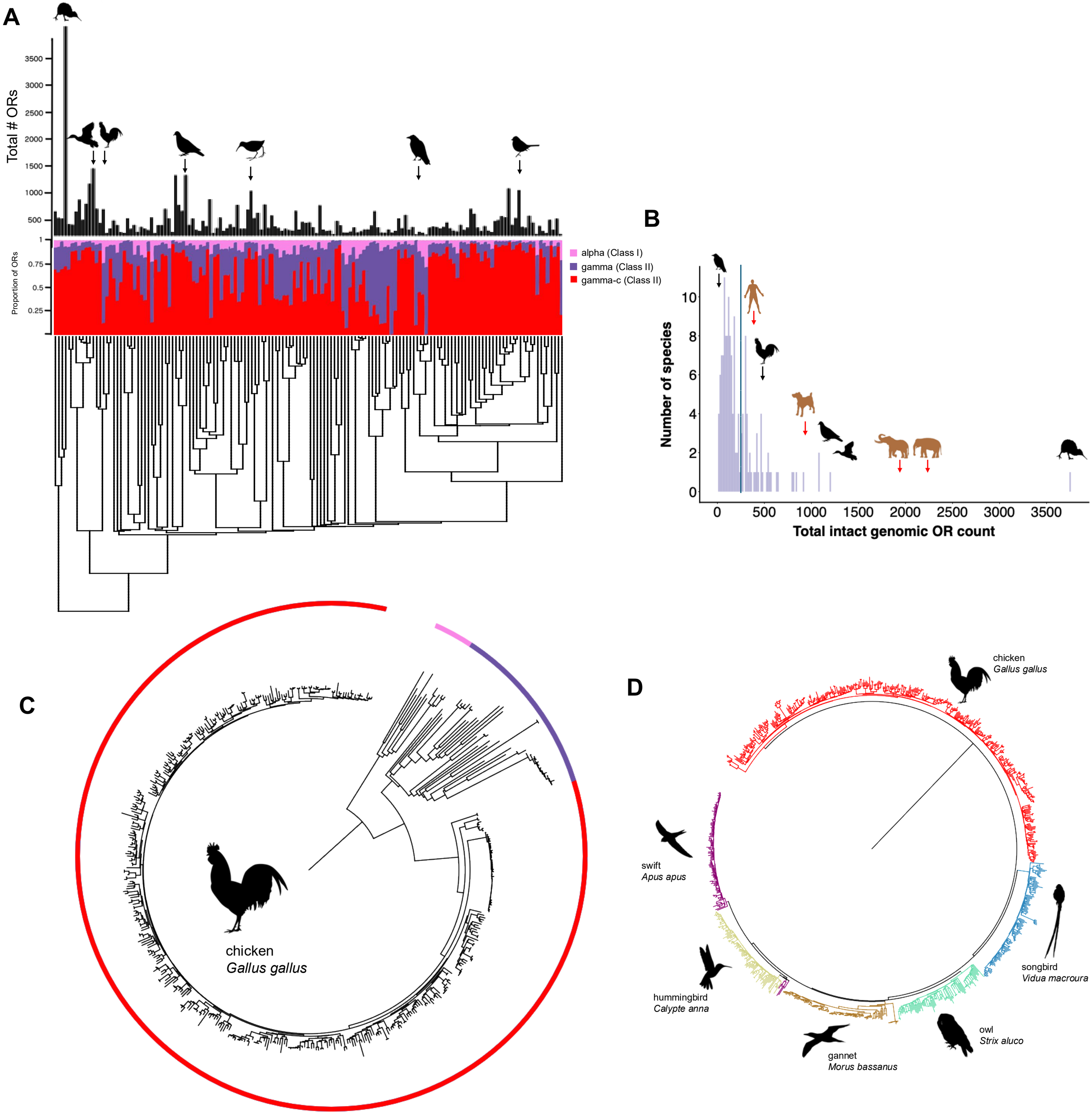
72. Zhou, X., and Vize, P.D. (2004). Proximo-distal specialization of epithelial transport processes within the Xenopus pronephric kidney tubules. Dev Biol *271*, 322–338. https://doi.org/10.1016/j.ydbio.2004.03.036.

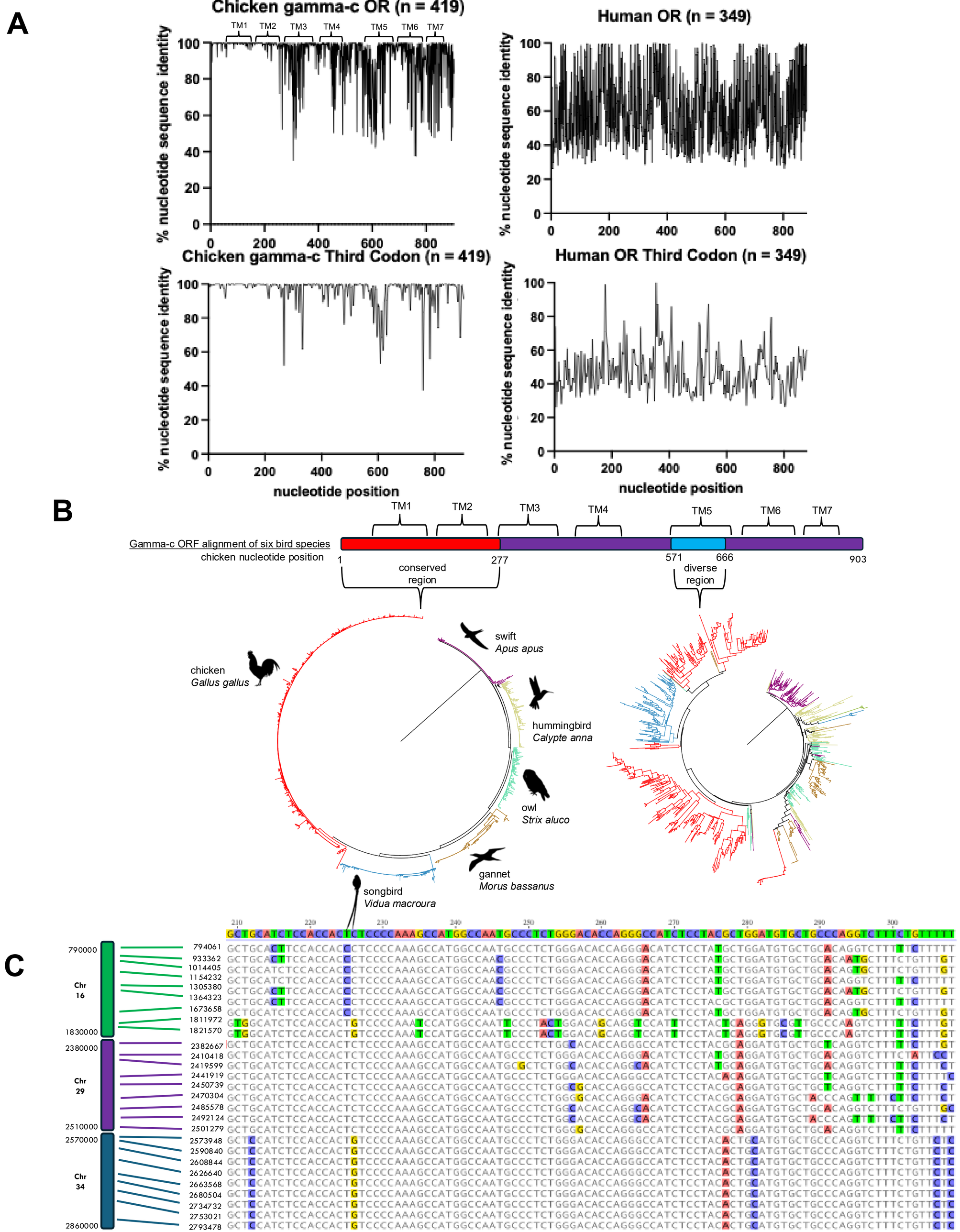
73. Shepard, B.D., Natarajan, N., Protzko, R.J., Acres, O.W., and Pluznick, J.L. (2013). A Cleavable N-Terminal Signal Peptide Promotes Widespread Olfactory Receptor Surface Expression in HEK293T Cells. PLOS ONE *8*, e68758. https://doi.org/10.1371/journal.pone.0068758.

74. Horak, M., and Wenthold, R.J. (2009). Different Roles of C-terminal Cassettes in the Trafficking of Full-length NR1 Subunits to the Cell Surface*. Journal of Biological Chemistry *284*, 9683–9691. https://doi.org/10.1074/jbc.M807050200.

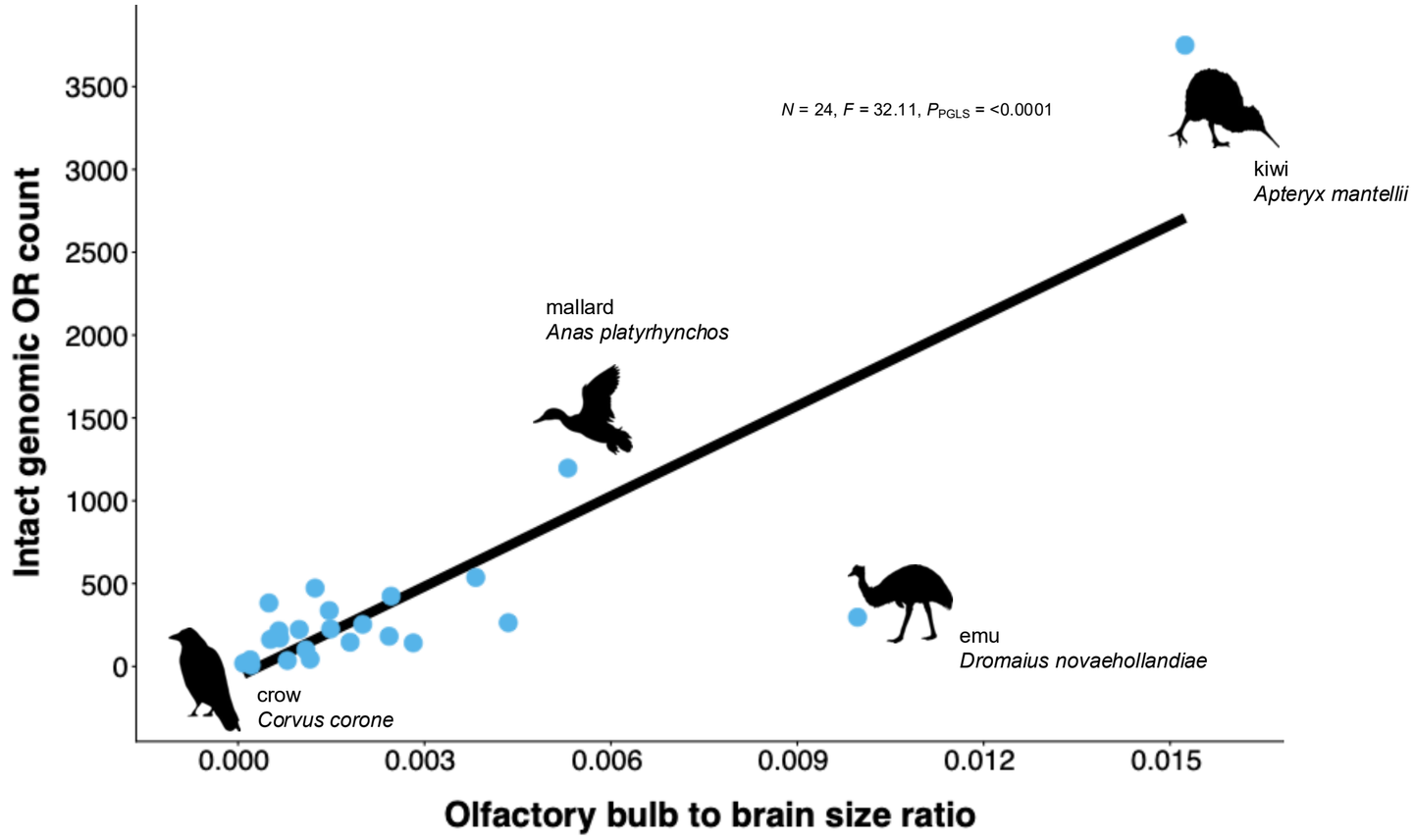
75. Tewari, J., and Matsunami, H. (2023). Measuring Cell Surface Expression of Odorant Receptors via Flow Cytometry. In The Olfactory System, B. J. Goldstein and H. Matsunami, eds. (Springer US), pp. 99–109. https://doi.org/10.1007/978-1-0716-3425-7_8.

76. Wu, L., Pan, Y., Chen, G.-Q., Matsunami, H., and Zhuang, H. (2012). Receptor-transporting Protein 1 Short (RTP1S) Mediates Translocation and Activation of Odorant Receptors by Acting through Multiple Steps *. Journal of Biological Chemistry *287*, 22287–22294. https://doi.org/10.1074/jbc.M112.345884.

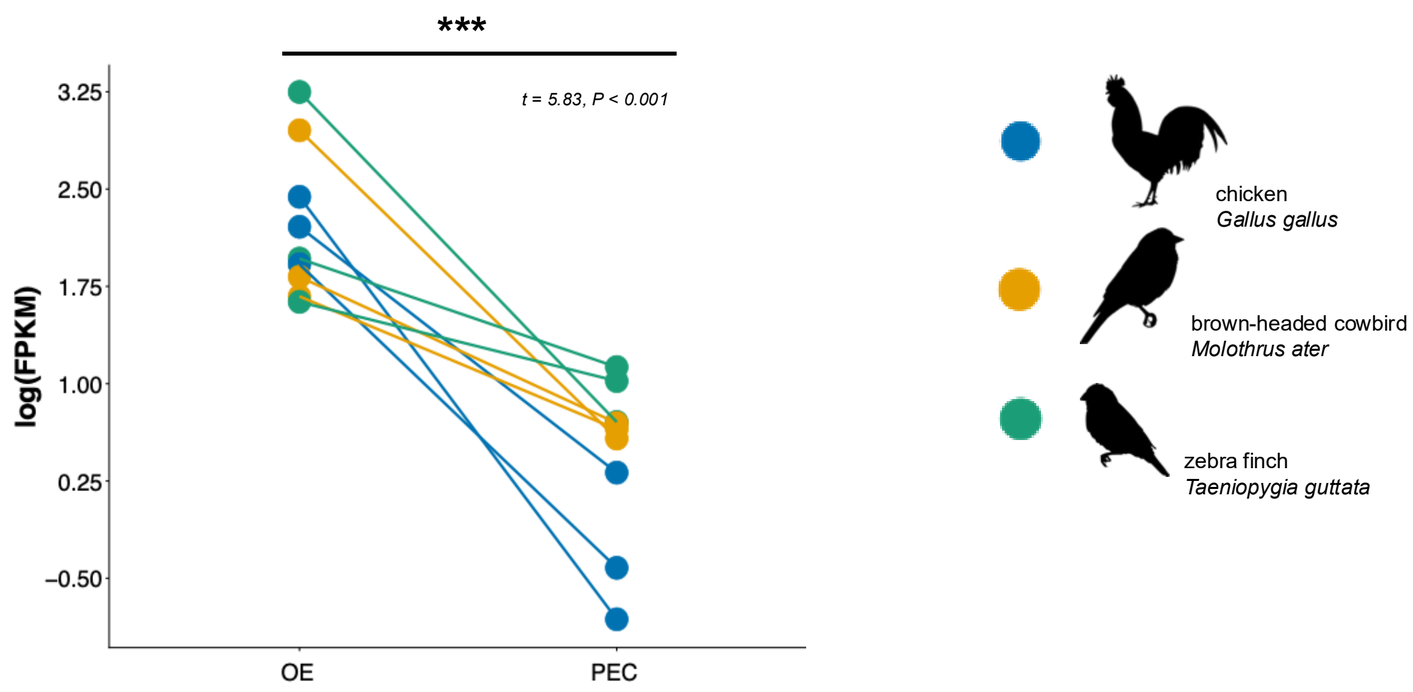

A
Total # ORs
3500
3000
2500
2000
1500
1000
500
1
Proportion of ORs
0.75
0.5
0.25
alpha (Class I)
gamma (Class II)
gamma-c (Class II)
B
Number of species
10
8
6
4
2
0
0
500
1000
1500
2000
2500
3000
3500
Total intact genomic OR count
C
chicken
Gallus gallus
D
chicken
Gallus gallus
swift
Apus apus
hummingbird
Calypte anna
gannet
Morus bassanus
owl
Strix aluco
songbird
Vidua macroura


A
Chicken gamma-c OR (n = 419)
Human OR (n = 349)
Chicken gamma-c Third Codon (n = 419)
Human OR Third Codon (n = 349)
TM1
TM2
TM3
TM4
TM5
TM6
TM7
% nucleotide sequence identity
nucleotide position
B
Gamma-c ORF alignment of six bird species
chicken nucleotide position
1
277
571
666
903
conserved region
diverse region
chicken
Gallus gallus
swift
Apus apus
hummingbird
Calypte anna
owl
Strix aluco
gannet
Morus bassanus
songbird
Vidua macroura
C
Chr 16
Chr 29
Chr 34

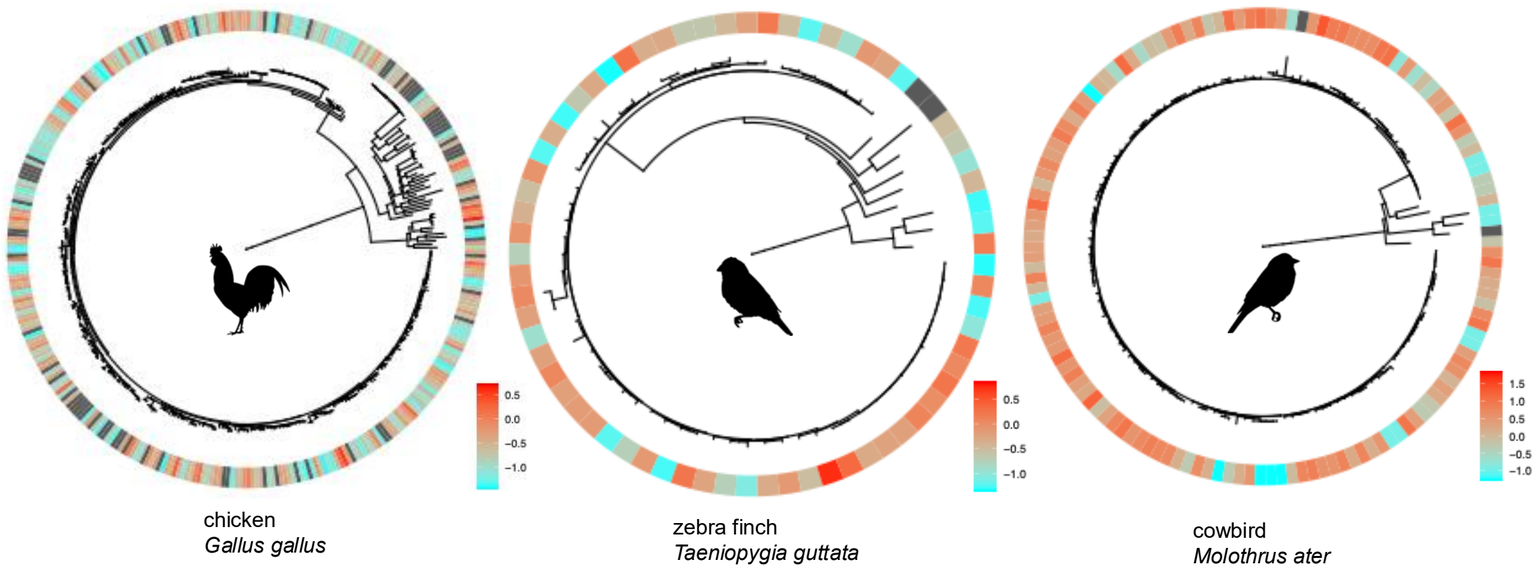
A


N = 24, F = 32.11, P_PGLS = <0.0001
kiwi
Apteryx mantellii
mallard
Anas platyrhynchos
emu
Dromaius novaehollandiae
crow
Corvus corone
Intact genomic OR count
3500
3000
2500
2000
1500
1000
500
0
0.000
0.003
0.006
0.009
0.012
0.015
Olfactory bulb to brain size ratio

B


***
t = 5.83, P < 0.001
log(FPKM)
3.25
2.50
1.75
1.00
0.25
−0.50
OE
PEC
chicken
Gallus gallus
brown-headed cowbird
Molothrus ater
zebra finch
Taeniopygia guttata

C


chicken
Gallus gallus
zebra finch
Taeniopygia guttata
cowbird
Molothrus ater

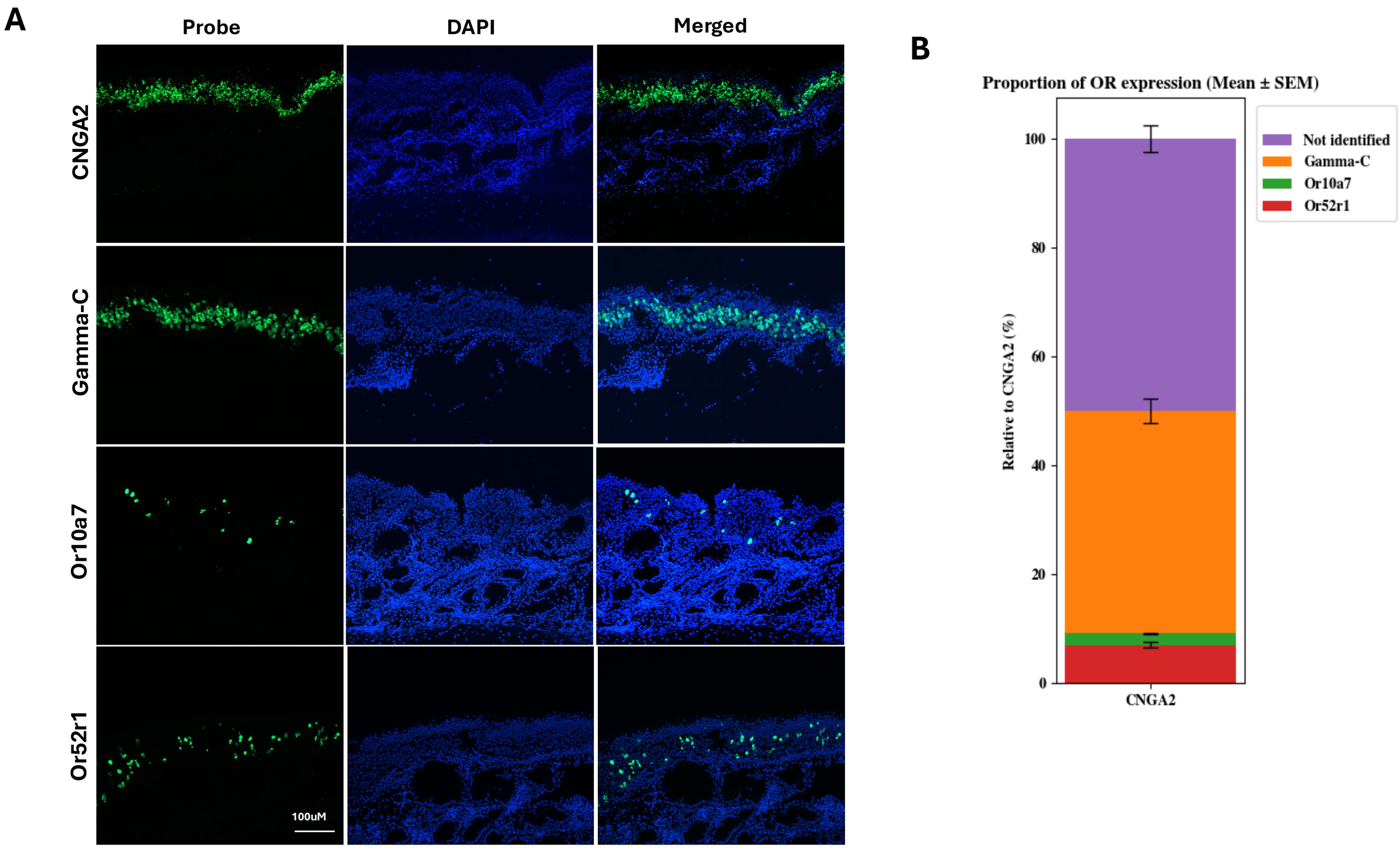
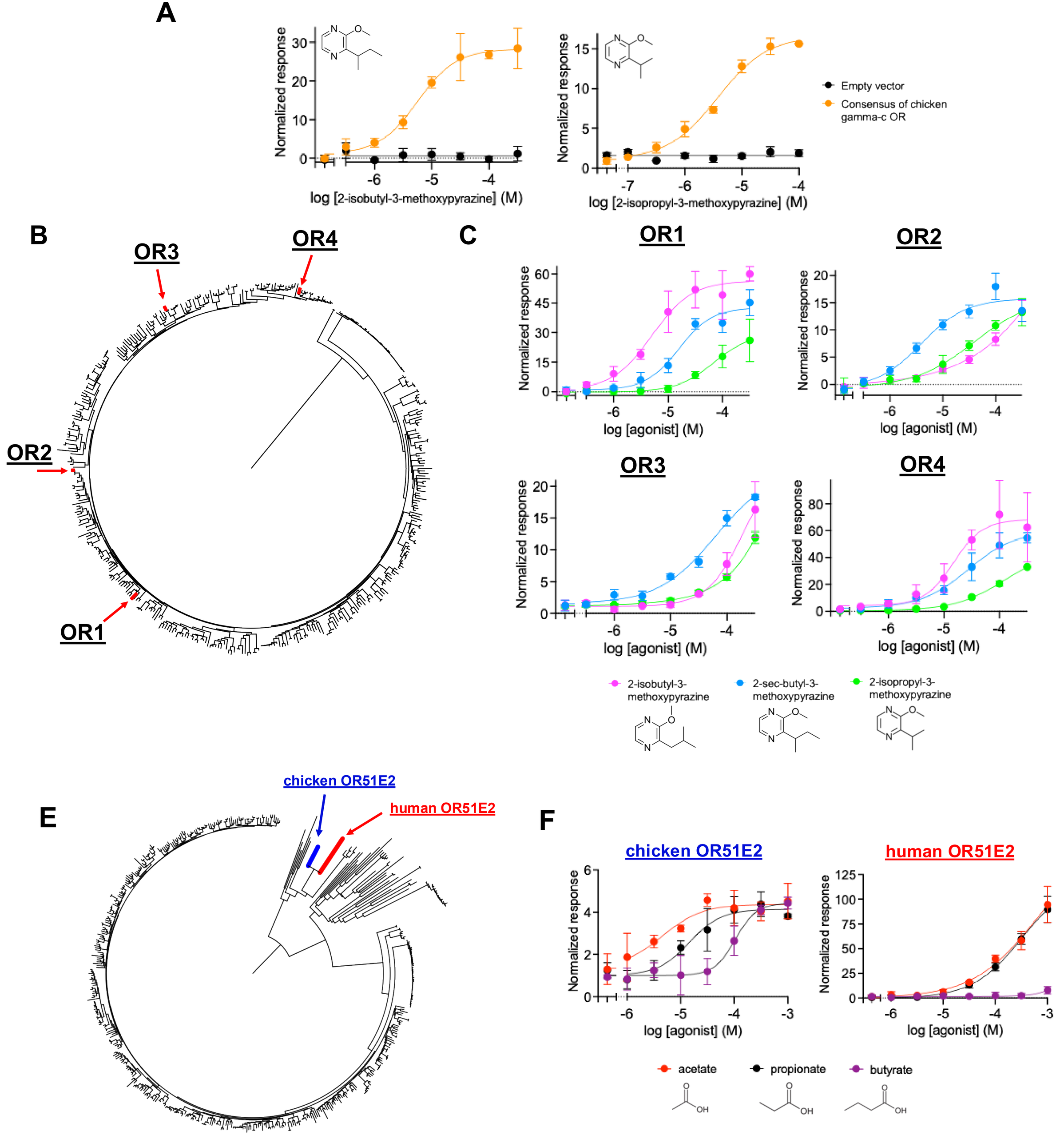
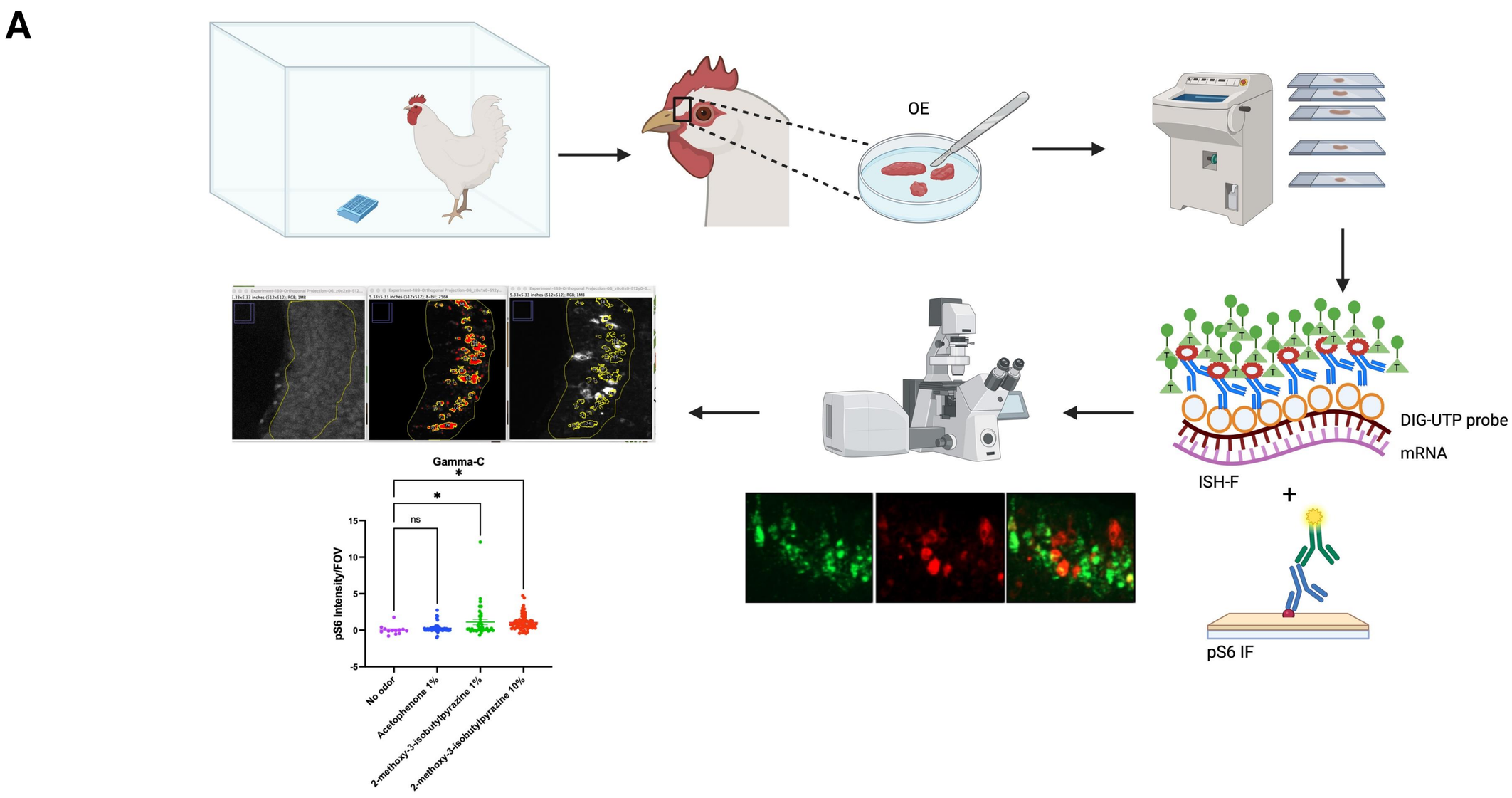
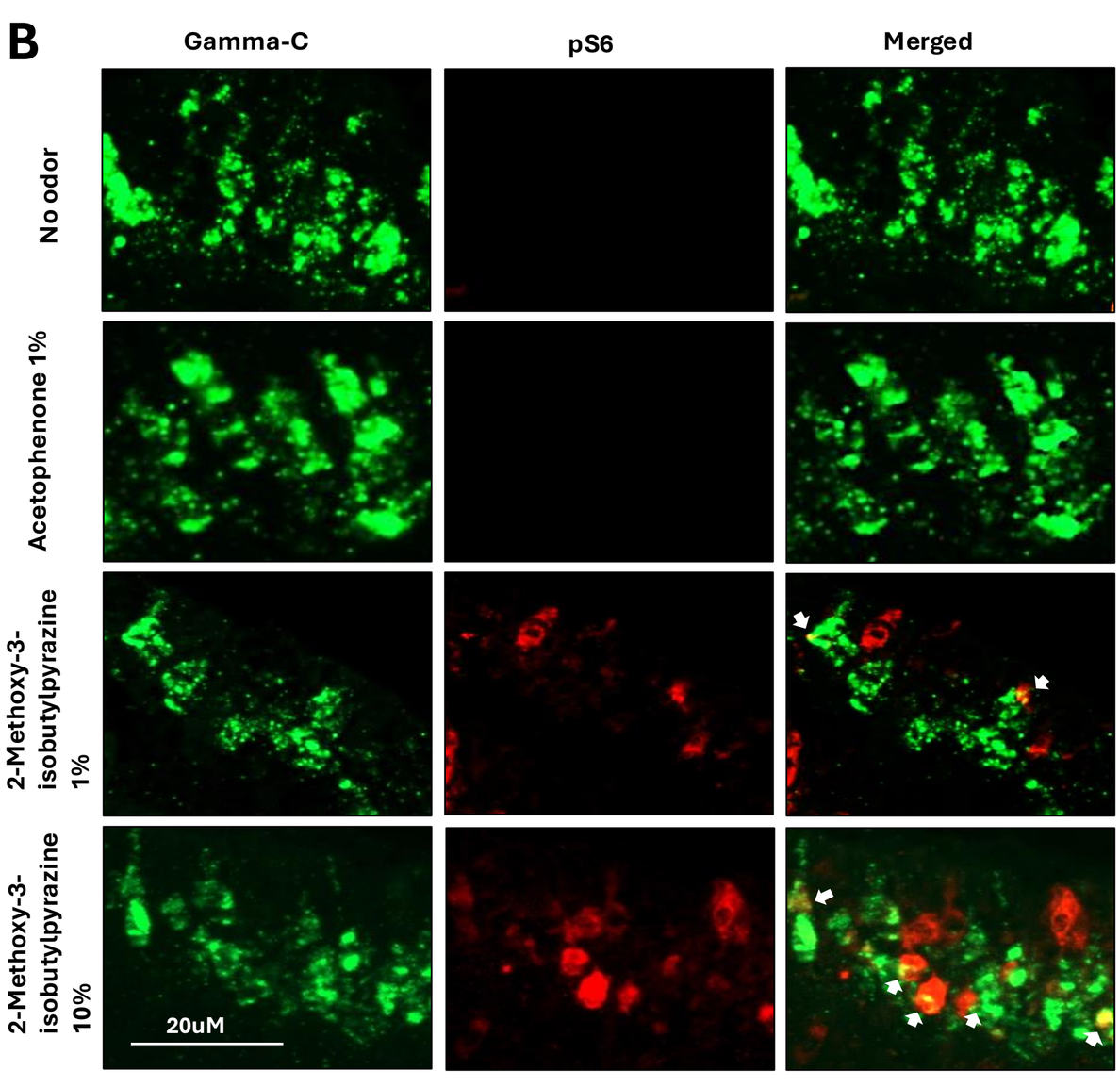
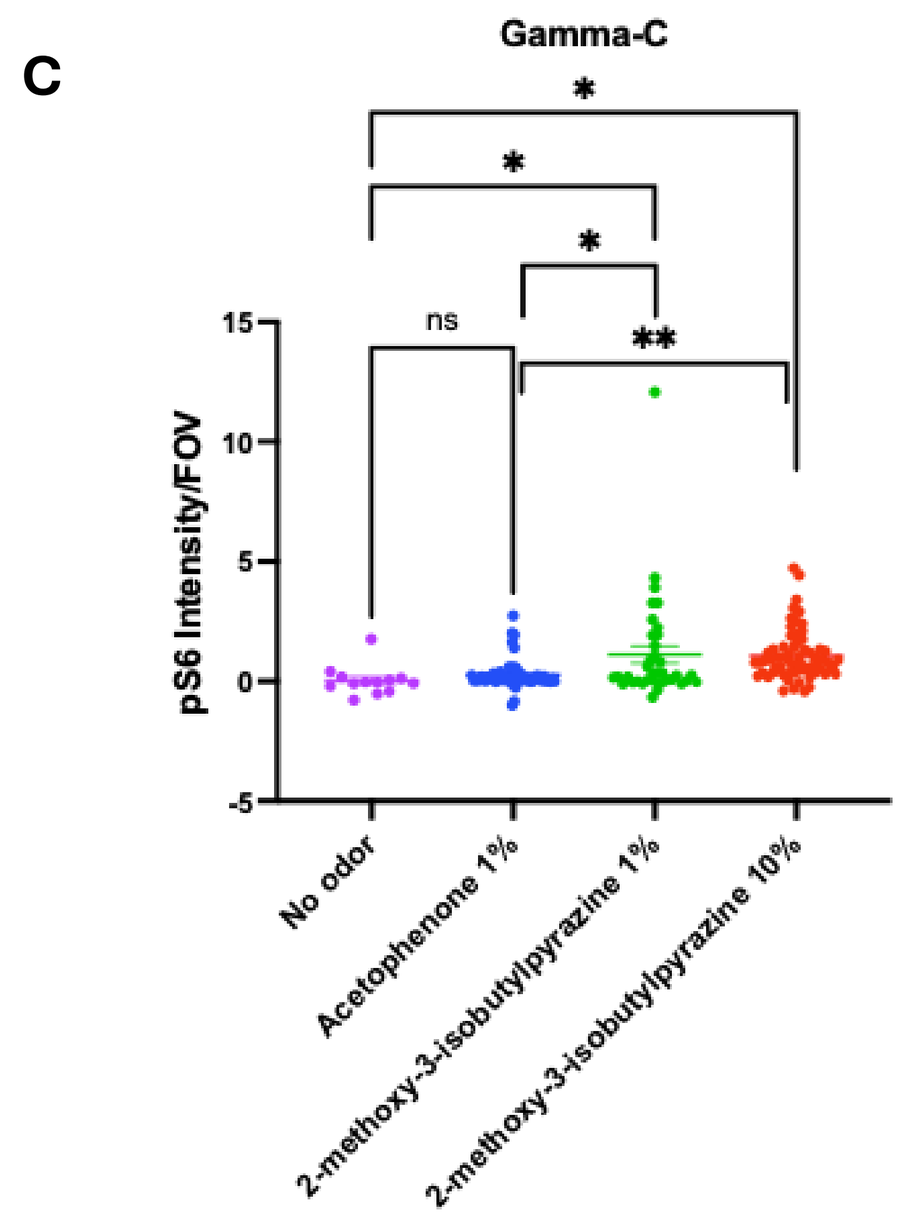


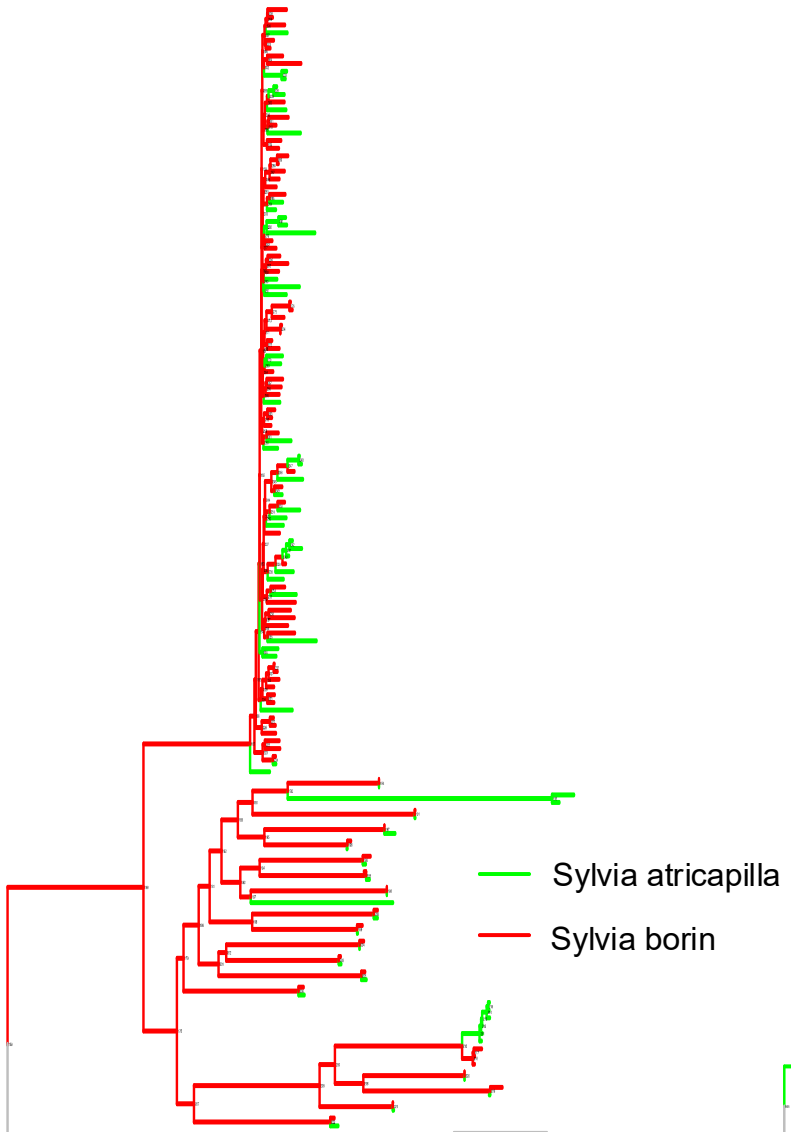
Fig. Odorant Receptors (Ors) distribution in chicken olfactory epithelium (OE). (A) In situ hybridization (ISH) for CNGA2, Gamma-C, Or10a7, and Or52r1 (top to bottom, respectively), and (B) Proportions for Ors expression relative to the OSN marker CNGA2. Scale bar shown in 100uM.

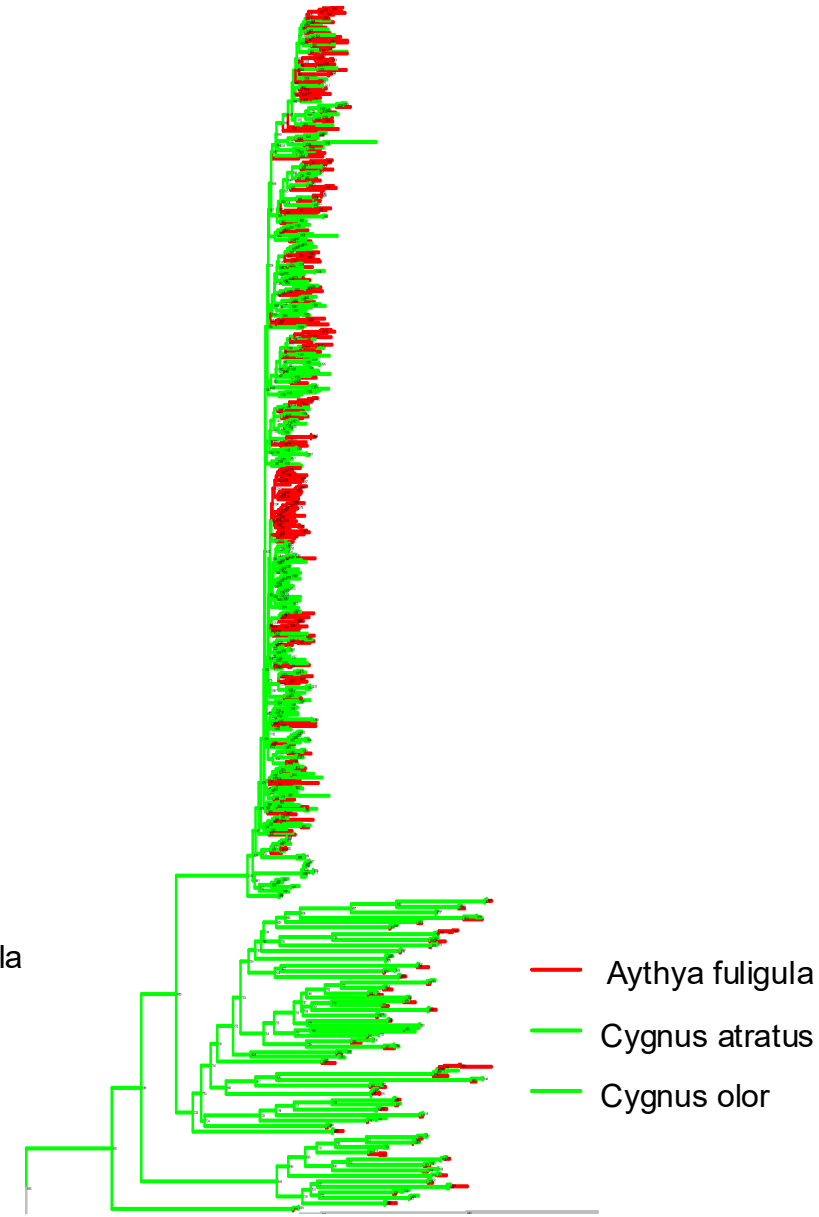

A
Normalized response
log [2-isobutyl-3-methoxypyrazine] (M)
log [2-isopropyl-3-methoxypyrazine] (M)
Empty vector
Consensus of chicken gamma-c OR
B
OR3
OR4
OR2
OR1
C
OR1
OR2
OR3
OR4
log [agonist] (M)
2-isobutyl-3-methoxypyrazine
2-sec-butyl-3-methoxypyrazine
2-isopropyl-3-methoxypyrazine
E
chicken OR51E2
human OR51E2
F
chicken OR51E2
human OR51E2
acetate
propionate
butyrate

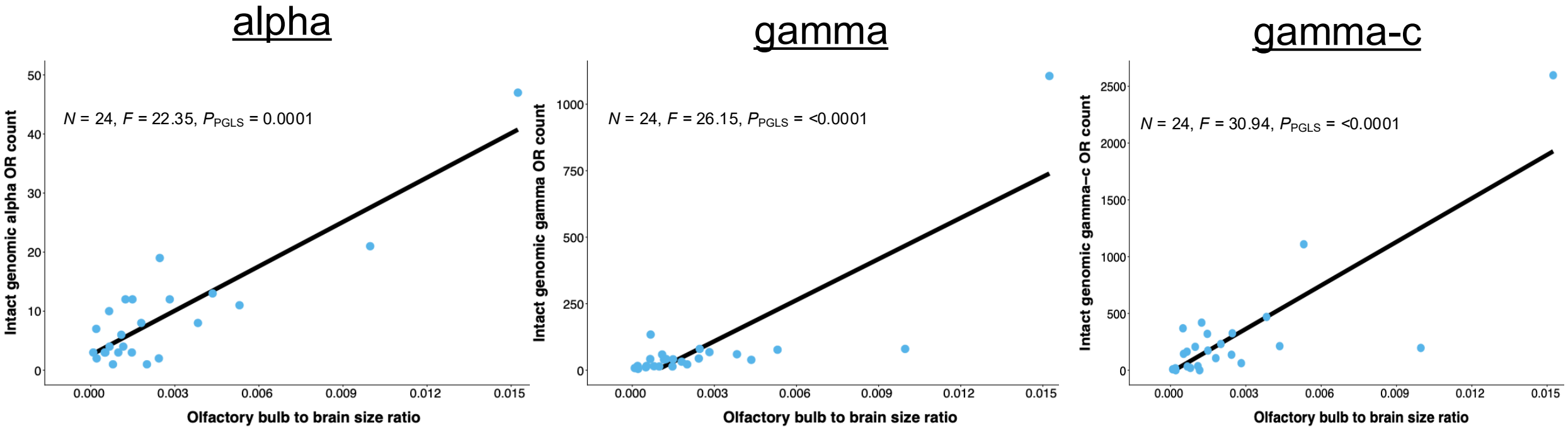

A
OE
DIG-UTP probe
mRNA
ISH-F
+
pS6 IF
Gamma-C
pS6 Intensity/FOV
No odor
Acetophenone 1%
2-methoxy-3-isobutylpyrazine 1%
2-methoxy-3-isobutylpyrazine 10%
ns
*

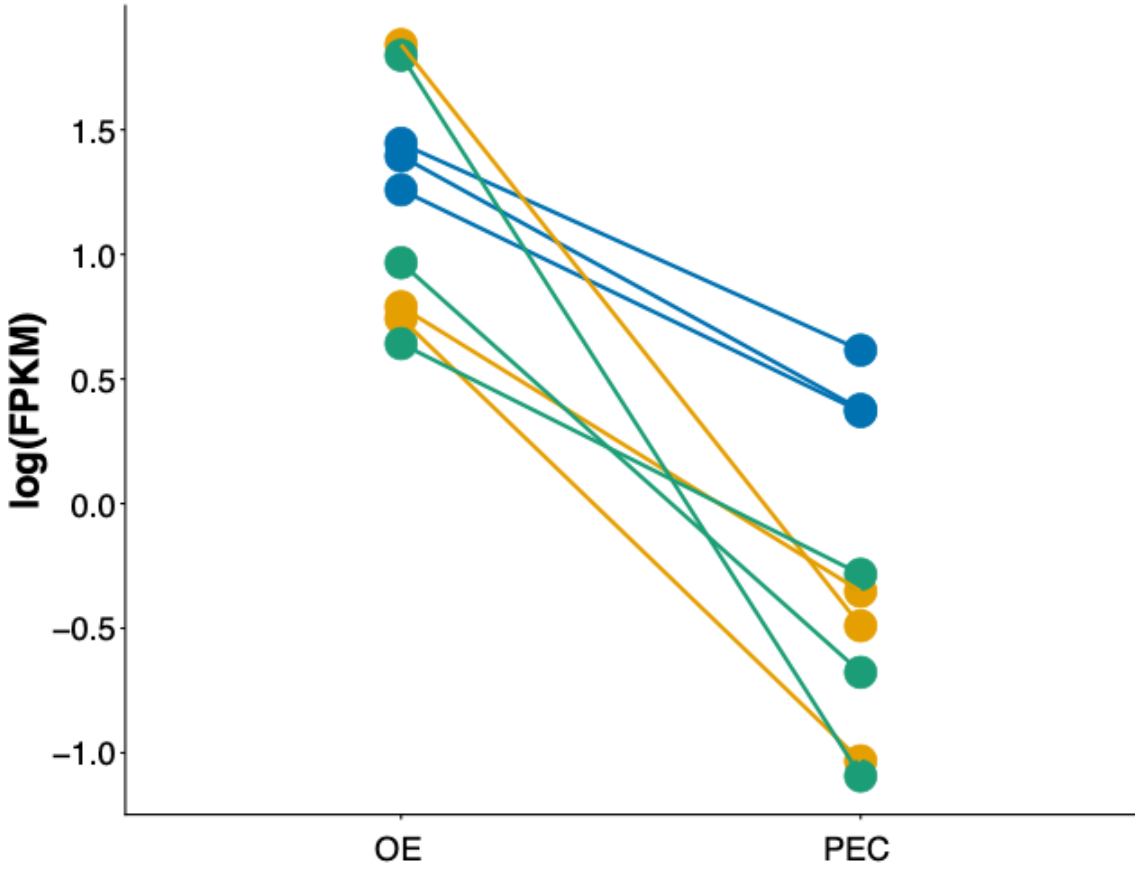

B
Gamma-C
pS6
Merged
No odor
Acetophenone 1%
2-Methoxy-3-isobutylpyrazine 1%
2-Methoxy-3-isobutylpyrazine 10%
20uM

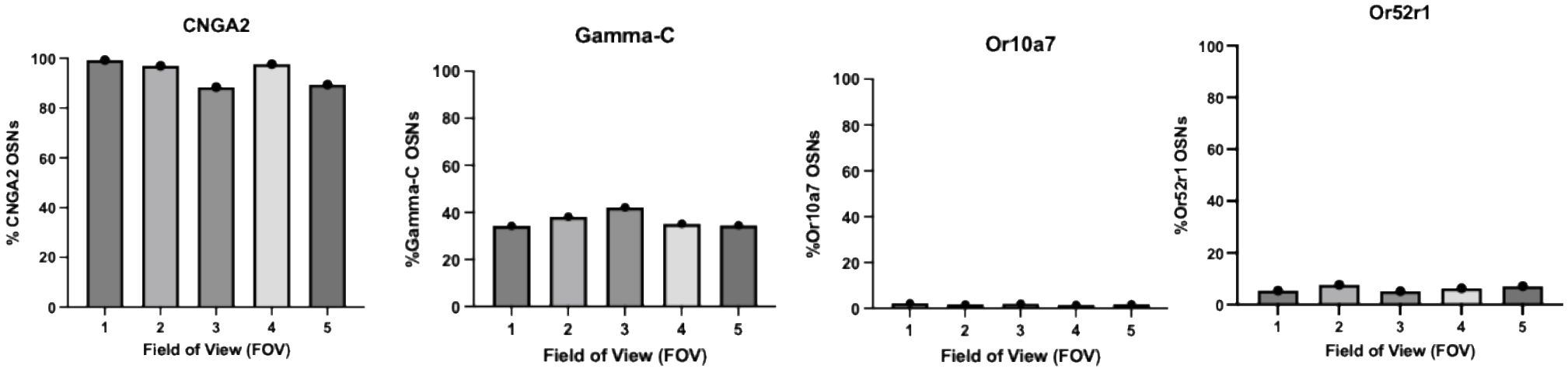

C
Gamma-C
pS6 Intensity/FOV
15
10
5
0
-5
ns
*
**
No odor
Acetophenone 1%
2-methoxy-3-isobutylpyrazine 1%
2-methoxy-3-isobutylpyrazine 10%


Sylvia atricapilla
Sylvia borin

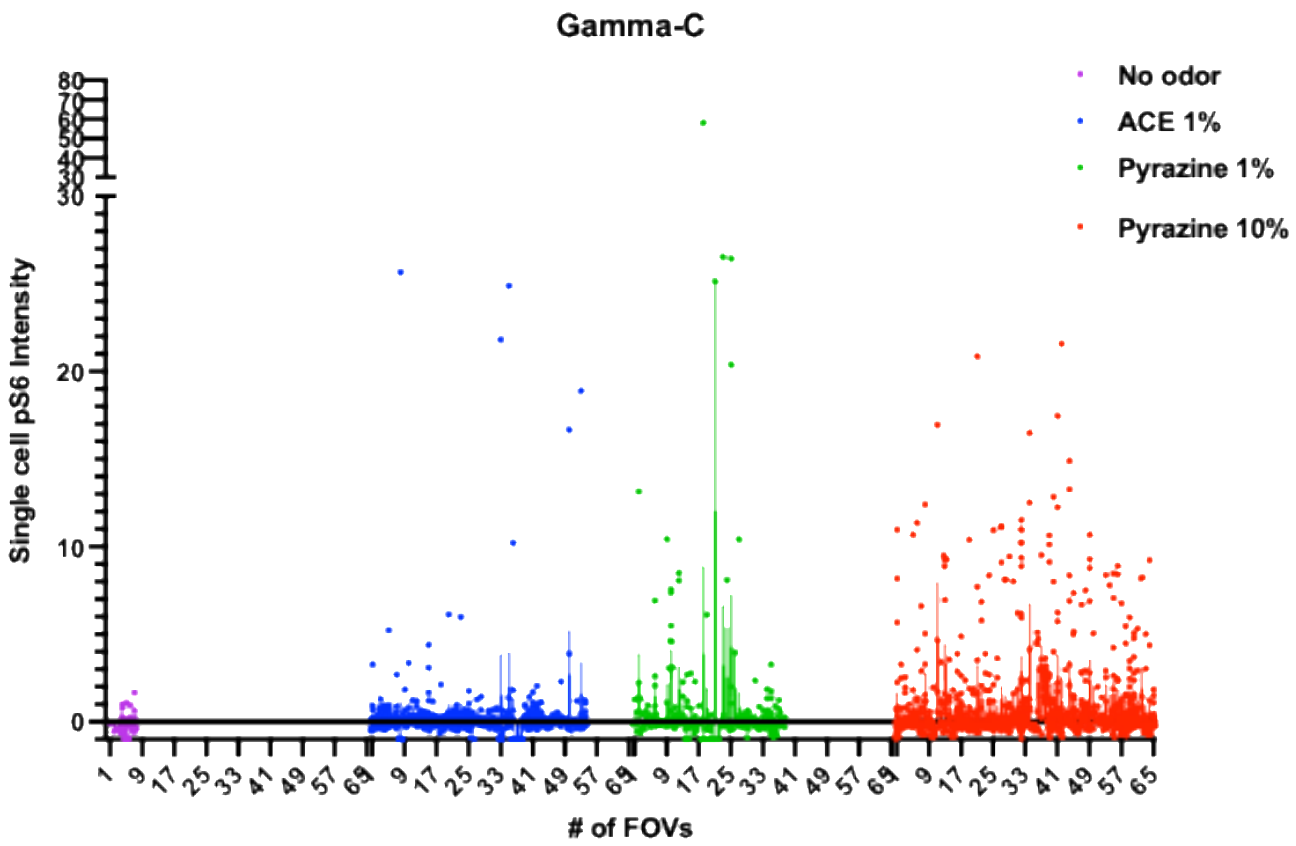

Aythya fuligula
Cygnus atratus
Cygnus olor


alpha
gamma
gamma-c
N = 24, F = 22.35, PPGLS = 0.0001
N = 24, F = 26.15, PPGLS = <0.0001
N = 24, F = 30.94, PPGLS = <0.0001
Intact genomic alpha OR count
Intact genomic gamma OR count
Intact genomic gamma–c OR count
Olfactory bulb to brain size ratio
Olfactory bulb to brain size ratio
Olfactory bulb to brain size ratio
0
10
20
30
40
50
0
250
500
750
1000
0
500
1000
1500
2000
2500
0.000
0.003
0.006
0.009
0.012
0.015


log(FPKM)
1.5
1.0
0.5
0.0
−0.5
−1.0
OE
PEC



5 randomly selected FOVs analyzed for semiquantitative analysis of ISH of the percent of OSN marker CNGA2 and different chicken Ors in the chicken OE. Analysis was done using ImageJ2, version 2.9.0/1.53t


Gamma-C
No odor
ACE 1%
Pyrazine 1%
Pyrazine 10%
Single cell pS6 Intensity
80
70
60
50
40
30
20
10
0
# of FOVs