Homeotic Regulation of Olfactory Receptor Choice via NFI-dependent Heterochromatic Silencing and Genomic Compartmentalization

Elizaveta Bashkirova[1,2], Kevin Monahan[3], Christine E. Campbell[4,5], Jason M. Osinski[4,6], Longzhi Tan[7], Ira Schieren[8], Gilad Barnea[9], X. Sunnie Xie[10,11], Richard M. Gronostajski[4,5] and Stavros Lomvardas[1,2,8,12,13*].

[1] Integrated Program in Cellular, Molecular and Biomedical Studies, Vangelos College of Physicians and Surgeons, Columbia University Irving Medical Center, Columbia University, New York, NY, 10032

[2] Department of Biochemistry and Molecular Biophysics, Vangelos College of Physicians and Surgeons, Columbia University Irving Medical Center, Columbia University, New York, NY, 10032

[3] Department of Biochemistry and Molecular Biology, Rutgers University, NJ

[4] Department of Biochemistry, New York State University at Buffalo, Buffalo, NY.

[5] University at Buffalo and New York State Center of Excellence in Bioinformatics and Life Sciences, Buffalo, NY

[6]Genetics, Genomics, and Bioinformatics Graduate Program, New York State University at Buffalo, Buffalo, NY.

[7] Department of Bioengineering, Stanford University, CA

[8] Zuckerman Mind, Brain, and Behavior Institute, Columbia University, New York, NY, 10027

[9] Department of Neuroscience, Division of Biology and Medicine, Brown University, Providence, RI

[10] Beijing Innovation Center for Genomics, Peking University, Beijing, China

[11] Biomedical Pioneering Innovation Center, Peking University, Beijing, China

[12] Department of Neuroscience, Vangelos College of Physicians and Surgeons, Columbia University Irving Medical Center, Columbia University, New York, NY

[13] Kavli Institute for Neurosciences at Columbia University, New York, NY, 10027

* Corresponding Author

Abstract
Expression of one out of >1000 olfactory receptor (OR) genes is stochastic but, yet, spatially organized in stereotypic anatomical zones of the olfactory epithelium. We discovered that transition to higher zones coincides with gradual increase of H3K9me3/H3K79me3 and expansion of long-range genomic interactions across OR genes with ascending zonal identities. Within a zone, however, heterochromatin levels, and genomic interaction frequencies progressively decline along increasing zonal OR identities. Consequently, in each differentiating neuron only ORs with the correct zonal identity combine chromatin state and genomic compartmentalization properties compatible with future gene choice; ORs with lower zonal identities have high H3K79me3/H3K9me3 enrichment and are silenced, while ORs with higher zonal identities are excluded from activating genomic interactions and, thus, cannot be chosen. This process is regulated by the gradient expression of NFIA, B, and X, triple deletion of which causes homeotic transformations on zonal OR expression, heterochromatin formation and genomic compartmentalization.

Introduction

Pattern formation governs fundamental developmental processes in plants and animals, vertebrates and invertebrates. From the classic theoretical predictions on the chemical basis of morphogenesis(Turing, 1953), to the seminal demonstration of the genetic control of segmentation in drosophila(Nusslein-Volhard and Wieschaus, 1980); and from the elucidation of motor neuron pool identity in mammals (Dasen et al., 2005), to the explanation of photoreceptor patterning in invertebrate ommatidia(Johnston et al., 2011), the molecular and genetic underpinnings of developmental patterning are -at various extend- understood. However, the mechanisms that regulate developmental patterning in the mammalian olfactory system are completely unknown. In this sensory system, spatial patterns are apparent in the periphery, the main olfactory epithelium (MOE) and the olfactory bulb (OB)(Sullivan et al., 1995; Sullivan et al., 1994), and less so in higher order brain processing centers(Stettler and Axel, 2009). In the MOE expression of >1000 class II olfactory receptor (OR) genes(Buck and Axel, 1991) from 18 different chromosomes(Sullivan et al., 1996; Zhang and Firestein, 2009) is organized in continuous “zones”, that are faithfully represented by zonally-restricted neuronal projections to corresponding areas of the OB(Ressler et al., 1994). Although there is evidence that the zonal organization of sensory inputs has behavioral and perceptual significance, ignorance of the molecular underpinnings of zonal OR organization prevents the systematic evaluation of the importance of zonal sensory inputs in olfactory perception.

Zones are stereotypic and evolutionarily conserved subdivisions of the MOE that express preferentially 50-200 OR genes from the whole OR repertoire. Originally, RNA ISH experiments revealed that the MOE is divided in 4 broad zones (Ressler et al., 1993; Vassar et al., 1993), with highly overlapping borders(Miyamichi et al., 2005). Single cell RNA-seq experiments in micro-dissected MOE pieces assigned zonal coordinates to the majority of the ORs and redefined the anatomical subdivisions of the MOE into 5 continuous zones (Tan and Xie, 2018). The septal organ, a distinct anatomical structure of the MOE expressing predominantly a single OR could be viewed as a distinct MOE zone(Tian and Ma, 2004), and so could the cul-de-sac MOE segments that express MS4A chemoreceptors(Greer et al., 2016). In fact, comprehensive 3D imaging of OR expression by multiplex RNA ISH, raised the prospect of 9 or more OR zones in the MOE(Zapiec and Mombaerts, 2020). Beyond these anatomical distinctions that are established along the dorsomedial-ventrolateral axes of the MOE, additional OSN subdivisions exist that do not entail spatial segregation of molecularly distinct cell types. For example, OSNs expressing class I ORs (fish-like ORs)(Bozza et al., 2009), or trace amine associated receptors (TAARs)(Johnson et al., 2012), are committed to each of these chemoreceptor subfamilies but are intermingled with class II-expressing OSNs in their respective zones. Thus, the emerging picture regarding chemoreceptor gene expression in the MOE suggests a hierarchical process that starts with broad zonal specification and commitment to a chemoreceptor family, followed by a singular stochastic choice of one of the zonally enabled receptor alleles. Although we made strides in the understanding of how this singular stochastic choice is accomplished, the molecular mechanisms that restricts it within a zonal repertoire are not understood.

The process of singular OR gene choice starts in OSN progenitors, where class II OR gene clusters become marked with the hallmarks of constitutive heterochromatin(Magklara et al., 2011), followed by the formation of OR-exclusive interchromosomal nuclear compartments(Armelin-Correa et al., 2014; Clowney et al., 2012; Tan et al., 2019). OR compartments, which assemble due to downregulation of Lamin b receptor (Lbr) in OSN progenitors, serve two opposing regulatory purposes: they further silence OR transcription, by insulating ORs from the transcription machinery, but they also promote the assembly of a juxtaposed multi-chromosomal enhancer hub that associates with one of the compartmentalized ORs, activating its transcription (Clowney et al., 2012; Markenscoff-Papadimitriou et al., 2014). This multi-enhancer hub consists of intergenic OR enhancers, the Greek Islands, which are bound by Lhx2, Ebf and Ldb1(Markenscoff-Papadimitriou et al., 2014; Monahan et al., 2019; Monahan et al., 2017). This tripartite protein complex promotes the highly specific interchromosomal contacts between Greek Islands, creating a strong transactivating complex responsible for the de-silencing and robust transcriptional activation of only one of the compartmentalized ORs. Translation of the chosen OR in the ER induces the unfolded protein response pathway(Dalton et al., 2013) and generates a feedback signal(Lewcock and Reed, 2004; Serizawa et al., 2003; Shykind et al., 2004) that hinders the ability of the Greek Island hub to de-silence additional OR alleles, stabilizing singular OR gene choice for the life of the OSN(Lyons et al., 2013).

The aforementioned complexity in OR gene regulation likely compensates for the fact that traditional mechanisms of combinatorial gene regulation cannot determine which OR is activated in each OSN. Numerous lines of evidence support the notion that OR gene choice is a stochastic process. First, “allelic exclusion”, expression from only the maternal or paternal allele in each OSN, suggests that *cis* and *trans* regulatory information cannot sufficiently define which OR will be expressed in each cell(Chess et al., 1994). In this note, transgenic ORs, from minigenes(Vassalli et al., 2002) to YAC transgenes(Ebrahimi et al., 2000; Serizawa et al., 2000), are never co-expressed with the endogenous OR alleles from which they were derived, arguing against a deterministic mode of OR regulation. Second, OR gene switching, a low frequency switch from the original chosen OR allele to a different OR identity, further supports the OR multipotency of each OSN (Shykind et al., 2004). Third, sc-RNA-seq studies revealed that there is a transient differentiation stage where OSN progenitors co-transcribe seemingly random combinations of 10-15 OR genes(Hanchate et al., 2015; Saraiva et al., 2015; Tan et al., 2015). These results taken together suggests that an unknown mechanism integrates the position of the OSN with the identity of each OR to restrict this stochastic choice within a spatially-determined OR sub-repertoire.

To obtain insight at the remarkable problem of spatially determined randomness, we performed a comparative analysis of the transcriptome, chromatin modification landscape, and nuclear architecture of various cell populations from different MOE zones. Our data reveal that early OSN progenitors from every zone co-express predominantly zone 1 ORs. However, as we transition from zone 1 to zone 5, developmental progression results in silencing of these ORs and activation of ORs from higher zonal indexes, culminating in the singular expression of ORs with the correct zonal identity. These spatiotemporal transcriptional transitions are regulated by two mechanistically linked processes: marking of an increasing number of OR genes with H3K9me3 and H3K79me3, and gradual recruitment of an increasing number of OR genes to OR compartments. While heterochromatin formation and OR compartmentalization progressively increase in higher zonal segments and expand to ascending zonal identities, within a zone, ORs with lower zonal identities have the highest levels of H3K9me3/H3K79me3 and genomic contacts. Consequently, in each zone there are three types of ORs: ORs expressed in previous zones that have high levels of H3K9me3/H3K79me3 that keep them completely silent; ORs expressed in subsequent zones that lack both H3K9me3/H3K79me3 and activating long-range genomic interactions; and ORs with the proper zonal identity that have intermediate enrichment of H3K9me3/H3K79me3 and HiC contact frequencies. Thus, only ORs with the correct zonal identity combine reversible levels of heterochromatin with the ability to interact with other ORs and the Greek Island hub. This process is controlled by three paralogous transcription factors, NFI A, B, and X whose gradient expression in OSN progenitors increases towards zone 5, expanding both the number of ORs marked by H3K9me3/H3K79me3 and the zonal constitution of OR compartments. Triple NFI deletion in OSN progenitors transforms the heterochromatic landscape and the nuclear compartmentalization of zone 4/5 OSNs to patterns reminiscent of zone 2/3 OSNs, while it induces ectopic expression of zone 2/3 ORs in place of zone 4/5 ORs. Thus, NFI genes act as homeotic regulators of chromatin-mediated silencing, nuclear architecture, and OR gene choice in the MOE.

Results

OR transcription undergoes a zonal transformation during OSN differentiation

Before the onset of singular OR expression, mitotically active OSN progenitors co-express multiple OR genes at levels that are high enough for detection by scRNA-seq(Hanchate et al., 2015; Saraiva et al., 2015; Tan et al., 2015). To identify the developmental stage that zonal restrictions are established we asked if the co-transcribed ORs follow zonal patterns of expression in these early OSN progenitors. We analyzed specific OSN progenitor populations isolated by FACS from Mash1CreER; TdTomato (a Cre-inducible reporter); Ngn1-GFP triple transgenic mice. Mash1CreER is expressed in pluripotent stem cells of the MOE(GBCs)(Rodriguez-Gil et al., 2015), whereas Ngn1-GFP is predominantly expressed in immediate neuronal precursors (INPs)(Fletcher et al., 2017) and immature OSNs (iOSNs)(Magklara et al., 2011). This genetic strategy allows also isolation of newly differentiated mature OSNs (mOSNs), which turned off Ngn1-GFP but still have detectable levels of GFP protein (Figure 1A). We injected P4 mice with tamoxifen, inducing permanent Tomato expression, and collected cells 48 hours later (Figure 1A). To assure that we could collect sufficient numbers of cells from each population we dissected the developing MOE into pieces containing predominantly zones 1-2 and 3-5, rather than dissecting individual zones. From each dissection we isolated four major cellular populations: Tomato^{+}/GFP^{-}, Tomato^{+}/GFP^{+}, Tomato^{-}/GFP$^{+bright}$, Tomato^{-}/GFP^{+dim} (Figure 1A), which represent GBCs, INPs, iOSNs, and

mOSNs respectively. scRNA-seq of these FAC-sorted populations confirmed that our genetic strategy enriches for the 4 major differentiation stages, with INP cells further divided into 3 subpopulations, INP1-3 (Supplemental Figure S1A). Moreover, tSNE plots confirmed in, an unbiased fashion, the developmental relationships of the 4 FAC-sorted cell populations (Figure 1B). OR mRNAs are first detected in INP3 cells (Figure 1C), which express multiple ORs, consistent with previously reported OR co-expression at the INP stage. Surprisingly, if we plot the zonal identity of the most highly expressed OR in each cell, we find that INP3 cells from both zonal dissections predominantly express zone 1/2 ORs, with zone 3-5 OR identities taking over transcriptionally only in OSNs of zone 3-5 dissections (Figure 1C).

To further characterize this intriguing zonal transition, we determined the identity of OR mRNAs in each cell. Consistent with previous reports, mOSNs predominantly express a single OR with a zonal identity that is in accordance with the zonal origin of the cell (Figure 1D). Similarly, iOSNs also express ORs with proper zonal identities, i.e. zone 1-2 ORs in zone 1-2 dissections and zone 3-5 ORs in zone 3-5 dissections (Figure 1E). However, in INP3 cells, where polygenic OR expression is common, we observe a striking departure from zonal restrictions in zones 3-5. Specifically, while INP3 cells from zone 1-2 co-express almost exclusively zone 1-2 ORs, INP3 cells from zones 3-5 co-express ORs from every zone. (Figure1F, Supplemental Figure S1B). As eluted in the tSNE plots (Figure 1C), zone 1-2 OR mRNAs are the most abundant in zone 3-5 INP3s, but zone 3-5 OR transcripts are also detected in most of these cells (Figure 1F). Specifically, we detect zone 1-2 ORs in 43 INP3 cells from zones 3-5, and zone 3-5 ORs in 29 of them, with only 5 cells expressing exclusively zone 3-5 ORs in this stage. This paradox cannot be attributed to dissection errors, since OSNs, both immature and mature, from the same dissections express predominantly zone 3-5 ORs. Thus, although zonal differences are apparent at the early stages of OSN differentiation, commitment to a zonal identity does not take place at the same developmental stages across every zone; INP3 cells are already committed to their zonal identity in zones 1-2, but appear zonally multipotent in zones 3-5, where commitment to a zonal identity occurs later. Finally, to obtain higher zonal resolution in the OR transcriptome of early progenitors, we trisected MOEs and performed bulk RNA-seq in INPs and mOSNs from zones 1, 2/3 and 4/5. Although this analysis does not provide single cell information, it confirms that INPs from each zone express predominantly zone 1 ORs, with an increasing number of zonal identities being added to the INP OR repertoire as we transition to higher zonal indexes (Supplemental Figure S1C). Thus, for each zone, ORs that are most highly expressed in INPs, are lowly expressed in mOSNs from the same zone, and *vice versa* (Supplemental Figure S1D).

Zonal heterochromatin deposition silences earlier zonal OR identities in later zones

We previously showed that OSN differentiation coincides with heterochromatin-mediated OR silencing(Magklara et al., 2011). If OR heterochromatinization is the reason for which zone 1 OR transcription stops in zone 2-5 OSNs, then zone 1 OR genes should have stronger enrichment of heterochromatic marks than the other ORs in the whole MOE. Similarly, zone 2 ORs, which will be repressed in iOSNs of zones 3-5, should have higher levels of heterochromatin than zone 3-5 ORs, and so forth. Indeed, visual inspection of H3K9me3 distribution across mixed OR clusters confirmed that zone 1 ORs have higher levels of H3K9me3 than zone 5 ORs in mOSNs isolated from the whole MOE (Figure 2A). Aggregate analysis of H3K9me3 distribution across all the ORs grouped by zonal identity, corroborates the gradual reduction of this modification from zone 1 to zone 5 ORs (Figure 2B). The only exception from this pattern regards the type I ORs, which although they are expressed in zone 1, they follow their own regulatory principles(Enomoto et al., 2019; Hirota et al., 2007).

Searching for additional histone marks with zonal correlations, we identified an even stronger zonal influence in the distribution of H3K79me3, the enrichment levels of which decline sharply with ascending zonal identities (Figure 2A, C). H3K79me3 is an intriguing core histone modification that may play opposing roles in gene regulation(Farooq et al., 2016). It was originally described as a heterochromatic mark enriched on mammalian pericentric and subtelomeric repeats, with an active role in establishment of heterochromatin and transcriptional silencing (Jones et al., 2008). However, H3K79me3 may also play a role transcriptional activation(Steger et al., 2008), as it is enriched in developmentally regulated and cell type specific enhancers(Bonn et al., 2012; Godfrey et al., 2019), and facilitates long-range enhancer-promoter interactions(Godfrey et al., 2020). Consistent with this, we previously showed that Greek Islands are surrounded by high levels of H3K79me3, is such a specific fashion that it was used as a criterium for the distinction of OR enhancers from other OSN enhancers(Markenscoff-Papadimitriou et al., 2014). Interestingly, the levels of

H3K79me3 and H3K9me3 at the sequences flanking OR enhancers show little variation between zones but follow the general trend of gradual increase with ascending zonal indexes (Supplemental Figure S2A).

A second prediction from our scRNA-seq data is that deposition of heterochromatic marks should be synchronous with the transcriptional commitment to a zonal identity, i.e. should be occurring precisely at the INP to iOSN transition. To define the timing of heterochromatin formation, we performed ChIP-seq for H3K79me3/H3K9me3 on FAC-sorted GBCs, INPs and iOSNs from the whole MOE. This analysis revealed that H3K9 and H3K79 trimethylation is first detected at low levels in INPs, specifically on zone 1 ORs. Since we don't have genetic markers to sort INP subpopulations we assume that the weak signal stems from the most differentiated INP3 population where OR co-expression is also first detected. Consistent with these marks being deposited at more differentiated stages, the levels of both H3K79me3 and H3K9me3 become stronger in iOSNs, the cell stage at which zone 1 ORs are transcriptionally replaced by ORs with the proper zonal identity (Figure 2D, Supplemental Figure S2B). Moreover, as in mOSNs, we detect intermediate enrichment of H3K79me3/H3K9me3 on zone 2 and 3 ORs and lowest levels of heterochromatic marks in zone 4/5 ORs.

The distribution of heterochromatic marks across the OR sub-genome suggests that ORs with higher zonal identities may become increasingly silenced as we transition to higher zones. To test this, we performed ChIP-seq in mOSNs isolated from micro-dissected MOE zones 1, 2/3 and 4/5, as previously described. Interestingly, this analysis shows that the distribution of heterochromatic modifications correlates with both the zonal identity of each OR and the zonal positioning of the mOSN. Thus, in zone 1 OSNs, zone 1 ORs are predominantly marked by H3K79me3/H3K9me3; in zone 2/3 OSNs, zone 1 OR genes have the highest levels of these marks but zone 2 and -to a lesser degree- zone 3 ORs also have strong enrichment; in zone 4/5 OSNs, OR genes with zone 1-4 identities have equally strong H3K79me3/H3K9me3 enrichment, and zone 5 ORs have more H3K79me3/H3K9me3 than in other zones (Figure 2E, Supplemental Figure S2C). A question emerging from these observations is why zone 3 and 5 ORs have lower levels of H3K79me3/H3K9me3 than zone 2 and 4 ORs in zones 2/3 and 4/5, respectively. According to the pattern that emerges from our studies, zone 3 ORs would not be heterochromatic in zone 2 mOSNs, and zone 5 ORs would not be heterochromatic in zone 4 mOSNs. Thus, because these two dissections contain mOSNs from two different zones, the ChIP signal on ORs with the higher zonal identities (zones 3 and 5 ORs) would be diluted. We, therefore, predicted that if we were to interrogate exclusively zone 3 and 5 OSNs, we would find increased levels of H3K79me3 on zone 3 and 5 ORs compared to the levels in the mixed zones 2/3 and 4/5. To test this, we performed H3K79me3 ChIP-seq in FAC-sorted Olfr1507^{+} OSNs (from Olfr1507iresGFP mice), which provide a pure zone 5 OSN population. In these cells, H3K79me3 levels are further increased in zone 5 ORs when compared to the mixed zone 4/5 OSNs and reach the levels of enrichment observed for zone 4 ORs (Figure 2E). Interestingly, the transcriptionally active OR of this population, Olfr1507, has reduced levels of this modification compared to its neighboring zone 5 ORs (Supplemental Figure S2D), consistent with an overall negative correlation between transcription and H3K79me3 enrichment genome-wide.

The demonstration that ORs from lower zonal indexes have high levels of H3K79me3/H3K9me3 in each zone explains why during the INP to iOSN transition these ORs become transcriptionally silenced. However, the realization that also ORs with the proper zonal identity have high levels of these repressive marks appears counterintuitive. It should be emphasized, however, that our zonal ChIP-seq data are obtained from mOSNs, which have completed the process of OR gene choice. Thus, we hypothesized that at less differentiated cells, which have not yet activated a single OR, the levels of H3K79me3/H3K9me3 will be lower on ORs from the associated zone than ORs from a previous zone. To test this, we isolated iOSNs from zone 4/5, and performed H3K79me3 ChIP-seq. Indeed, this experiment shows that the enrichment levels of H3K79me3 on zone 4/5 ORs are lower than the levels observed for these ORs in mOSNs and the levels observed for zone 1-3 ORs in both iOSNs and mOSNs (Supplemental Figure S2E). Thus, although zone 4/5 ORs eventually obtain high levels of heterochromatic marks in their zone of expression, these enrichment levels are reached after the completion of OR choice and only at the ORs that were not chosen for transcription. In contrast, ORs from zone 1-3 with already high H3K79me3 enrichment levels are likely excluded from the process of OR gene choice in this zone. Expanding this pattern for zone 1 and 2/3 dissections, we expect that zone 1 and 2/3 ORs have lower enrichment levels for H3K79me3/H3K9me3 in iOSNs of their respective zones, permitting their transcriptional activation. Thus, it appears that in every zone, ORs from a previous zone obtain high levels of H3K9me3/H3K79me3 before the onset of OR gene choice, while ORs with the proper zonal identity have intermediate levels during OR choice, which further increase after a single OR is transcriptionally activated. This pattern raises a new question: Why

ORs from ensuing zones, which have negligible levels of H3K9med3/H3K79me3, are not chosen in OSNs from preceding zones?

Zonal genomic compartmentalization excludes later zonal OR identities from earlier zones

To understand what keeps ORs from high zonal indexes from being expressed in lower zones, we searched for an orthogonal mechanism to the aforementioned heterochromatin-mediated silencing. OR gene activation requires formation of OR compartments, assembly of a multi-enhancer hub next to the OR compartment, and association of this hub with one of the proximal OR alleles. To test if zonal positioning and identity influence this process, we performed *in situ* HiC in FAC-sorted OSNs from zones 1, 2/3, and 4/5. This analysis revealed that long-range *cis* and *trans* genomic interactions between ORs are strongly influenced by the identity of the OR and the position of the OSN along the zonal coordinates of the MOE, mimicking the deposition of heterochromatic marks. For example, visual inspection of the genomic interactions between 3 OR clusters in chromosomes 2, shows that a cluster that contains mostly zone 4/5 ORs makes contacts with the other clusters only in zone 4/5 OSNs. In contrast, the other 2 OR clusters, which are either enriched for zone 1-3 ORs, or contain ORs from every zonal index, make strong genomic contacts with each other in all three zones (Figure 3A). To expand this analysis to every OR in the mouse genome, we measured the frequency of interchromosomal OR-OR interactions across zone 1, 2-3, and 4-5 ORs in OSNs from the 3 zonal dissections, reaching a similar conclusion: In zone1 OSNs, only zone 1 ORs have strong *trans* interactions; in zone 2/3 OSNs interactions between zone 1 ORs are even stronger but there is significant increase in *trans* contacts between zone 2/3 ORs; in zone 4/5 OSNs there is further increase in *trans* OR-OR interactions between every OR identity, including zone 4/5 ORs (Figure 3B). Intriguingly, this analysis suggests that the group of ORs that can be chosen for expression in each zone has almost identical normalized frequencies of *trans* HiC contacts within the group (Figure 3B). Similar observations are observed in iOSNs, with differences in genomic contacts between different OR identities being sharper than in mOSNs, especially in *trans* (Supplemental Figure S3A, B). Interestingly, at this developmental stage there is a strong correlation between the levels of H3K79me3 and the frequency of HiC contacts (Supplemental Figure S3C), suggesting that the two processes may be mechanistically linked, as recently proposed for heterochromatin and genomic compartmentalization.

We then asked if there are zonal differences between the long-range contacts made by Greek Islands. Consistent with the small variations in flanking H3K79me3/H3K9me3, t*rans* HiC contacts between Greek Islands are comparable between zones 1 and 4/5 mOSNs, with a small increase towards zone 4/5 (Figure 3C). One exception is detected on enhancers residing in OR clusters that contain predominantly zone 4-5 ORs. As these OR clusters make very infrequent contacts with other OR clusters in zone 1 (Figure 3A), their local enhancers also have reduced contacts with Greek Islands, consistent with the notion that OR compartmentalization promotes enhancer compartmentalization. Thus, not surprisingly, contacts between Greek Islands are stronger in zone 4/5 than zone 1 iOSNs, following sharper differences OR compartmentalization and H3K79me3 deposition between the two zones (Supplemental Figure S3D, E). Given that at the iOSN stage, zone 4/5 ORs have lower levels of heterochromatin than zone 1-3 ORs, it is possible that only interactions between the Greek Island hub and zone 4/5 ORs result in transcriptional activation.

To obtain insight to the zonal differences in OR compartmentalization at single cell resolution, we performed Dip-C, a high coverage version of single cell HiC, in 96 mature OSNs, 48 each from zones 1 and 4/5. To retain allelic information, we crossed OmpiresGFP mice (mixture of Bl6/123 strains) to Castaneous (Cas) mice and used Cas-specific SNPs to distinguish Cas from non-Cas alleles. After filtering for high quality cells, we excluded 4 cells and generated contact maps from a median of over 400,000 contacts per cell. We used the haplotype resolved data to compute distances of all genomic loci at 20kb resolution and generated 3D models for all cells, as previously described(Tan et al., 2019). In these models, we were able to detect formation of interchromosomal compartments in mOSNs from both zones (Figure 4A, B). Analyzing distances between pairs of loci in the 3D model we were able to determine the size and complexity of OR compartments in each cell. Consistent with predictions from our bulk HiC, zone 4/5 OR compartments are larger and contain more ORs from more chromosomes that zone 1 compartments (Figure 4C, D, Supplemental Figure S4A-D). Interestingly, despite the fact that we can distinguish zone 1 from zone 4/5 mOSNs based on OR compartmentalization, within a zonal mOSN population every cell has a distinct patter of long range *cis* and *trans* OR contacts Supplemental Figure S4A, B). As previously shown, Greek Islands have stronger interchromosomal contacts than OR genes, that are less influenced by zonal identity than OR-OR interactions (Supplemental Figure S4E, F). Thus, in a

given zone4/5 OSN, on average around 27% of enhancers were engaged in interchromosomal contacts within 2.5 particle radii compared to 22% in zone1. But the size of the largest hub (ranging from 4-8 enhancers), presumably the site of OR transcription, was comparable in the zone1 and zone4/5 cells (Supplemental Figure S4G). Importantly, although Greek Islands make strong contacts with ORs from every zone in both mOSN types, in zone 4/5 they have a significant preference for zone 4/5 ORs compared to zone 1, and zone 2/3 ORs. We do not detect a significant association with zone 1 ORs in zone 1 mOSNs, most likely because almost half of these cells express class I ORs, which are not regulated by the multi-enhancer hub.

Differential expression of NFI paralogues implicates them in zonal OR gene regulation

The realization that both deposition of H3K9me3/H3K79me3 on OR genes, and the increase in OR compartmentalization occur gradually during the transition from zone 1 to zone 5 has important regulatory implications. Instead of deploying distinct, zone-specific transcription factor combinations, developing OSNs could establish zonal restrictions with shared factors expressed in a graded fashion. The transcription factors with an established role in OR gene choice, Ebf 1-4 and Lhx2, do not exhibit zonal differences in expression at any developmental stage, whereas Ldb1 has somewhat lower expression in zone 1 (Supplemental Figure 5A). Thus, we searched for additional transcription factors that may specifically contribute to the zonal restrictions of OR gene choice. Taking into account the early developmental onset by which the aforementioned zonal differences are established, we searched for transcription factors that meet two criteria: their expression levels in both INPs and iOSNs are significantly different between zone 1 and zone 4/5, while their expression in zone 2/3 is intermediate between the two other zones. From the 19 transcription factors that are differentially expressed between zones 1 and 4/5, only three satisfy both criteria, NFIA, B, and X (Figure 5A-D). These three members of the nuclear factor I (NFI) family of transcription factors control a plethora of developmental and cell specification processes(Gronostajski, 2000; Zenker et al., 2019) and were previously implicated OSN differentiation(Baumeister et al., 1999; Behrens et al., 2000). are expressed at higher levels in zone 4/5 cells, from INP to mOSN, with NFIA and NFIB being expressed higher in INPs and iOSNs, and NFIX being upregulated in mOSNs (Figure 5E). Immunofluorescence confirmed that NFIA protein is absent from the neuronal lineages of zone 1 while it is detected in zone 5 INPs and iOSNs, but it also revealed that it is expressed in Sustentacular cells in a non-zonal fashion (Supplemental Figure 5B).

To explore whether NFI gene expression differences between zones has functional consequences in genome occupancy, we performed omni-ATAC(Corces et al., 2017) on iOSNs from zones 1 and 5. This experiment revealed the existence of >5,000 zone 4/5 specific peaks that have predicted NFI DNA binding sequences as the most significantly enriched motifs (Figure 5C). If we focus our omni-ATAC analysis on OR promoters, which have weak ATAC signal that does not reach peak-calling threshold, we detect stronger signal on zone 1 promoters in both zone 1 and 4/5 iOSNs. Interestingly, as it is the case for H3K79me3/H3K9me3 distribution and OR compartmentalization, ATAC signal increases across most OR promoters in zone 4/5 iOSNs, expanding to zone 4/5 promoters, which have the weakest signal in both zones. Consistent with an increase of NFI expression in zone 4/5 iOSNs, NFI motifs are twice as frequent on zone 4/5 than zone 1 promoters (Supplemental Figure S5C). In contrast zone 5 promoters have low quality Ebf motifs, suggesting the possibility that NFI binding compensates for weaker Ebf binding on these sequences (Supplemental Figure S5D). To further explore the possibility that NFI proteins bind directly on OR promoters we performed CUT&RUN(Meers et al., 2019a; Meers et al., 2019b) with an antibody against NFIA. Although we obtained ~2,000 peaks genome wide, with NFI motifs significantly enriched in these peaks, we did not detect NFI binding on zone 4/5 OR promoters (data not shown). This, however, is not surprising, since due to cellular heterogeneity none of the transcription factors known to regulate OR expression can be detected on OR promoters (Monahan et al., 2017). Interestingly, however, we detect NFI signal on Greek Islands despite the lack of full NFI motifs on these enhancers (Supplemental Figure S5E). Consistent with binding of additional factors, like NfI proteins, to the core of Lhx2/Ebf/Ldb1 in zone 4/5, ATAC-seq signal is increased on Greek Islands, especially the ones that reside on OR clusters the contain exclusively zone 4/5 ORs (Supplemental Figure S5F-G). These observations, taken together, suggest that NFI proteins may contribute to the aforementioned zonal differences of heterochromatin assembly, OR compartmentalization, and Greek Island hub formation, as well as in the selectivity by which Greek Islands interact with each other and with ORs with various zonal identities, determining which ORs can be expressed in zone 4/5 OSNs.

NFI A, B, and X regulate zonal OR expression

To interrogate the potential role of NFI A, B and X in zonal OR expression, we adopted a loss-of-function approach. To overcome possible redundancy between the NFI paralogues, we deleted all three genes simultaneously. To assure complete removal of NfI activity before the iOSN stage, where zonal differences are first detected, we used the Krt5CreER drive. This Cre driver is expressed in HBCs, the quiescent stem cells of the MOE. We crossed Krt5CreER; TdTomato mice to NFIA,B,X fl/fl mice(Clark et al., 2019), and induced Cre-mediated recombination with tamoxifen. To force the quiescent HBCs to differentiate into OSNs we ablated the MOE with methimazol and allowed 40 days for a complete restoration by the marked progeny of the NFI KO or control HBCs (Supplemental Figure S6A). To explore the role of NfI proteins in zonal OR expression, we first performed IF in MOE sections of control and triple NFI KO mice using antibodies for ORs Olfr1507(zone 5 OR) and Olfr17 (zone 2 OR). This analysis revealed a strong reduction in the number of Olfr1507-expressing OSNs (Figure 6A), and ectopic expression of Olfr17 in zone 5 OSNs (Figure 6B). To explore the extent of this zonal transcriptional transformation we FAC-sorted Tomato$^+$ control and KO OSNs from the whole MOE and performed RNA-seq. This analysis shows strong reduction of all zone 4/5 OR mRNAs and a concomitant increase of zone 2/3 mRNAs in the triple KO OSNs (Figure 6C).

The observations afforded by IF and whole MOE RNA-seq analyses suggest that zone 2/3 ORs substitute zone 4/5 ORs in their cognate zone. To directly test this prediction, we performed RNA-seq analysis in Tomato$^+$ OSNs that were FAC-sorted specifically from zones 4/5. This experiment confirms the strong transcriptional downregulation of every zone 4/5 OR, and the frequent ectopic expression of zone 2/3 ORs in zone 4/5 (Figure 6D). This transcriptional transformation can be defined as homeotic, since the general mOSN identity is not altered by the triple NFI deletion: only 13 out of ~200 OSN-specific markers are significantly different between control and KO OSNs (Supplemental Figure S6B). Moreover, from the ~207 zone 4/5-specific markers, 90 are significantly downregulated in the triple KO, suggesting that NFI proteins control additional components of zonal identity (Supplemental Figure S6B). Importantly, detection of most zone 4/5-specific markers in the triple KO provides independent confirmation for the accuracy of our zone 4/5 dissections and suggests that additional transcription factors may contribute to non-OR related zonal gene regulation.

Having established a robust phenotype from the triple NFI KO, we explored how redundant the three NfI proteins are. In this note, we analyzed separately double (NFIA, B) and single (NFIX) KO OSNs generated with the same strategy described for the triple KO. Intriguingly, in double (NFIA, B) KO mice, zone 4/5 OSNs express predominantly zone 3/4 ORs instead of zone 2/3 ORs (Figure 6D). Although expression of zone 5 ORs dissipates in the double KO, only 37% of zone 2 ORs are significantly upregulated in zones 4/5 and at lower levels than the triple KO (Figure 6D, Supplemental Figure S6D). In contrast the effects of the double vs triple deletion are almost identical for zone 3 ORs (Supplemental Figure S6D). Importantly, the majority of zone 4 ORs are not affected by the double deletion, suggesting that NfiX is sufficient to preserve zone 4 OR expression in zones 4/5 (Supplemental Figure S6D). Finally, in the single NfiX KO mice we do not detect significant transcriptional changes: zone 4/5 OSNs still express zone 4/5 ORs (Figure 6D), with little variability between control and KO dissections (Supplemental Figure S6D). The severity of the zonal transformation in OR expression correlates with the overall change in zonal OSN markers (Supplemental Figure S6C). Thus, although there is redundancy between the 3 NfI proteins, it appears that different NFI genes have different contribution to the establishment of zonal boundaries according to their sequence, levels and, possibly timing of expression, since NFIX is lowly transcribed in INPs. Reciprocally, ORs with different zonal identities exhibit differential sensitivity to the reduction of NFI gene levels, with zone 2 and 4 ORs requiring a triple Nfi deletion, while zone 3 and 5 ORs being mis-regulated even when only two NFI genes are deleted.

NFI A, B, and X govern heterochromatin deposition and OR compartmentalization

We searched for regulators of OR expression that have gradient expression across the 5 zones, instead of zone-specific transcription factor combinations, because both OR heterochromatinization and compartmentalization progressively increase and expand across increasing zonal indexes. Thus, upon establishing a key role of NFI genes in zonal OR expression we asked if NFIA, B and X regulate H3K79me3/H3K9me3 deposition and OR compartmentalization. Deploying the genetic strategy described previously (Supplemental Figure S6A), we isolated control and 3ple NFI KO Tomato$^+$ OSNs from zone 4/5 dissections, which were used for ChIP-seq and *in situ* HiC. ChIP-seq revealed a striking transformation in the

pattern of H3K79me3/H3K9me3 across zonal OR identities with strong reduction of both histone modifications in zone 3-5 ORs, and little change in zone 1-2 ORs (Figure 7A, B), mimicking the pattern observed in zones 2/3 (Supplemental Figure S7A). Furthermore, if we group zone 2 and 3 ORs in accordance to the significance of their upregulation in zones 4/5 (Figure 7C), we make an intriguing observation in regards of H3K79me3 enrichment: zone 2 ORs that are most significantly upregulated have the lowest levels of H3K79me3 between zone 2 ORs, whereas for zone 3 ORs the trend is the opposite (Figure 7D). Thus, as it was already suggested by our zonal analysis in wild type MOEs, there is a non-linear correlation between levels of H3K79me3/H3K9me3 enrichment and transcriptional outcomes, with intermediate enrichment levels being "optimal" for OR gene choice.

If levels of H3K79me3/H3K9me3 enrichment influence OR compartmentalization (or vice versa), then we would expect strong changes in long-range genomic interactions upon NFI deletion. Indeed, *in situ* HiC in control and 3ple Nfi KO Tomato^{+} OSNs from zone 4/5 revealed strong reduction in the long-range genomic contacts made by zone 4/5 ORs. This reduction applies to both long range *cis* contacts between OR gene clusters (Figure 7E), but also for genomewide *trans* interactions (Figure 7F). In fact, direct comparison of the patterns of zonal interchromosomal interactions between the three dissected zones suggests that 3ple Nfi KO OSNs from zone 4/5 adopt a nuclear architecture that is most similar with the nuclear architecture of zone 2/3 OSNs (Supplemental Figure S7B). In support of this, if we plot interchromosomal OR interactions across a large mixed cluster in chromosome 2, we find that the distribution of *trans* OR contacts between ORs from different zones is quite different between zones 2/3 and 4/5; yet in the triple Nfi KO, we find that this distribution is transformed in 4/5 OSNs to an almost identical pattern with zone 2/3 OSNs (Supplemental Figure S7C). Finally, we asked if triple NFI deletion affects the frequency of *trans* contacts between Greek Islands. Overall, *trans* contacts between Greek Islands are not reduced in triple KO (Supplemental Figure S7D), consistent with the fact that we did not observe strong zonal differences for these genomic contacts. However, Greek Islands that reside on OR clusters that contain exclusively zone 4/5 ORs exhibit a reduction on trans contacts with other OR enhancers, consistent with the effects of NFI deletion in the genomic contacts of zone 4/5 ORs. Finally, CUT&RUN analysis in triple NFI KO and control OSNs shows reduction of the NFI signal on Greek Islands (Supplemental Figure S7E), but not complete elimination, most likely due to the fact that Greek Islands remain highly accessible. The lack of consensus NFI motifs from these enhancers, however, and our inability to detect NFI binding on Greek Islands by crosslinked ChIPs (data not shown), suggests that the CUT&RUN signal may reflect indirect recruitment of these factors via protein-protein interactions with Lhx2/Ebf/Ldb1. In either case, our data show that NFI proteins establish chromatin modification and nuclear organization landscapes that repress zone 2/3 ORs and enable the choice of zone 4/5 ORs in the developing zone 4/5 OSNs.

Discussion

Our experiments revealed a mechanism by which a random process becomes stochastic, transforming the relative position of a cell across the MOE axes into biased OR gene choices. The solution to the perplexing segmentation of the MOE into overlapping zones is stunningly simple: a gradual increase of heterochromatin deposition and genomic compartmentalization from the first to the last zone, alters, in a sequential fashion, the OR repertoire that can be activated in each OSN. ORs that are activatable in one zone, are silenced in subsequent ones, while ORs that cannot interact with the Greek Island hub in this zone, acquire this privilege in later MOE segments. This serial progression is conceptually reminiscent of the fascinating phenomenon of spatial collinearity in Hox gene regulation(Lewis, 1978; Noordermeer and Duboule, 2013), albeit with an important distinction: the order of Hox gene activation is determined by their linear order on the chromosome, whereas here this order is defined by the sequential OR recruitment in 3-dimensional genomic compartments. Consequently, whereas Hox transcription is highly stereotypic within segments, the activation of an OR allele with proper zonal credentials is not. As our DIP-C revealed, OR compartments have highly variable constitution, with the 3D maps of OR contacts being different even between OSNs from the same zone. This immediately suggests that OR compartments and Greek Island hubs assemble in an “opportunistic” fashion, influenced by both the relative affinities between OR loci and their proximity in the nuclear space. While the former is determined by the position of the OSN, the latter is random and dependent on the relative position of each chromosome in the interphase nucleus. With OR gene clusters being located in 18 different chromosomes, it is not surprising that the constitution of OR compartments and Greek Island hubs differs from cell to cell. Thus, although a mechanism evolved to establish some semblance of order in OR gene regulation, the concept of transcriptional activation via interchromosomal associations renders chance a key operating principle.

A quantitative model for opposing heterochromatin functions

The focus of this work was to determine the mechanism that skews OR gene expression to a specific zone. We showed that the concentration of NFI proteins in OSN progenitors determine the repertoire of ORs that can be expressed in mOSNs via regulation of heterochromatin assembly and nuclear architecture. A possible explanation of the homeotic transformations occurring upon NFI deletion is that different levels of H3K79me3/H3K9me3 enrichment have different consequences in chromatin compaction, genomic compartmentalization, and transcriptional activation. We propose that high levels of H3K79me3/H3K9me3 induce both chromatin compaction and formation of OR compartments; intermediate levels of H3K79me3/H3K9me3 suffice for OR compartmentalization but not for chromatin compaction, while low levels of these modifications do not promote either. Consequently, in each OSN ORs from earlier zonal indexes are not expressed because they are actively silenced by constitutive heterochromatin and ORs from subsequent zones are not expressed because they cannot associate with other ORs and, eventually, with the Greek Island hub. Thus, only ORs with the correct zonal identity are both in a partially accessible chromatin state and have the histone modification signature that enables interactions with Greek Islands.

The aforementioned model assumes that H3K79me3/H3K9me3 have independent functions in the regulation of chromatin compaction and 3D organization. The former likely depends on high density deposition of tri-methyl marks across numerous successive nucleosomes(Canzio et al., 2011; Canzio et al., 2013). This would allow Hp1 proteins and other components of heterochromatin to polymerize on the chromatin fiber, insulating regulatory sequences from transcriptional activators Canzio et al., 2013). The latter, which may be organized by homotypic chromatin contacts(Gibson et al., 2019; Narlikar, 2020) and the phase separation properties of heterochromatin(Mirny et al., 2019), could be accomplished at intermediate levels of enrichment that do not induce chromatin compaction. Consistent with this, recent reports provide functional evidence for a role of H3K79me3/me2 in the establishment of long-range enhancer-promoter interactions(Godfrey et al., 2020) and role of H3K9me3 in establishment of genomic compartments *in vivo* (Falk et al., 2019). Importantly, the differences in H3K79me3/H3K9me3 enrichment and frequency of HiC contacts are strongest in iOSNs during the process of OR gene choice. This suggests that there is a transient “bivalent” chromatin modification state during which converging ORs with intermediate levels of heterochromatin can either become fully heterochromatic and transcriptionally silent or associated with Greek Islands and transcriptionally chosen. The intriguing observation that H3K79me3 levels on ectopically expressed zone 2 ORs are reduced, while on

ectopically expressed zone 3 ORs are increased, is consistent with the notion of a “sweet spot” of heterochromatin that enables activating genomic interactions instead of inducing gene silencing.

In addition to the putative role of histone modifications in the establishment of repressive or activating genomic interactions, the repertoire of transcription factors bound on OR promoters and enhancers could also influence genomic compartmentalization. Increased abundance of NFI motifs on zone 4/5 OR promoters and detection of NFI proteins on Greek Islands by CUT&RUN are consistent with an NFI-dependent re-direction of the Greek Island hubs towards zone 4/5 ORs though homotypic interactions. Such homotypic interactions by DNA-bound transcription factors were recently shown to promote assembly of specific interchromosomal compartments in yeast(Kim et al., 2019; Kim and Shendure, 2019). That said, we have not demonstrated binding of NFI proteins on zone 4/5 OR promoters and we cannot unequivocally exclude that NFI occupancy of Greek Islands is a reflection of their highly accessible state, since these sequences do not contain *bona fide* NFI motifs and NFI deletion did not completely eliminate the CUT&RUN signal from these enhancers. Thus, indirect mechanisms by which NIF expression influences nuclear architecture and OR gene choice cannot be excluded.

Transcription (factor) mediated silencing as a means of establishing zonal restrictions

An important question raised by our findings regards the mechanism by which heterochromatin is distributed in a zonal fashion across different OR identities. The three types of transcription factors expected to bind on OR promoters, Ebf, Lhx2, and NFI, have thousands additional genomic targets, but they do not induce H3K79/H3K9 trimethylation outside of the OR clusters. However, an intriguing common theme in the distribution of these marks across OR genes and Greek Islands, is the local reduction of H3K79me3/H3K9me3 at the transcription start site (TSS) and core enhancer element, respectively. In fact, in the INP and iOSN stage, which coincide with the onset of heterochromatin assembly, there is stronger signal downstream of the TSS that declines after the transcription end site (TES), whereas on enhancers the flanking signal is symmetrical. This pattern is consistent with a mechanism of transcription-mediated heterochromatin assembly on OR genes and their enhancers, with the former being biased towards the gene body and the latter being bi-directional. This would be consistent with biochemical experiments reporting purification of mammalian elongating RNA polymerase II complexes containing Dot1l, the only lysine methyltransferase known to tri-methylate H3K79me3(Kim et al., 2012). Similarly, transcription-dependent deposition of H3K9me2/3 has been observed in *Sacharomyces pompe* and *Arabidopsis thaliana(Allshire and Madhani, 2018)*, providing precedent for our hypothesis. Notably, transcription mediated H3K79/H3K9 trimethylation assumes that nascent RNA produced by OR promoters and enhancers has distinct features from the rest of the genome. Indeed, as we previously described, OR gene clusters, except class I ORs which do not have H3K9me3/H3K79me3, represent regions of extreme AT-rich content, which correlates strongly with increased levels of heterochromatin (Clowney et al., 2011). Thus, it is possible the Lhx2-, Ebf-, and NFI-mediated transcription of AT-rich sequences induces heterochromatin formation in INPs and OSNs. In support of this, GFP reporter transgenes driven by one of the strongest OSN promoters, the Lhx2-, Ebf-. and NFI-bound OMP promoter(Behrens et al., 2000; Monahan et al., 2017), become transcriptionally silenced when they include OR sequences (Nguyen et al., 2007). Although OMP is a mature OSN marker, we detect weak expression of this gene in late INPs and iOSNs, at the stage of OR heterochromatin assembly. Therefore, the transcriptional silencing of OMP-driven OR transgenes is consistent the hypothesis that low level OR transcription in OSN progenitors induces heterochromatin formation. Finally, in support of this model is the observation that ORs that are most highly expressed in INPs, are not the ORs that are most frequently chosen in mOSNs and *vice versa*. For example, OR Olfr1507, which is the most frequently chosen OR in zone 5, is not detected in INPs isolated from this zone.

If relative promoter strength determines directly, or through a transcription-mediated deposition, the levels of heterochromatic marks on each OR, then why these levels increase for every OR in higher zones? Considering that Ebf and Lhx2 expression levels are homogeneous across zones, additional factors have to be deployed to strengthen OR promoter activity in higher zones. Dependence on such auxiliary factors may relate on the quality and relative arrangement of Lhx2 and Ebf motifs on each promoter. Thus, consistent with our scRNA-seq and omni-ATAC results, we propose that zone 1 promoters have strong enough Lhx2/Ebf binding sites that can induce low transcription and intermediate levels of H3K79me3/H3K9me3 in zone 1 INP3 cells. In subsequent zones, binding of auxiliary factors on zone 1 promoters would further increase their strength, resulting in increased deposition of heterochromatic marks, and silencing of zone 1 ORs in zones 2-5. A direct prediction of this hypothesis is that reducing the binding of Lhx2 and Ebf on a zone 1 OR promoter would result

in ectopic expression of this OR in higher zones. In fact, this is exactly what was reported for the targeted mutation of a pair of Ebf and Lhx2 binding sites from the promoter of Olfr151, which results in ectopic expression of this zone 1 OR in zones 2/3(Rothman et al., 2005). Thus, weakening the binding of the core transcription factors Ebf/Lhx2 reduces the "baseline" activity of a zone 1 OR promoter, rendering the binding of auxiliary factors in zones 2-3 an activating rather than a repressive event. In this note, *bona fide* zone 2-3 promoters, with naturally weaker Lhx2/Ebf binding would require additional auxiliary transcription factors to become transcriptionally enabled via intermediate H3K79me3/H3K9me3 deposition. If these factors are more highly expressed in zones 4/5, or Nfi proteins can also bind on zone 2-3 promoters, then the strength of these promoters will increase further in zones 4/5, resulting deposition of high levels of heterochromatin and transcriptional silencing in subsequent zones. Finally, zone 4/5 promoters with weakest (or completely absent) binding for Lhx2/Ebf, may require highest levels of Nfi proteins to obtain intermediate levels of H3K79me3/H3K9me3, enabling their future choice. The fact that zone 5 ORs are sensitive to both double and triple Nfi deletion, while zone 4 promoters, with stronger Ebf motifs, only respond to the triple NFi deletion, is consistent with this hypothesis. Importantly, there are additional genetic observations supporting this model: zone 5 ORs Olfr1507-1509 cannot be expressed as transgenes without the H enhancer(Serizawa et al., 2005), while zone 1 and 2 ORs can be expressed as minigenes with minimal promoter sequences(Vassalli et al., 2002). Furthermore, deletion of the H enhancer from the endogenous cluster impairs only the expression of zone 5 and not of zone 1 ORs(Fuss et al., 2007; Nishizumi et al., 2007), supporting the notion that OR promoter strength decreases with ascending zonal identity. Finally, the increased frequency of Greek Island interactions and the somewhat higher complexity of Greek Island hubs in zone 4/5 compared to zone 1, could represent a compensatory mechanism for the strong activation of weaker OR promoters.

Zonal OR expression preserves a wide receptive field

The final question we wish to address regards the potential benefit of this unusual regulatory mechanism and the merit of generating zonal boundaries in the olfactory epithelium. It is well established that the zonal topography of OR expression in the MOE is reflected in targeting of OSN axons in the olfactory bulb. Because this topography disintegrates in higher olfactory processing centers(Stettler and Axel, 2009), we propose that the major advantages of zonal OR expression relate to the process of OSN axon guidance, the singular nature of OR expression in mature OSNs, and the distributed representation of every OR identity. Containment of OSNs expressing the same OR at defined anatomical segments of the MOE, facilitates the fasciculation of their axons and reduces the number of different extracellular identities required for the formation of OR-specific glomeruli. In other words, there is no need for 1400 extracellular barcodes for the generation of unique glomeruli, but only 40-200 would be sufficient for the segregation of the axons emerging from a specific zone. Furthermore, limiting the process of OR gene choice to 40-200 ORs instead of 1400 different ORs, reduces the probability of choosing two different ORs at the same time. Finally, the aforementioned process provides a safeguard against the evolution of transcriptionally dominant OR alleles. If an OR promoter was to become too strong, it would be chosen first, suppressing the choice of other ORs. However, with this elegant regulatory mechanism in place, stronger promoters induce early transcription in INPs and accumulation of high levels of heterochromatin outside of zone 1, leading to their "ostracization" in zones 2-5. Of course, this system does not have an unlimited capacity to block transcriptional domination, since artificially powerful promoters can express OR transgenes in high OSN frequencies(D'Hulst et al., 2016; Fleischmann et al., 2008; Nguyen et al., 2007). On the other hand, weaker promoters can still be represented at considerable frequencies, since there are zones dedicated to promoters of various strengths, with zone 5 likely preserving frequent expression of the 40 weakest OR promoters. Importantly, differences in promoter strength become irrelevant in mature OSNs, where accumulation of numerous Ebf/Lhx2/Ldb1 bound enhancers assembles a transcriptional complex that supports equally high transcriptional levels for every chosen OR in each OSN. Thus, by combining developmental restrictions to the OR identities that can be chosen in each OSN, with a random selection from this spatially determined OR repertoire, the olfactory organ expresses thousands of OR alleles at meaningful frequencies, assuring sensitive detection of the highest number of volatile chemicals.

Acknowledgments
The authors have no competing interests for this work. Mice were treated in compliance with the rules and regulations of IACUC under protocol number AC-AAAT2450 and AC-AABG6553. We thank members of the Lomvardas lab for critical discussions and input on the manuscript. We thank Drs. Tom Maniatis and Richard Axel for critical reading of the manuscript. Stavros Lomvardas acknowledges support from the National Institutes

of Health Common Fund 4D Nucleome Program (Grant 5U01DA040582), and the National Institute of Deafness and Communications Disorders (Grant 5R01DC018745). SL was also supported by the HHMI Faculty Scholar Award (#55108540). RMG was supported by NYSTEM contracts C030133 and C30290GG. GB was supported by 5R01DC013561 (NIH).We are in the process of depositing sequencing data at GEO (RNA-seq, scRNA-seq, ChIP-seq, Omni-ATAC, and CUT&RUN), and the 4DN portal (for HiC and Dip-C).

Materials and Methods

Zonal Annotation:
Zonal annotation from "A Near-Complete Spatial Map of Olfactory Receptors in the Mouse Main Olfactory Epithelium"(Tan and Xie, 2018). We generated bins from the continuous data rounded to the nearest integer. There were a total of 1011 ORs with known zonal annotation. Of these 115 are ClassI ORs, of which nearly all are expressed in zone1 and 896 ClassII ORs. Of these 261 are expressed in zone1, 283 in zone2, 164 in zone3, 144 in zone4 and 44 in zone5.

Zonal Dissections:
We used Olfr16iresGFP (zone 1 OR), Olfr17iresGFP (zone 2 OR), and Olfr1507iresGFP (zone 5 OR) to practice dissections of zones 1, 2/3, and 4/5, respectively. Upon obtaining an accurate understanding of the zonal boundaries in the MOE we performed zonal dissections without the use of these fiduciary markers. Accuracy of dissections was confirmed by RNA-seq.

Single cell RNA-seq
Mash1-CreER; tdTomato; Neurogenin1-GFP pups were injected with tamoxifen at P4 and olfactory epithelium collected after 48 hours. Tissue was dissected into ventral (mostly zone 4-5) and dorsal OE (mostly zone1) sections and from which cells were sorted into 384 well plates with unique cell and molecular barcodes. Library preparation and sequencing was performed in collaboration with the New York Genome Center (NYGC) using a TSO approach similar to SMARTer cDNA synthesis and sequenced on HiSeq2500. Reads were aligned according to the drop-seq platform using STAR for alignment and discarding multi mapped reads. Experiment was performed in biological replicates. Out of 764 cells, we discarded cells with less than 1000 genes and 20,000 UMIs, resulting in 669 cells and further filtered for cells that contained less than 5% mitochondrial reads, resulting in 591 cells used for analysis. Clustering was performed in Seurat (TMI: used filtering default of a gene being expressed in at least 3 cells to include in data for clustering and differential expression analysis as it is a cleaner way to look at cell-type specific genes, but turned the filtering default to one 1 cell for OR expression analysis). For all OR expression analysis we used a threshold of 3UMI for an OR to be considered expressed.

Seurat clustering resulted in 5 populations. To identify markers for each cluster we performed differential expression analysis using, Wilcoxon rank sum test. We looked at expression of selected olfactory lineage markers at different differentiation states to annotate the cell type. The heatmap in supplementary figure1 plots the logFoldChange for the different markers relative to the average expression between the two groups. From this analysis we determined that Mash1+Neurog1^{neg} cells correspond to cycling GBCs. Mash1+Neurog1^{bright} cells correspond to the INP1-3 populations, of which INP1 is high for cell cycle markers. Mash1^{neg}Neurog1^{bright} are predominately iOSNs but also contain some INP3 cells, and Mash1^{neg}Neurog1^{dim} cells correspond to OSNs.

Bulk RNAseq in OE cell types:
Mash1-CreER; tdTomato; Neurogenin1-GFP pups were injected with tamoxifen at P4 and olfactory epithelium collected after 48 hours. Tissue was dissected into a ventral (mostly zone 4-5), dorsal OE (mostly zone1) and a central section (that is enriched for zone2-3). Cell-types were sorted by the fluorescent markers into 4 populations: Tomato^{+}/GFP^{-}, Tomato^{+}/GFP^{+}, Tomato^{-}/GFP$^{+bright}$, Tomato^{-}/GFP^{+dim}, that we will refer to as GBC, INP, iOSN and OSN, based on single cell RNAseq analysis. RNAseq libraries were prepared with Nugen NuQuant RNA-seq library system and sequenced PE on HiSeq2500 and NextSeq. Reads were aligned with STAR and normalization and differential expression analysis was performed with DEseq.

To find zone5 enriched transcription factors we found significantly differentially expressed transcription factors (from the GO database) between zone5 and zone1 cells with an adjust p value less than 0.05 and at least a two fold change in expression. There were 30 DE TFs in GBCs, 31 in INPs, 44 in iOSNs and 18 in OSNs. To get

the most likely candidates driving zonal identity we further selected TFs that were expressed with a TPM of at least 10 in their respective zone, of which there were 15 in GBC, 21 in INP, 24 in iOSNs, and 8 in OSNs. NFI transcription factors were among the group of zonal TFs at every developmental stage.

Mature zonal markers vs non-zonal markers:
To find mature zone5 markers, we looked at non-OR genes differentially expressed with adjusted p value less than 0.05, and at least a two fold change in expression, and an expression of at least 5 TPM in zone5 and found 128 genes upregulated between zone5 and zone1, and 58 genes differentially expressed between zone5 and zone2. From this we made a list of 139 genes that were upregulated in zone5 relative to either zone1 or zone2.

Zone5 early NFI knockout RNA-seq
Triple knockout (NFI A,B,X flox/flox, tdTom, OMP-gfp, Krt5-CreER), AB only knockout (NFI A,B flox/flox, tdTom, OMP-gfp, Krt5-CreER), X only knockout (NFI X flox/flox, tdTom, OMP-gfp, Krt5-CreER) or wt (tdTom, OMP-gfp, Krt5-CreER) adult mice were injected with tamoxifen 3x and 10 days later olfactory epithelium was ablated with Methimazole. After allowing the OE to rebuild for 30 days we collected zone5 OMP-gfp+ cells for RNAseq in the 4 genotypes. Libraries were made with Nugen Nuquant RNA-seq kit and sequenced on Nextseq.

Volcano plots:
To test if mature markers change in Nfi knockout zone5, we made a list of top 200 significantly differentially expressed genes that increased (with an adjusted p value < 0.05, and change greater than 2 fold) in Ngn1(GFP)^{dim} cells compared to Ngn1(GFP)^{bright} cells. A second condition for non-zonal mOSN markers was that they were not significantly differentially expressed between zone5 and z1 Ngn1(GFP)^{dim} cells. We also made a list of zone5 enriched mature markers that were significantly differentially expressed between z5 and z1 or z2 Ngn1(GFP)^{dim} cells (adjusted p-value less than 0.05 and greater than 2 fold change), of which there were 208. And similarly made a list of z1 or z2 Ngn1(GFP)^{dim} enriched markers, of which there were 141 genes. For the zone5 markers there were 90 differentially expressed genes out of 207 total. For the zone1 markers there were 29 differentially expressed out of 138 total. And for the OSN markers there were 13 differentially expressed out of 200.

Native chip:
Native chip was performed as described in Monahan et al. 2017. Libraries were prepared with NuGEN Ovation V2 DNA-Seq Library Preparation Kit, and sequenced 50PE on PE on HiSeq2500 and NextSeq or 75PE on NextSeq respectively.

Omni ATAC:
Performed omni-ATAC seq in replicate from zonally dissected iOSNs (Mash1-CreER negative, Ngn-GFP bright positive, i.e. Tomato^{-}/GFP$^{+bright}$ cells). Aligned with bowtie, peak-calling with HOMER and differential peak analysis with Diffbind. There were 14,169 peaks differentially expressed between Z1Ngn and Z5Ngn. Of these 8,635 were Z1 differentially enriched and 5,534 were Z5OMP differentially enriched with P-value 0.05. We ran Homer for ne-novo motif discovery.

In situ Hi-C:
In situ Hi-C was performed on FAC-sorted cells as described in Monahan et al. 2019. Cell inputs ranged from approximately 150,000 to 500,000 cells and each experiment was performed with two biological replicates. Hi-C library preparation was also performed as described in Monahan et al. 2019 and prepared Hi-C libraries were sequenced 75PE on NextSeq 500. Raw fastq files were processed using the Distiller pipeline. Analysis was performed on data pooled from two biological replicates, after confirming that the results of analysis of individual replicates were similar. There were an average of 370 million reads per sample, and all data analysis of OR gene clusters was performed on contact natrices binned at 50kb.

Dip-C analysis:
To isolate mature olfactory sensory neurons, Castaneous (Cas) mice were crossed to OMPiresGPF mice. MOE was collected from adult mice resulting from this cross, dissected into zones and FAC-sorted to isolate mature neurons. Dip-C was performed as described in Tan et al. 2019 on 96 mature OSNs: 48 each from zones 1 and zones 4/5. Libraries were sequenced 150PE on NextSeq 500. Aligned reads were processed

according to the Dip-C pipeline (Tan et al., 2019). Since OMPiresGFP mice were a mixture of Bl6/129 strains, we only included SNPs that were unique to Cas mice to distinguish homologs. After alignment, cells were filtered using several quality control metrics described in Tan et al. 2019: We excluded cells that had less than 20,000 reads, cells that had a low contact-to-read ratio, and cells that had a high variability in 3D structure across computational replicates. Only 4 of 96 cells failed these metrics. Overall, the median number of contacts across cells was over 400,000. Computational analysis, including computing average contact densities of certain pairs of regions, and generation of 3D genome structures, was performed using the Dip-C pipeline. Pairwise distance heatmaps were reordered using hierarchical clustering.

Mice:

Mice were treated in compliance with the rules and regulations of IACUC under protocol number AC-AAAT2450 and AC-AABG6553. Mice were sacrificed using CO_2 following cervical dislocation. All the mouse strains described here were described in detail in (Monahan et al., 2019) except of the NFI KO strains which were described in (Clark et al., 2019)

Antibodies:

Olfr17 antibody were raised in rabbits against epitope RRIIHRTLGPQKL located at the C-terminus of the OR protein. Olfr1507 antibody was described in (Barnea et al., 2004). NFIA antibody is from Active Motif (39397). The following antibodies were used for native ChIP: H3K79me3 (abcam ab2621) and H3K9me3 (ab8898).

Figure Legends

Figure 1: Single cell RNA-seq analysis reveals that developing olfactory progenitors undergo a timed transition in their repertoire of OR expression.

(A) Schematic of the olfactory epithelium (OE) and the genetic strategy for isolating four cell populations along the developmental lineage from the same tissue. Mice heterozygous for Mash1-CreER, TdTomato (a cre inducible reporter) and the Ngn1-GFP transgene were injected at P4 with tamoxifen. After 48 hours the tissue was microdissected into crude zones (zone 1-2 and zone 3-5), and FAC-sorted to isolate 4 cell populations: $Tomato^{+}/GFP^{-}$, $Tomato^{+}/GFP^{+}$, $Tomato^{-}/GFP^{+bright}$, $Tomato^{-}/GFP^{+dim}$ that were used for bulk and single-cells RNAseq. A representative FACS plot is shown.

(B) tSNE clustering of FAC-sorted cell populations with Seurat based on the most variable genes shows separation of single cells into 6 populations, which we associated with cell identities using expression of known OE markers (See Sup1). Clustering shows the relationship between the sorted populations (left) and lineage (right).

(C) tSNE clustering from B, separated by zonal origin colored according to most expressed OR.

(D) Normalized OR expression (normalised UMI counts) within single cells from zone 1-2 tissue (D,E,F) and zone 3-5 tissue (G,H,I). While mOSN (D,G) and iOSN (E,F) cells from both zones are committed to their respective zonal OR repertoires and generally express one dominant OR gene, INP3 cells (F,I) co-express many ORs (up to 17 in our data). Unlike cells from zone 1-2, INP3 cells from zone 3-5 (I) are not restricted in their zonal repertoire and can express ORs from all 5 zonal indexes in the same cell.

Sup 1: OR genes highly expressed in progenitors are less frequently chosen in mature OSNs.

(A) Heatmap shows the scaled expression of known olfactory lineage markers in single cells ordered into columns by their t-SNE cluster identity.

(B) Bulk RNA-seq of cell populations isolated using the same genetic strategy outlined in Figure 1A dissected into zone 1 (red OE schematic), zone 2/3 (yellow/green OE schematic), and zone 4/5 (blue OE schematic). While INP cells from all zones are biased toward expressing zone1 ORs, there is increased expression of higher zonal indexes in zone 2/3 and zone 4/5 cells. mOSN cells from the same zonal tissue are committed to their "correct" zonal OR repertoire.

(C) Aggregate OR expression per zonal cell population in single cells (normalized for the number of cells per category) shows OR expression in both zones begins in INP3 cells, where zone 3-5 cells have more total OR transcription. Cells from zone 3-5 undergo a commitment to their correct zonal OR repertoire at the iOSN developmental stage.

(D) Plots of expression in INP and mOSNs from the RNA-seq data in B shows an anticorrelation between ORs preferentially expressed in INP and mOSN cells.

Figure 2: Heterochromatin deposition silences OR genes from lower zones.

(A) Signal tracks of H3K9me3 and H3K79me3 native ChIP-seq from the whole MOE over two representative OR clusters: one on chr 14 (top) and a portion of a larger OR cluster on chr 2 (bottom). Signal is shown in reads per 10 million. Below the signal track OR genes are colored according to their zone of expression: zone1 ORs in red, zone2 ORs in yellow and zone 5 ORs in blue. Both histone marks appear deposited in a zonal fashion and can also be seen flanking the H Greek Island enhancer (purple, top).

(B) Signal density of H3K9me3 native ChIP-seq in the whole MOE is calculated for all OR gene bodies and scaled to the size of the OR gene. Box plots of signal density over OR genes separated by their zone of expression depict a pattern of deposition that is high on zone 1 OR genes, progressively decreases with higher zonal OR indexes, and is absent on class I ORs.

(C) The same analysis as in B but for H3K79me3 native ChIP shows an even stronger zonal pattern of deposition.

(D) H3K79me3 native ChIP seq in GBC, INP, iOSN and mOSN populations (isolated with FAC-sorting cells from transgenic mice described in Figure 1A show onset of H3K79me3 deposition in INP cells. Each row of the heatmaps shows coverage over an OR gene body (separated into categories by their zone of expression) rescaled to 6kb and 2kb flanking on each side. Heatmaps are sorted in the same order and the average signal per zonal category is shown above (colored by zonal annotation), in reads per 10 million.

(E) H3K79me3 native ChIP-seq in mOSNs isolated from zonal dissections of the MOE in OMP-ires-GFP mice, and Olfr1507+ mOSNs isolated by FAC-sorting MOE from Olfr1507-ires-GFP mice. Colored MOE schematics are shown above each heatmap (red for zone1, yellow/green for zone 2/3 and blue for zone 4/5) depicting the zone of dissection.

Supplementary Figure 2: Progressive accumulation of heterochromatin on OR genes occurs in developing OSNs.
(A) H3K9me3 and H3K79me3 native ChIP-seq signal over 63 OR enhancers (Greek Islands) in mOSNs from three zonal dissections (red for zone1, yellow/green for zone 2/3 and blue for zone 4/5 (as described in Figure2). Heatmaps are centered on Greek Islands and show 6kb flanking on either side. High levels of H3K79me3 and H3K9me3 are flanking the Greek Islands in mOSNs from all zones.
(B) Analogous analysis to Figure 2D for H3K9me3 native ChiP-seq shows a similar onset of deposition in the olfactory lineage.
(C) Analogous analysis to Figure 2E for H3K9me3 native ChiP-seq shows a similar zonal pattern of deposition over OR genes in mOSNs isolated from different zones.
(D) Comparison of H3K79me3 native ChIP-seq in zone 4/5 iOSNs and mOSNs (isolated by FAC-sorting cells from zone 4-5 tissue in Ngn1-GFP and OMP-ires-GFP mice, respectively). Heatmaps showing coverage over OR gene bodies (scaled to 6kb, with 2kb flanking on either side) for iOSNs (right) and mOSNs (left) show H3K79me3 signal increasing on zone 4/5 ORs as the cells mature.
(E) H3K79me3 native ChIP signal track (in reads per 10 million) over the OR gene cluster containing Olfr1507 in zone1 mOSNs, zone 4/5 mOSNs, and Olfr1507-expressing zone5 mOSNs. Results show absence of H3K79me3 over the active gene in Olfr1507+ mOSNs and near equal signal over all other OR genes regardless of their zone.

Figure 3: Zonal OR compartmentalization permits OR genes from later zones to be recruited into the OR compartment.
(A) In situ Hi-C contact matrices of a 90Mb region of chr 2 from mOSNs isolated from zone1 (red), zone2 (yellow), and zone 4/5 (blue) show long range cis interactions between 3 large OR gene clusters. For zone2 OSNs instead of whole zonal dissection we used OSNs expressing the zone 2 OR Olfr17. The cluster on the right (blue) is enriched for zone 4/5 OR genes, does not contain any zone1 ORs, and only makes strong long range interactions with the other clusters in zone 4-5 mOSNs.
(B) Heatmaps of mean normalized *trans* HiC contacts between OR genes annotated by their zonal index at 50Kb resolution (excluding bins that contain ORs genes expressed in multiple zones and within 50kb of Greek Islands) shows increased trans contacts in cells from higher zones (zone1 mOSNs < zone2 mOSNs < zone 4/5 mOSNs).
(C) In aggregate, Greek Islands make a similar fraction of their total contacts with Greek islands on other chromosomes in zone1 and zone4/5 mOSNs.
(D) Analysis of contacts made by individual Greek Islands show that 6 out of 59 Greek Islands have differential frequency of trans interactions between the zone1 and zone5 mOSNs: 5 form significantly more interactions in zone 4/5 cells and one in zone1 (Wilcoxon rank sum test: padj < 0.05). Greek Islands residing in OR clusters with zone 4-5 ORs and no zone1 ORs (blue) appear relatively less engaged in zone1 mOSNs. Notably all the Greek Islands residing in the zone 4-5 rich OR cluster described in A form significantly fewer interactions in zone1 mOSNs. All analysis was performed at 50kb resolution and shows merged data from two biological replicates that yielded similar results when analyzed separately.

Sup 3: HiC in iOSNs shows nuclear structure during OR choice: isolation of ORs from lower zones and increased enhancer interaction
(A) Analogous analysis to Figure 3A in iOSNs. In situ Hi-C contact matrices of a 90Mb region of chr 2 in iOSNs (isolated by expression of Ngn1-GFP transgenic reporter) from zone1(red) and zone4-5 (blue) dissected OE, show long range cis interactions between three large OR gene clusters. The OR cluster on the right (blue) enriched for zone 4/5 ORs described in Figure 3 only makes strong interactions with other clusters in zone4 /5 iOSNs.
(B) Heatmaps of mean normalized trans HiC contacts between OR genes annotated by their zonal index described in Figure 3B show increased trans contacts in zone 4/5 iOSNs than zone1 iOSNs. In comparison to zone 4/5 mOSNs in Figure 3, iOSNs show weaker contacts made by zone 4/5 ORs.
(C) Interchromosomal interactions between OR gene loci are correlated with the amount of H3K79me3 deposition. The heatmap of normalized trans HiC contacts between the zonal OR cluster bins described in B (annotated in the zonal color bar) and ordered by H3K79me3 ChIP signal (gray color bar on top) shows that OR

cluster regions that have higher levels of H3K79me3 and are enriched for zone1 ORs have increased HiC contacts.
(D) Greek Islands form stronger interchromosomal interactions in zone4/5 iOSNs (blue) compared to zone1 iOSNs (red) (Wilcoxon rank sum test p-value = 8.12x10-8).
(E) Analogous to Figure 3D for iOSNs. Comparison of interchromosomal interactions made by individual Greek Islands show 47 out of 59 form significantly more contacts in zone 4-5 (Wilcoxon rank sum test padj < 0.5). Greek Islands residing in OR clusters with zone 4-5 ORs and no zone1 ORs (blue) are 5 of the top 6 differentially less-interacting Greek Islands. All analysis was performed at 50kb resolution and shows merged data from two biological replicates that yielded similar results when analyzed separately.

Figure 4: Diploid chromatin conformation capture (Dip-C) reveals differences in OR compartments in single zonal mOSNs.
(A,B) 3D genome structures in single cells generated from Diploid chromatin conformation capture (Dip-C) in 96 mOSNs--48 cells from each zone1 and zone4/5. Cells were isolated from OMP-ires-GFP mice crossed to Castaneous.
(A) Representative 3D structures of OR compartments in a zone1 and a zone4/5 mOSN colored by the chromosome, shows increased complexity of OR compartments in the zone 4/5 mOSN. Serial sections at 15 particle radii (approximately ~900nm) show several OR aggregates in each cell.
(B) The same 3D structures in A. colored by zonal annotation of each particle (analogous to 20kb of chromatin) shows OR genes of all zonal indexes intermingling.
(C,D) Single cell heatmaps of pairwise distances between OR genes generated from 3D nuclear structure in a zone1 mOSN (C) and zone4/5 mOSN (D) show an intermingling of OR genes from different chromosomes. Heatmaps sorted by chromosome order (right) or via hierarchical clustering (left) show increased interchromosomal contacts in the zone4/5 mOSN. The heatmap displaying contacts at any distance (top) and contacts within 10 particle radii (approximately ~600nm) shows larger aggregates in the zone 4-5 mOSN. Heatmaps of contacts within 10 particle radii are shown for all cells in Supplementary Figure 4. Only OR genes for which the zonal annotation is known (and that do not reside within 20kb of an OR of another zonal index or within 50kb of a Greek Island) were used to generate the heatmaps.
(E) Analysis of contact densities of interchromosomal contacts (the average contact density computed from single cell contact matrices) of all ORs (black), zone1 ORs (red), zone 2/3 ORs (green) and zone4/5 ORs (blue) with other ORs confirms that zone4/5 mOSNs (blue cell) have an increased density of interchromosomal contacts between OR genes of all zonal indexes, relative to zone1 mOSNs (red cell).
(F) A similar comparison of Greek Island contact densities with all ORs (black), zone1 ORs (red), zone2/3ORs (green) and zone4/5 ORs (blue) shows that in zone4/5 cells Greek Islands form stronger interchromosomal interactions with zone4-5 OR genes relative to zone1 and zone2-3 OR genes (Wilcoxon rank sum test: p-values = 1.078x10-6 and 0.002, respectively).

Supplementary Figure 4: OR compartments are highly variable between single cells but show a consistent difference between cells of different zones.
(A,B) Hierarchically clustered heatmaps of distances between zonal ORs (described in Figure 4) computed from 3D genome structures show OR genes within 10 particle radii (~600nm) in each zone1 mOSN (A) and zone4/5 mOSN (B).
(C,D) Analysis of all OR genes within 2.5, 5 and 10 particle radii (analogous to ~150, 300 and 600nm) of one another shows that OR genes in zone 4-5 mOSNs are in proximity to a greater number of OR genes on different chromosomes (C) and form more complex compartments (a higher number of chromosomes within a given distance) (D).
(E) On average a larger fraction of Greek Islands within zone 4-5 mOSNs is found near (within 2.5 particle radii) of Greek Islands from other chromosomes (mean of 27% in zone 4-5 cells, and 22% of zone 1 cells, Wilcoxon rank sum test p-value 0.003).
(F) The size of the largest hub (defined as the maximum site of interchromosomal interactions between Greek Islands)--which ranges from 4-8 interchromosomal Greek Islands--is similar in zone 1 and zone 4-5 mOSNs.

Figure 5: NFI family transcription factors have a graded expression in cells from different zones.

(A,B,C,D) Heatmaps of differentially expressed transcription factors between zone1, zone 2/3 and zone 4/5 cells at different developmental stages. GBC, INP, iOSN and mOSN cells were isolated from Mash1-CreER; Ngn1-GFP; TdTomato mice as described in Figure 1A. Significantly differentially expressed transcription factors (from the GO database) between zone5 and zone1 cells were defined as having an padj<0.05 and at least a two fold change in expression. By these metrics there were 30 differentially expressed TFs in GBCs (A), 31 in INPs (B), 44 in iOSNs (C) and 18 in OSNs (D). The heatmaps showing levels of transcription of these genes in TMP are sorted based on expression in zone 4/5 cells. A color bar above each heatmap shows the log2 fold change in zone 4/5 cells relative to zone 1 cells (blue values represent genes higher in zone 4/5).
(E) Transcription levels are shown for each of the members of the NFI family: NFIA, NFIB, and NFIX (NFIC, which is not expressed in the MOE, is excluded). Plots of expression at the four developmental stages in zone 1 cells (red), zone 2/3 cells (green) and zone 4/5 cells (blue) show higher levels of transcription of all three factors in zone 4/5 and show different patterns of expression during the course of development (NFIA and NFIB decrease in mOSNs, while NFIX increases).
(F) Omni-ATAC peaks in zone1 and zone 4/5 iONS were called using HOMER. Differential accessibility analysis with Diffbind identified 14,169 differentially accessible peaks (with a p-value < 0.05) between the two zones. Of these peaks 8,635 were more accessible in zone1 iOSNs and 5,534 were more accessible in zone 4/5 iOSNs. Top sequence motifs identified with HOMER show the NFI motif as the most highly enriched motif in zone4/5 unique peaks (p-value=1x10-1449).

Supplementary Figure 5: Increased expression of NFI factors in higher zones may compensate for their relatively weaker OR promoters.
(A) Transcription levels of core OSN lineage transcription factors EBF1-4, Lhx2 and Ldb1 at different stages of OSN development in zone1 cells (red), zone 2/3 cells (green) and zone 4/5 cells blue, isolated as described in Figure 1A. Only Ldb1 was differentially upregulated in zone 4/5 in INP and iOSN cells (padj =0.007 and 2.08x10-7, respectively).
(B) NFIA immunofluorescence staining in zones 1 and 5 (location in the MOE depicted in the schematic). Strong nuclear staining is only observed in zone 5 neuronal cells outlined on the top and bottom dashed lines. Staining is also observed in non-neuronal sustentacular cells (above the dashed line) in both zones.
(C) Heatmaps of omni-ATAC peaks in zone 1 and zone 4/5 iOSNs (described in Figure 5F) show increased coverage over OR promoters in zone 4/5 iOSNs. Heatmaps are centered on the TSS with 2kb flanking on either side and are sorted in the same order. The plots above the heatmap showing the average signal per zonal category are colored by zonal annotation. The motifs shown were identified with HOMER to be differentially enriched on 259 zone 1 OR promoters relative to 187 zone 4/5 promoters and vice versa. Promoters are defined as 500bp upstream and downstream of the TSS.
(D) Density plot of the best EBF motif found in each OR promoter separated by zone.
(E) Heatmap of omni-ATAC signal on each of the 63 Greek Islands in iOSNs and mOSNs. Heatmaps are centered on Greek Islands and show a 1kb flanking region on either side. A plot of omni-ATAC signal on each Greek Island is plotted for zone 1 iOSNs vs zone 4/5 iOSNs. Greek Islands are colored based on the types of ORs in their resident OR clusters. Greek Islands in clusters that have a 4/5 OR but no zone 1 ORs are colored in blue, and Greek Islands in clusters with zone 1 ORs and no zone 5 ORs are shown in red. Significantly differentially accessible Greek Islands (p-value<0.05 determined with Diffbind) are indicated with a star.
(F) Heatmaps of NFIA Cut&Run signal on Greek Islands show enrichment of signal relative to a control antibody. Heatmaps are centered on Greek Islands and show a 1kb flanking region on either side.

Figure 6: Triple knockout of NFI ABX in progenitors transforms zone4/5 olfactory epithelium into zone2/3.
(A) Olfr1507 immunofluorescence staining in MOE sections (magenta) of adult Nfi A,B,X triple knockout mice and age matched wt mice. Nuclei are stained with Dapi (blue). (See supplementary Figure 6A for description of the genetics and experimental design for inducing Nfi A,B,X knockout. Images were taken in the same spot in the MOE, shown in the schematic.
(B) Olfr17 Immunofluorescence staining in MOE sections (green) of adult NFI ABX triple knockout mice and age matched wt mice shows normal Olfr17 expression continues in NFI ABX knockout mice in zone 2. Olfr17 also spreads into higher zones and is detected in zone 5. Images were taken in the same spot in the MOE, shown in the schematic. Nuclei are stained with Dapi (blue).
(C) RNA-seq in NFI ABX knockout cells sorted from the whole MOE (described in Supplementary Figure 6A). Boxplot of expression of ORs of different zonal indexes shows increased expression of zone 2 and 3 ORs and a decreased expression of zone 4 and 5 ORs.

(D) RNA-seq was performed in mOSNs from zone 4/5 dissected MOE from NFI ABX triple knockout (NFI Af/f; Bf/f,Xf/f; Krt5-CreER;tdTomato; Omp-gfp), NFI AB double knockout (NFI Af/f; Bf/f; Krt5-CreER;tdTomato; Omp-gfp), NFIX knockout (NFI Xf/f; Krt5-CreER;tdTomato; Omp-gfp) and age matched (Krt5-CreER;tdTomato; Omp-gfp) according to the induction strategy described in Supplementary Figure S6A). Plots show a different pattern of transcription of OR genes of different zonal indexes in the different genotypes: zone 4/5 mOSNs in NFIABX triple knockout express zone 2/3 ORs, and in NFIAB double knockout express zone 3/4 ORs. Quantification of the number of differentially expressed ORs in each genotype is shown in Supplementary Figure S6D).

Supplementary Figure 6: OE in Nfi ABX knockout zone5 undergoes a homeotic transition in their OR repertoire while maintaining their cellular identity.
(A) Nfi ABX triple knockout cells were isolated by injecting adult NFI Af/f; Bf/f,Xf/f; Krt5-CreER; tdTomato mice with tamoxifen 3x, to induce knockout of Nfi A, B and X and to label knockout cells with the cre inducible reporter in quiescent horizontal basal cells (HBCs). After 10 days MOE was ablated with methimazole and rebuilt for 40 days followed by zonal dissection and FAC-sorting.
(B) Differential expression analysis of mature OSN markers, zone 4/5 mOSN markers and zone 1 mOSN markers in NFIABX knockout zone 4/5. We saw only 13/200 mature markers significantly downregulated in NFI ABX knockout, compared to 90/207 zone 5 markers significantly downregulated and 29/138 zone1 markers that were significantly upregulated.
(C) Transcriptional change in zone 4/5 mOSN markers varied between the NFI knockout genotypes described in Figure 6D.
(D) Differential expression analysis of all ORs in the different NFI knockout genotypes. Percentages of significantly upregulated (red) and downregulated (blue) ORs are shown.

Figure 7: Nfi A, B and X regulate chromatin state and OR compartment formation
(A,B) Native ChIP-seq for H3K79me3 (A) and H3K9me3 (B) in NFIABX knockout zone 5 cells (isolated as described in Supplementary Figure S7A). Heatmaps of ChIP signal over OR gene bodies (scaled to 6kb, with 2kb flanking on either side) show a dramatic decrease of both histone marks on zone 3-5 ORs compared to wt.
(C) Comparison of H3K79me3 on zone 2 and 3 OR genes that are ectopically expressed in zone 4/5 in NFIABX triple knockout. All zone 2 and zone 3 ORs were classified based on their expression difference in NFIABX knockout and wt as either not significant (gray), significant (light yellow for zone 2 and light green for zone 3) with a padj<0.05, or significant & most enriched (in dark yellow for zone 2 and dark green for zone 3) with padj<0.01 and a fold increase > 18 in the knockout.
(D) H3K79me3 signal density over the gene bodies of OR gene categories described in C (ChIP coverage scaled to the size of the OR) shows less deposition on zone 2 significant & very enriched genes compared to other non-significant and significant zone 2 ORs (Wilcoxon rank sum test p-values = 4.823e-07 and 0.001 respectively), and more deposition on zone 3 significant and very enriched genes compared to not significant and significant zone 3 ORs (Wilcoxon rank sum test p-values =1.649e-05 and 0.005 respectively) suggesting there is an optimum level of H3K79me3 that is consistent with high transcription.
(E) In situ Hi-C contact matrices of a 90Mb region of chr 2 from wt (left) and NFI ABX triple knockout zone 4/5 mOSNs (right) show long range cis interactions between 3 large OR gene clusters. The cluster on the right (blue)--enriched for zone 4-5 OR genes--shows almost no long range contacts with the other two clusters in NFI ABX knockout HiC.
(F) Heatmaps of mean normalised trans HiC contacts between OR genes annotated by their zonal index (as described in Figure 3B) show that in NFIABX zone 4/5 mOSNs (bottom) zone 4/5 ORs make proportionately fewer interchromosomal contacts compared to zone 1 and zone 4/5 ORs. In contrast wt zone 4/5 mOSNs have a more homogenous pattern of interchromosomal contacts between ORs.
(G) Comparison of interchromosomal interactions made by individual Greek Islands shows that 3 Greek Islands made significantly fewer contacts in zone 4/5 Nfi knockout mOSNs compared to wt (Wilcoxon rank sum test padj < 0.5) and one Greek Island (residing in an OR cluster with one zone 3 OR showed more contacts).

Supplementary Figure 7: Nfi ABX knockout zone 4/5 cells closely resemble zone 2/3 cells in chromatin state, compartment formation and OR expression
(A) Heatmap showing side by side comparison of H3K79me3 ChIP-seq signal in mOSNs from zone 1, zone 2/3, zone 4/5 and Nfi knockout zone 4/5 (previously described in Figure 2E and Figure 7A) shows NFI ABX knockout mOSNs have a chromatin state most similar to zone 2/3 mOSNs.

(B) Side by side heatmaps of trans HiC contacts between OR genes of different zonal indexes (previously described in Figures 3B and 7E) show the nuclear OR interactome in zone 4/5 NFI ABX knockout mOSNs is most similar to that of zone 2/3 mOSN Olfr17.
(C) Density plot of trans HiC contacts made by zone1 (red), zone 2-3 (green) and zone 4-5 (blue) ORs with a large OR cluster on chr 2 shows a different pattern of interactions. Color bar on the bottom shows the zonal annotation of OR genes residing in the cluster. Zone 5 mOSNs have a more level pattern of contacts, while both zone 2/3 mOSNs and zone 4/5 mOSNs show preferential interactions with regions that are enriched for zone 1 ORs.
(D) In aggregate, there is no significant difference between Greek Island contacts in zone 4/5 wt and NFI ABX triple knockout mOSNs.
(E) NFIA Cut&Run in cells isolated from all zones of the MOE in NFI ABX knockout and wt mice.

Bibliography

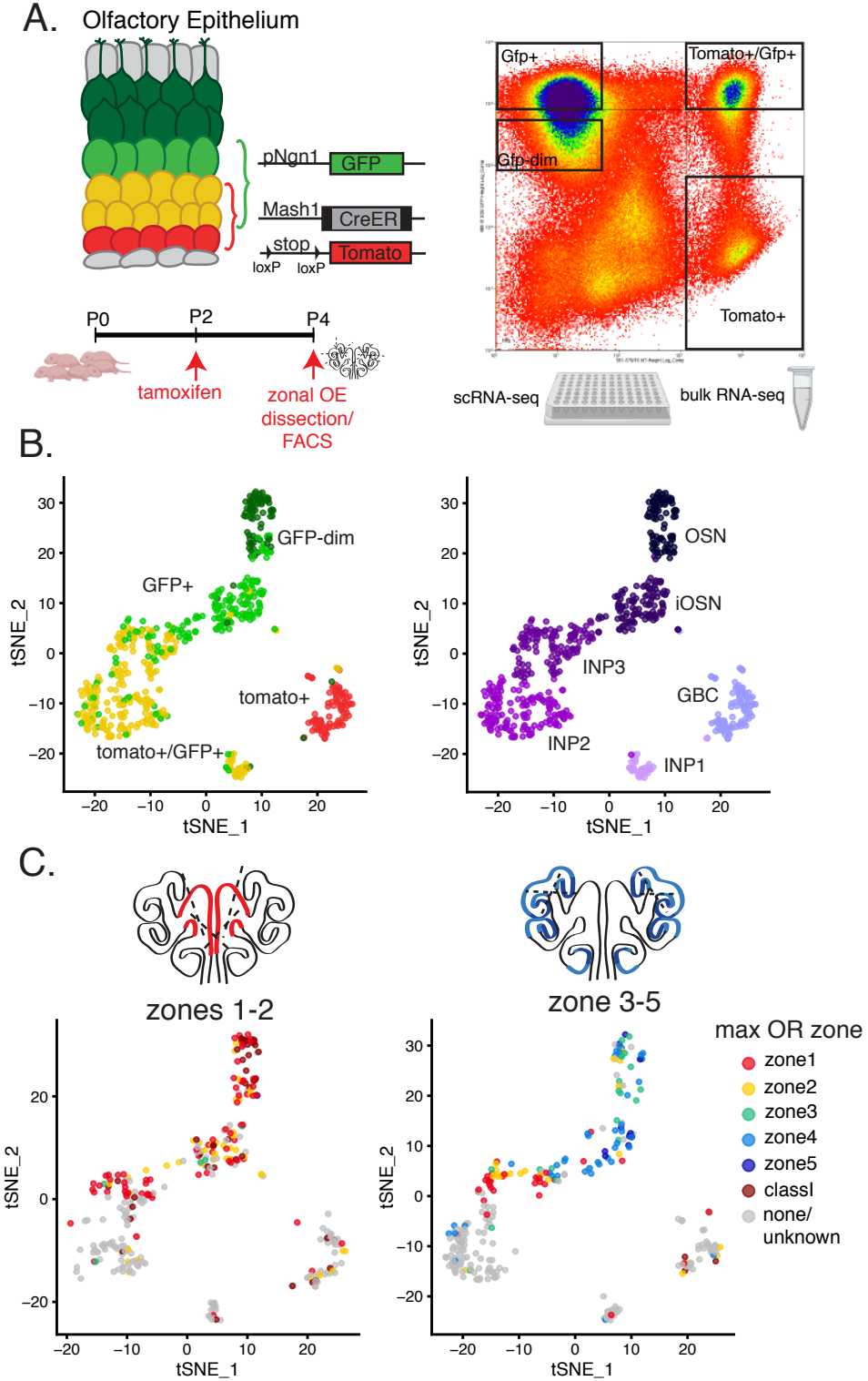
Allshire, R.C., and Madhani, H.D. (2018). Ten principles of heterochromatin formation and function. Nat Rev Mol Cell Biol *19*, 229-244.
Armelin-Correa, L.M., Gutiyama, L.M., Brandt, D.Y., and Malnic, B. (2014). Nuclear compartmentalization of odorant receptor genes. Proceedings of the National Academy of Sciences of the United States of America *111*, 2782-2787.
Barnea, G., O'Donnell, S., Mancia, F., Sun, X., Nemes, A., Mendelsohn, M., and Axel, R. (2004). Odorant receptors on axon termini in the brain. Science *304*, 1468.
Baumeister, H., Gronostajski, R.M., Lyons, G.E., and Margolis, F.L. (1999). Identification of NFI-binding sites and cloning of NFI-cDNAs suggest a regulatory role for NFI transcription factors in olfactory neuron gene expression. Brain Res Mol Brain Res *72*, 65-79.
Behrens, M., Venkatraman, G., Gronostajski, R.M., Reed, R.R., and Margolis, F.L. (2000). NFI in the development of the olfactory neuroepithelium and the regulation of olfactory marker protein gene expression. The European journal of neuroscience *12*, 1372-1384.
Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczynski, B., Riddell, A., and Furlong, E.E. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. Nature genetics *44*, 148-156.
Bozza, T., Vassalli, A., Fuss, S., Zhang, J.J., Weiland, B., Pacifico, R., Feinstein, P., and Mombaerts, P. (2009). Mapping of class I and class II odorant receptors to glomerular domains by two distinct types of olfactory sensory neurons in the mouse. Neuron *61*, 220-233.
Buck, L., and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell *65*, 175-187.
Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M.D., Madhani, H.D., Narlikar, G.J., and Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. Molecular cell *41*, 67-81.
Canzio, D., Liao, M., Naber, N., Pate, E., Larson, A., Wu, S., Marina, D.B., Garcia, J.F., Madhani, H.D., Cooke, R.*, et al.* (2013). A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. Nature.
Chess, A., Simon, I., Cedar, H., and Axel, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. Cell *78*, 823-834.
Clark, B.S., Stein-O'Brien, G.L., Shiau, F., Cannon, G.H., Davis-Marcisak, E., Sherman, T., Santiago, C.P., Hoang, T.V., Rajaii, F., James-Esposito, R.E.*, et al.* (2019). Single-Cell RNA-Seq Analysis of Retinal Development Identifies NFI Factors as Regulating Mitotic Exit and Late-Born Cell Specification. Neuron *102*, 1111-1126 e1115.
Clowney, E.J., Legros, M.A., Mosley, C.P., Clowney, F.G., Markenskoff-Papadimitriou, E.C., Myllys, M., Barnea, G., Larabell, C.A., and Lomvardas, S. (2012). Nuclear aggregation of olfactory receptor genes governs their monogenic expression. Cell *151*, 724-737.
Clowney, E.J., Magklara, A., Colquitt, B.M., Pathak, N., Lane, R.P., and Lomvardas, S. (2011). High-throughput mapping of the promoters of the mouse olfactory receptor genes reveals a new type of mammalian promoter and provides insight into olfactory receptor gene regulation. Genome research *21*, 1249-1259.
Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B.*, et al.* (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods *14*, 959-962.
D'Hulst, C., Mina, R.B., Gershon, Z., Jamet, S., Cerullo, A., Tomoiaga, D., Bai, L., Belluscio, L., Rogers, M.E., Sirotin, Y.*, et al.* (2016). MouSensor: A Versatile Genetic Platform to Create Super Sniffer Mice for Studying Human Odor Coding. Cell reports *16*, 1115-1125.
Dalton, R.P., Lyons, D.B., and Lomvardas, S. (2013). Co-opting the unfolded protein response to elicit olfactory receptor feedback. Cell *155*, 321-332.

Dasen, J.S., Tice, B.C., Brenner-Morton, S., and Jessell, T.M. (2005). A Hox regulatory network establishes motor neuron pool identity and target-muscle connectivity. Cell *123*, 477-491.
Ebrahimi, F.A., Edmondson, J., Rothstein, R., and Chess, A. (2000). YAC transgene-mediated olfactory receptor gene choice. Dev Dyn *217*, 225-231.
Enomoto, T., Nishida, H., Iwata, T., Fujita, A., Nakayama, K., Kashiwagi, T., Hatanaka, Y., Kondo, H., Kajitani, R., Itoh, T.*, et al.* (2019). Bcl11b controls odorant receptor class choice in mice. Commun Biol *2*, 296.
Falk, M., Feodorova, Y., Naumova, N., Imakaev, M., Lajoie, B.R., Leonhardt, H., Joffe, B., Dekker, J., Fudenberg, G., Solovei, I.*, et al.* (2019). Heterochromatin drives compartmentalization of inverted and conventional nuclei. Nature *570*, 395-399.
Farooq, Z., Banday, S., Pandita, T.K., and Altaf, M. (2016). The many faces of histone H3K79 methylation. Mutat Res Rev Mutat Res *768*, 46-52.
Fleischmann, A., Shykind, B.M., Sosulski, D.L., Franks, K.M., Glinka, M.E., Mei, D.F., Sun, Y., Kirkland, J., Mendelsohn, M., Albers, M.W.*, et al.* (2008). Mice with a "monoclonal nose": perturbations in an olfactory map impair odor discrimination. Neuron *60*, 1068-1081.
Fletcher, R.B., Das, D., Gadye, L., Street, K.N., Baudhuin, A., Wagner, A., Cole, M.B., Flores, Q., Choi, Y.G., Yosef, N.*, et al.* (2017). Deconstructing Olfactory Stem Cell Trajectories at Single-Cell Resolution. Cell Stem Cell *20*, 817-830 e818.
Fuss, S.H., Omura, M., and Mombaerts, P. (2007). Local and cis effects of the H element on expression of odorant receptor genes in mouse. Cell *130*, 373-384.
Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell *179*, 470-484 e421.
Godfrey, L., Crump, N.T., O'Byrne, S., Lau, I.J., Rice, S., Harman, J.R., Jackson, T., Elliott, N., Buck, G., Connor, C.*, et al.* (2020). H3K79me2/3 controls enhancer-promoter interactions and activation of the pan-cancer stem cell marker PROM1/CD133 in MLL-AF4 leukemia cells. Leukemia.
Godfrey, L., Crump, N.T., Thorne, R., Lau, I.J., Repapi, E., Dimou, D., Smith, A.L., Harman, J.R., Telenius, J.M., Oudelaar, A.M.*, et al.* (2019). DOT1L inhibition reveals a distinct subset of enhancers dependent on H3K79 methylation. Nature communications *10*, 2803.
Greer, P.L., Bear, D.M., Lassance, J.M., Bloom, M.L., Tsukahara, T., Pashkovski, S.L., Masuda, F.K., Nowlan, A.C., Kirchner, R., Hoekstra, H.E.*, et al.* (2016). A Family of non-GPCR Chemosensors Defines an Alternative Logic for Mammalian Olfaction. Cell *165*, 1734-1748.
Gronostajski, R.M. (2000). Roles of the NFI/CTF gene family in transcription and development. Gene *249*, 31-45.
Hanchate, N.K., Kondoh, K., Lu, Z., Kuang, D., Ye, X., Qiu, X., Pachter, L., Trapnell, C., and Buck, L.B. (2015). Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. Science *350*, 1251-1255.
Hirota, J., Omura, M., and Mombaerts, P. (2007). Differential impact of Lhx2 deficiency on expression of class I and class II odorant receptor genes in mouse. Molecular and cellular neurosciences *34*, 679-688.
Johnson, M.A., Tsai, L., Roy, D.S., Valenzuela, D.H., Mosley, C., Magklara, A., Lomvardas, S., Liberles, S.D., and Barnea, G. (2012). Neurons expressing trace amine-associated receptors project to discrete glomeruli and constitute an olfactory subsystem. Proceedings of the National Academy of Sciences of the United States of America *109*, 13410-13415.
Johnston, R.J., Jr., Otake, Y., Sood, P., Vogt, N., Behnia, R., Vasiliauskas, D., McDonald, E., Xie, B., Koenig, S., Wolf, R.*, et al.* (2011). Interlocked feedforward loops control cell-type-specific Rhodopsin expression in the Drosophila eye. Cell *145*, 956-968.
Jones, B., Su, H., Bhat, A., Lei, H., Bajko, J., Hevi, S., Baltus, G.A., Kadam, S., Zhai, H., Valdez, R.*, et al.* (2008). The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. PLoS genetics *4*, e1000190.
Kim, S., Dunham, M.J., and Shendure, J. (2019). A combination of transcription factors mediates inducible interchromosomal contacts. eLife *8*.
Kim, S., and Shendure, J. (2019). Mechanisms of Interplay between Transcription Factors and the 3D Genome. Molecular cell *76*, 306-319.
Kim, S.K., Jung, I., Lee, H., Kang, K., Kim, M., Jeong, K., Kwon, C.S., Han, Y.M., Kim, Y.S., Kim, D.*, et al.* (2012). Human histone H3K79 methyltransferase DOT1L protein [corrected] binds actively transcribing RNA polymerase II to regulate gene expression. The Journal of biological chemistry *287*, 39698-39709.

Lewcock, J.W., and Reed, R.R. (2004). A feedback mechanism regulates monoallelic odorant receptor expression. Proceedings of the National Academy of Sciences of the United States of America *101*, 1069-1074.
Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565-570.
Lyons, D.B., Allen, W.E., Goh, T., Tsai, L., Barnea, G., and Lomvardas, S. (2013). An epigenetic trap stabilizes singular olfactory receptor expression. Cell *154*, 325-336.
Magklara, A., Yen, A., Colquitt, B.M., Clowney, E.J., Allen, W., Markenscoff-Papadimitriou, E., Evans, Z.A., Kheradpour, P., Mountoufaris, G., Carey, C.*, et al.* (2011). An epigenetic signature for monoallelic olfactory receptor expression. Cell *145*, 555-570.
Markenscoff-Papadimitriou, E., Allen, W.E., Colquitt, B.M., Goh, T., Murphy, K.K., Monahan, K., Mosley, C.P., Ahituv, N., and Lomvardas, S. (2014). Enhancer interaction networks as a means for singular olfactory receptor expression. Cell *159*, 543-557.
Meers, M.P., Bryson, T.D., Henikoff, J.G., and Henikoff, S. (2019a). Improved CUT&RUN chromatin profiling tools. eLife *8*.
Meers, M.P., Tenenbaum, D., and Henikoff, S. (2019b). Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. Epigenetics Chromatin *12*, 42.
Mirny, L.A., Imakaev, M., and Abdennur, N. (2019). Two major mechanisms of chromosome organization. Curr Opin Cell Biol *58*, 142-152.
Miyamichi, K., Serizawa, S., Kimura, H.M., and Sakano, H. (2005). Continuous and overlapping expression domains of odorant receptor genes in the olfactory epithelium determine the dorsal/ventral positioning of glomeruli in the olfactory bulb. J Neurosci *25*, 3586-3592.
Monahan, K., Horta, A., and Lomvardas, S. (2019). LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. Nature *565*, 448-453.
Monahan, K., Schieren, I., Cheung, J., Mumbey-Wafula, A., Monuki, E.S., and Lomvardas, S. (2017). Cooperative interactions enable singular olfactory receptor expression in mouse olfactory neurons. eLife *6*.
Narlikar, G.J. (2020). Phase-separation in chromatin organization. Journal of biosciences *45*.
Nguyen, M.Q., Zhou, Z., Marks, C.A., Ryba, N.J., and Belluscio, L. (2007). Prominent roles for odorant receptor coding sequences in allelic exclusion. Cell *131*, 1009-1017.
Nishizumi, H., Kumasaka, K., Inoue, N., Nakashima, A., and Sakano, H. (2007). Deletion of the core-H region in mice abolishes the expression of three proximal odorant receptor genes in cis. Proceedings of the National Academy of Sciences of the United States of America *104*, 20067-20072.
Noordermeer, D., and Duboule, D. (2013). Chromatin architectures and Hox gene collinearity. Current topics in developmental biology *104*, 113-148.
Nusslein-Volhard, C., and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. Nature *287*, 795-801.
Ressler, K.J., Sullivan, S.L., and Buck, L.B. (1993). A zonal organization of odorant receptor gene expression in the olfactory epithelium. Cell *73*, 597-609.
Ressler, K.J., Sullivan, S.L., and Buck, L.B. (1994). Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. Cell *79*, 1245-1255.
Rodriguez-Gil, D.J., Bartel, D.L., Jaspers, A.W., Mobley, A.S., Imamura, F., and Greer, C.A. (2015). Odorant receptors regulate the final glomerular coalescence of olfactory sensory neuron axons. Proceedings of the National Academy of Sciences of the United States of America *112*, 5821-5826.
Rothman, A., Feinstein, P., Hirota, J., and Mombaerts, P. (2005). The promoter of the mouse odorant receptor gene M71. Molecular and cellular neurosciences *28*, 535-546.
Saraiva, L.R., Ibarra-Soria, X., Khan, M., Omura, M., Scialdone, A., Mombaerts, P., Marioni, J.C., and Logan, D.W. (2015). Hierarchical deconstruction of mouse olfactory sensory neurons: from whole mucosa to single-cell RNA-seq. Scientific reports *5*, 18178.
Serizawa, S., Ishii, T., Nakatani, H., Tsuboi, A., Nagawa, F., Asano, M., Sudo, K., Sakagami, J., Sakano, H., Ijiri, T.*, et al.* (2000). Mutually exclusive expression of odorant receptor transgenes. Nature neuroscience *3*, 687-693.
Serizawa, S., Miyamichi, K., Nakatani, H., Suzuki, M., Saito, M., Yoshihara, Y., and Sakano, H. (2003). Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. Science *302*, 2088-2094.
Serizawa, S., Miyamichi, K., and Sakano, H. (2005). Negative feedback regulation ensures the one neuron-one receptor rule in the mouse olfactory system. Chemical senses *30 Suppl 1*, i99-100.

Shykind, B.M., Rohani, S.C., O'Donnell, S., Nemes, A., Mendelsohn, M., Sun, Y., Axel, R., and Barnea, G. (2004). Gene switching and the stability of odorant receptor gene choice. Cell *117*, 801-815.
Steger, D.J., Lefterova, M.I., Ying, L., Stonestrom, A.J., Schupp, M., Zhuo, D., Vakoc, A.L., Kim, J.E., Chen, J., Lazar, M.A.*, et al.* (2008). DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. Molecular and cellular biology *28*, 2825-2839.
Stettler, D.D., and Axel, R. (2009). Representations of odor in the piriform cortex. Neuron *63*, 854-864.
Sullivan, S.L., Adamson, M.C., Ressler, K.J., Kozak, C.A., and Buck, L.B. (1996). The chromosomal distribution of mouse odorant receptor genes. Proceedings of the National Academy of Sciences of the United States of America *93*, 884-888.
Sullivan, S.L., Bohm, S., Ressler, K.J., Horowitz, L.F., and Buck, L.B. (1995). Target-independent pattern specification in the olfactory epithelium. Neuron *15*, 779-789.
Sullivan, S.L., Ressler, K.J., and Buck, L.B. (1994). Odorant receptor diversity and patterned gene expression in the mammalian olfactory epithelium. Prog Clin Biol Res *390*, 75-84.
Tan, L., Li, Q., and Xie, X.S. (2015). Olfactory sensory neurons transiently express multiple olfactory receptors during development. Molecular systems biology *11*, 844.
Tan, L., and Xie, X.S. (2018). A Near-Complete Spatial Map of Olfactory Receptors in the Mouse Main Olfactory Epithelium. Chemical senses *43*, 427-432.
Tan, L., Xing, D., Daley, N., and Xie, X.S. (2019). Three-dimensional genome structures of single sensory neurons in mouse visual and olfactory systems. Nature structural & molecular biology *26*, 297-307.
Tian, H., and Ma, M. (2004). Molecular organization of the olfactory septal organ. J Neurosci *24*, 8383-8390.
Turing, A.M. (1953). The Chemical Basis of Morphogenesis. Philosophical Transactions of the Royal Society *237*, 37-73.
Vassalli, A., Rothman, A., Feinstein, P., Zapotocky, M., and Mombaerts, P. (2002). Minigenes impart odorant receptor-specific axon guidance in the olfactory bulb. Neuron *35*, 681-696.
Vassar, R., Ngai, J., and Axel, R. (1993). Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. Cell *74*, 309-318.
Zapiec, B., and Mombaerts, P. (2020). The Zonal Organization of Odorant Receptor Gene Choice in the Main Olfactory Epithelium of the Mouse. Cell reports *30*, 4220-4234 e4225.
Zenker, M., Bunt, J., Schanze, I., Schanze, D., Piper, M., Priolo, M., Gerkes, E.H., Gronostajski, R.M., Richards, L.J., Vogt, J.*, et al.* (2019). Variants in nuclear factor I genes influence growth and development. Am J Med Genet C Semin Med Genet *181*, 611-626.
Zhang, X., and Firestein, S. (2009). Genomics of olfactory receptors. Results Probl Cell Differ *47*, 25-36.

Figure1


A.
Olfactory Epithelium
pNgn1
GFP
Mash1
CreER
stop
Tomato
loxP
loxP
P0
P2
P4
tamoxifen
zonal OE dissection/ FACS
Gfp+
Tomato+/Gfp+
Gfp-dim
Tomato+
scRNA-seq
bulk RNA-seq
B.
GFP-dim
GFP+
tomato+
tomato+/GFP+
tSNE_1
tSNE_2
OSN
iOSN
INP3
GBC
INP2
INP1
C.
zones 1-2
zone 3-5
max OR zone
zone1
zone2
zone3
zone4
zone5
classI
none/ unknown

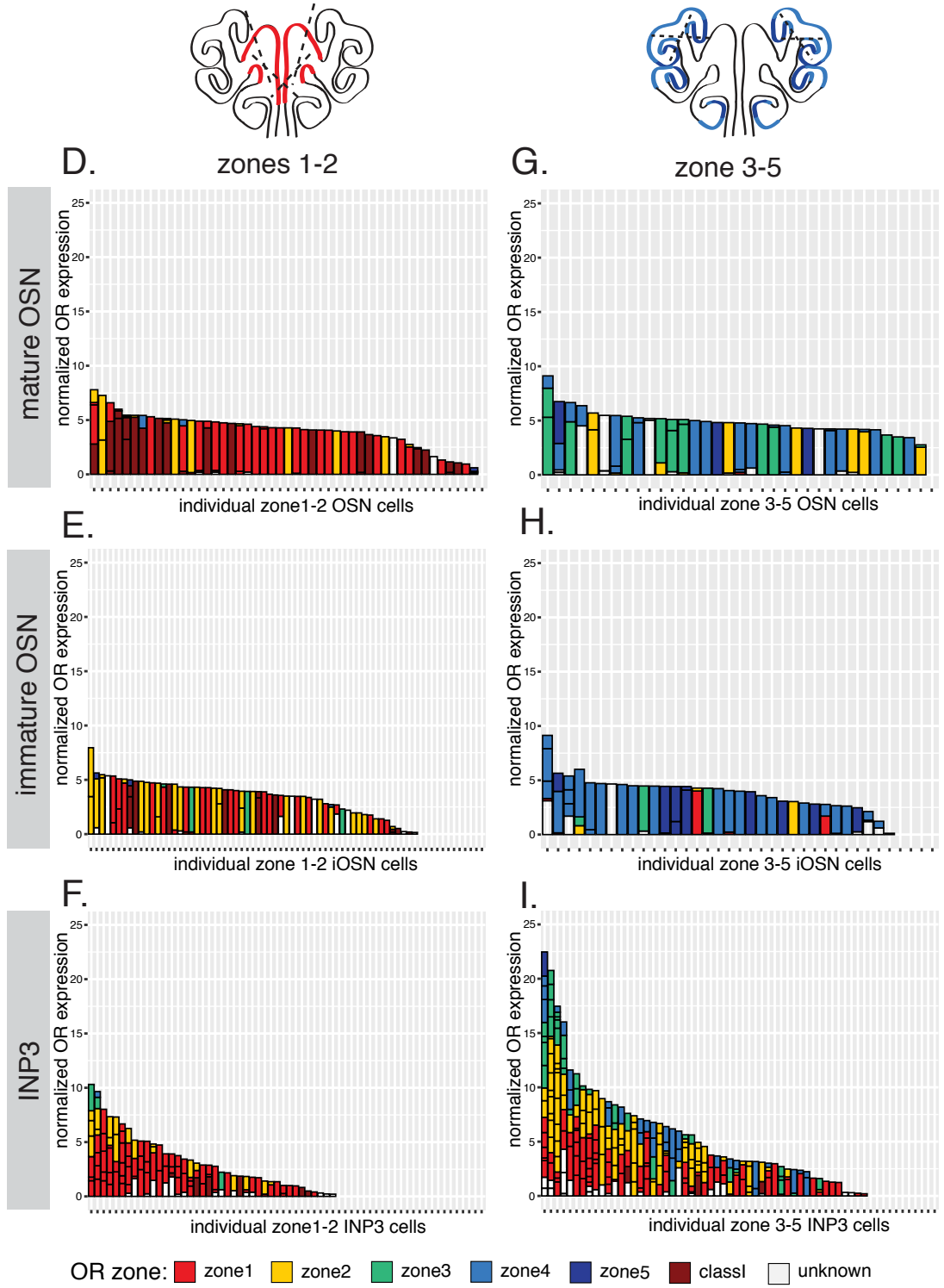

D.
zones 1-2
G.
zone 3-5
mature OSN
normalized OR expression
individual zone1-2 OSN cells
individual zone 3-5 OSN cells
E.
H.
immature OSN
individual zone 1-2 iOSN cells
individual zone 3-5 iOSN cells
F.
I.
INP3
individual zone1-2 INP3 cells
individual zone 3-5 INP3 cells
OR zone:
zone1
zone2
zone3
zone4
zone5
classI
unknown

Figure 2

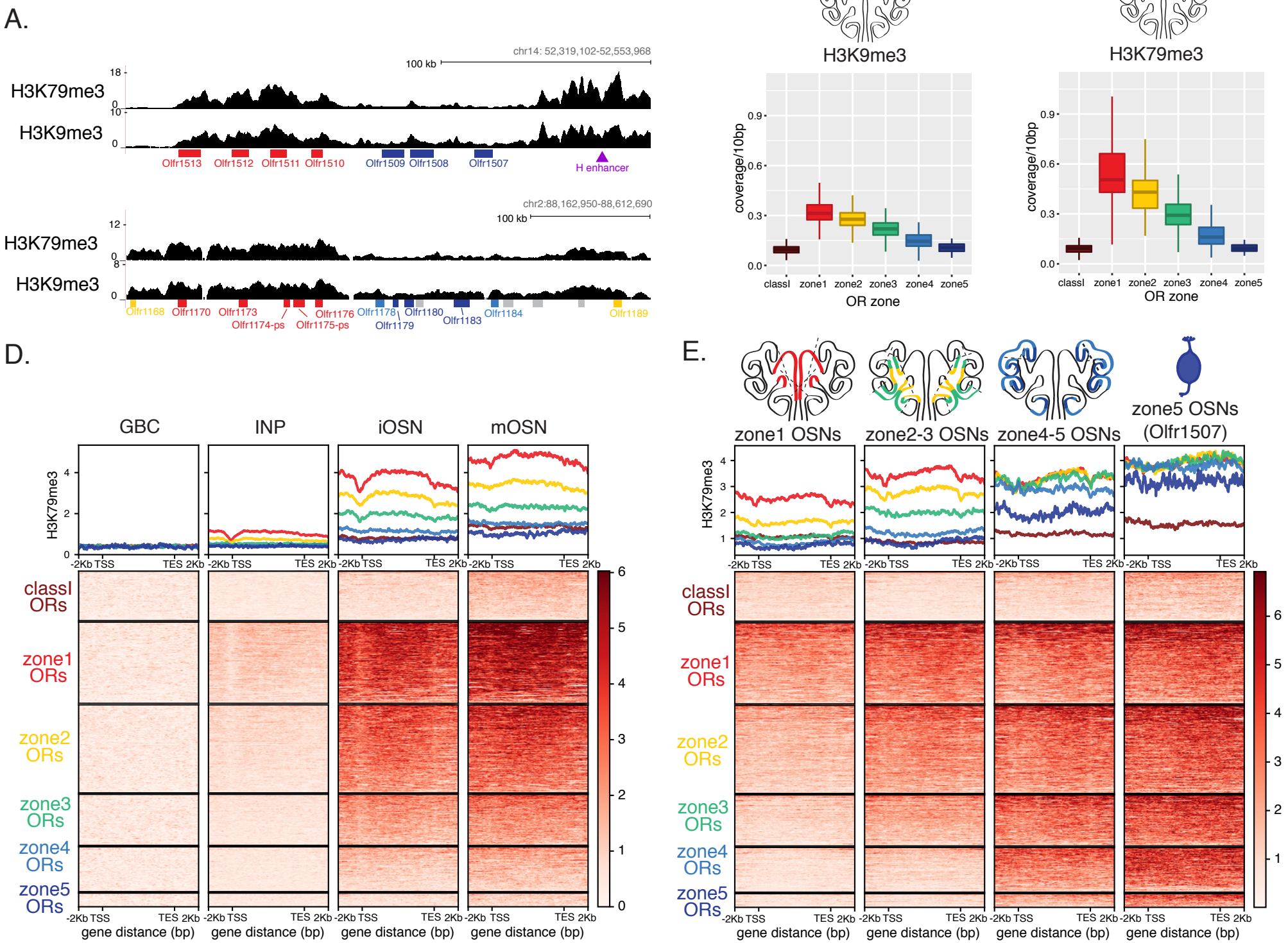

A.
chr14: 52,319,102-52,553,968
100 kb
H3K79me3
H3K9me3
Olfr1513
Olfr1512
Olfr1511
Olfr1510
Olfr1509
Olfr1508
Olfr1507
H enhancer
chr2:88,162,950-88,612,690
Olfr1168
Olfr1170
Olfr1173
Olfr1174-ps
Olfr1175-ps
Olfr1176
Olfr1178
Olfr1179
Olfr1180
Olfr1183
Olfr1184
Olfr1189
B.
H3K9me3
coverage/10bp
OR zone
classI
zone1
zone2
zone3
zone4
zone5
C.
H3K79me3
D.
GBC
INP
iOSN
mOSN
-2Kb TSS
TES 2Kb
gene distance (bp)
classI ORs
zone1 ORs
zone2 ORs
zone3 ORs
zone4 ORs
zone5 ORs
E.
zone1 OSNs
zone2-3 OSNs
zone4-5 OSNs
zone5 OSNs (Olfr1507)

Figure 3

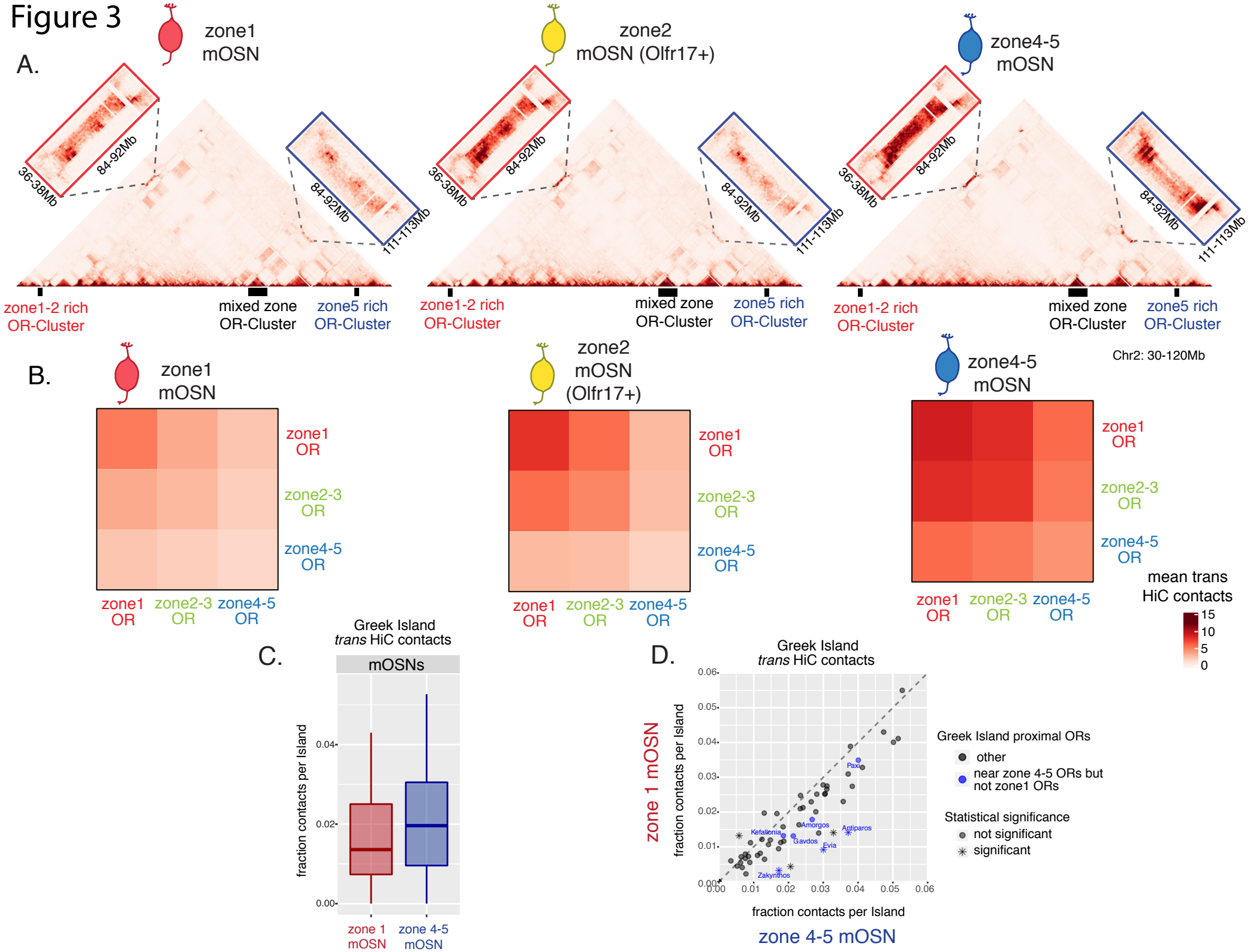

A.
zone1
mOSN
zone2
mOSN (Olfr17+)
zone4-5
mOSN
36-38Mb
84-92Mb
111-113Mb
zone1-2 rich
OR-Cluster
mixed zone
OR-Cluster
zone5 rich
OR-Cluster
B.
Chr2: 30-120Mb
zone1
OR
zone2-3
OR
zone4-5
OR
mean trans
HiC contacts
15
10
5
0
C.
Greek Island
trans HiC contacts
mOSNs
fraction contacts per Island
0.04
0.02
0.00
zone 1
mOSN
zone 4-5
mOSN
D.
Greek Island
trans HiC contacts
zone 1 mOSN
fraction contacts per Island
0.06
0.05
0.04
0.03
0.02
0.01
0.00
Paxi
Amorgos
Antiparos
Kefallonia
Gavdos
Evia
Zakynthos
fraction contacts per Island
zone 4-5 mOSN
Greek Island proximal ORs
other
near zone 4-5 ORs but
not zone1 ORs
Statistical significance
not significant
significant

Figure 4

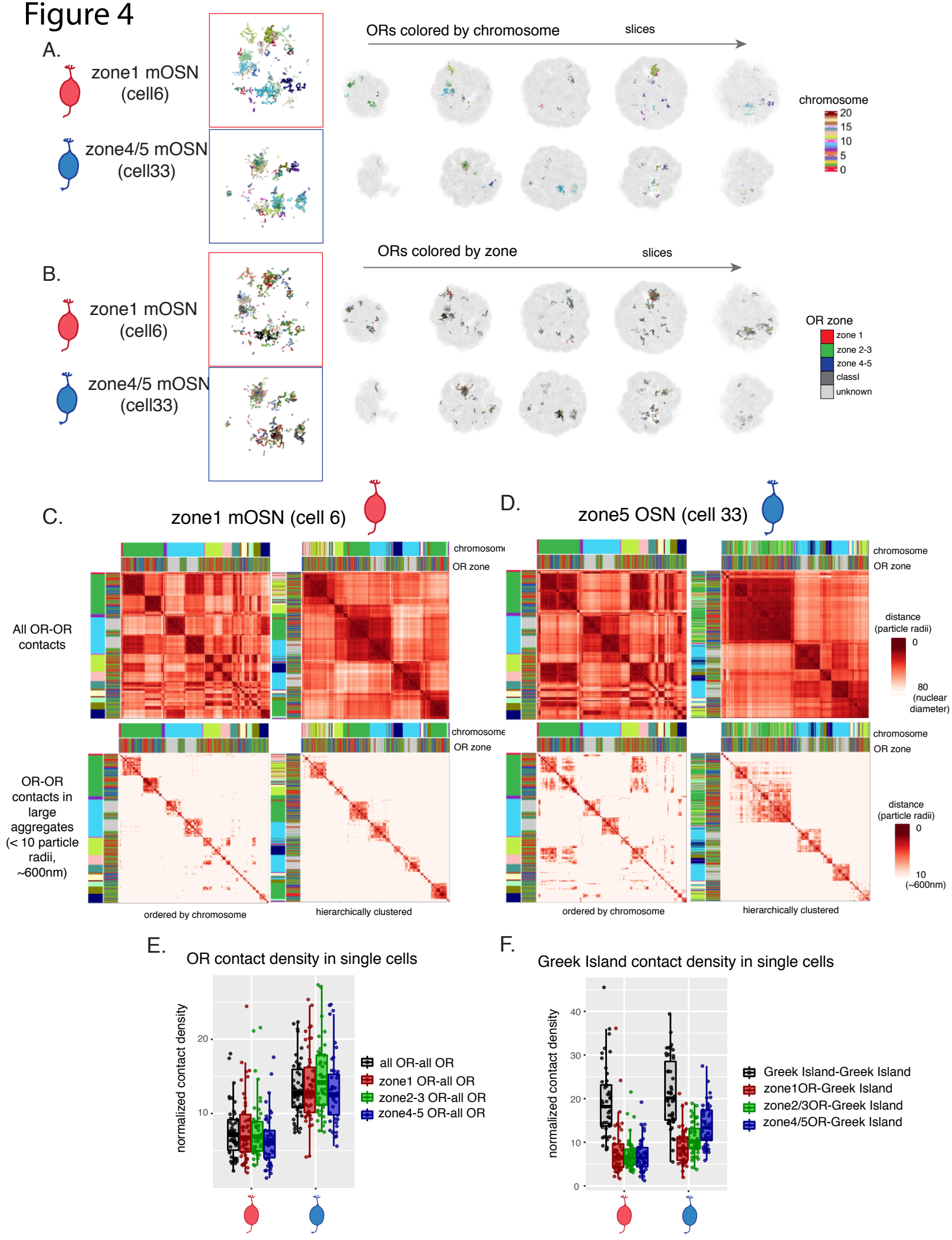

A.
zone1 mOSN (cell6)
zone4/5 mOSN (cell33)
ORs colored by chromosome
slices
chromosome
20
15
10
5
0
B.
zone1 mOSN (cell6)
zone4/5 mOSN (cell33)
ORs colored by zone
slices
OR zone
zone 1
zone 2-3
zone 4-5
classI
unknown
C.
zone1 mOSN (cell 6)
D.
zone5 OSN (cell 33)
All OR-OR contacts
OR-OR contacts in large aggregates (< 10 particle radii, ~600nm)
chromosome
OR zone
distance (particle radii)
0
80 (nuclear diameter)
10 (~600nm)
ordered by chromosome
hierarchically clustered
E.
OR contact density in single cells
normalized contact density
all OR-all OR
zone1 OR-all OR
zone2-3 OR-all OR
zone4-5 OR-all OR
F.
Greek Island contact density in single cells
normalized contact density
Greek Island-Greek Island
zone1OR-Greek Island
zone2/3OR-Greek Island
zone4/5OR-Greek Island

Figure 5

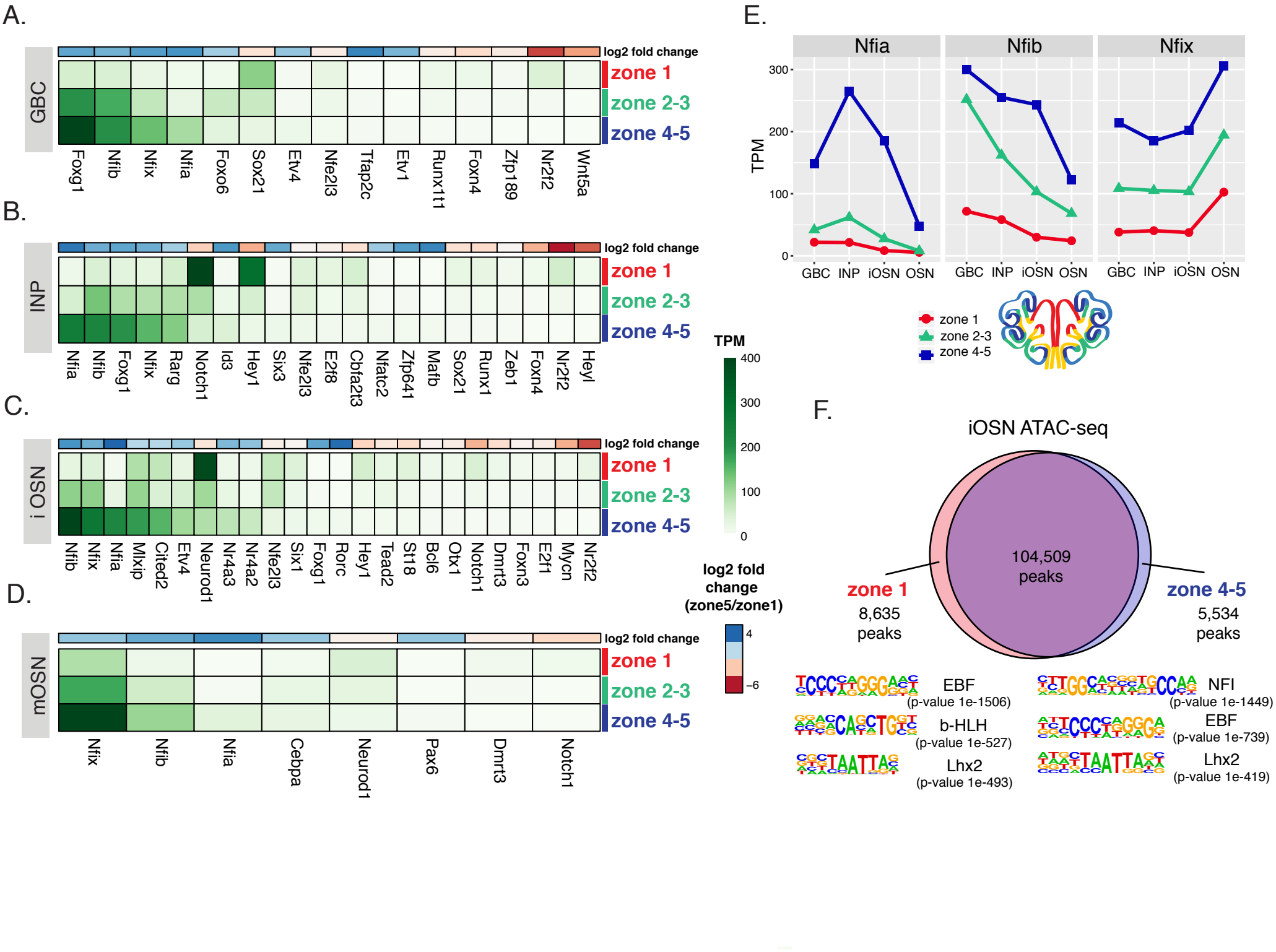

A.
GBC
log2 fold change
zone 1
zone 2-3
zone 4-5
Foxg1
Nfib
Nfix
Nfia
Foxo6
Sox21
Etv4
Nfe2l3
Tfap2c
Etv1
Runx1t1
Foxn4
Zfp189
Nr2f2
Wnt5a
B.
INP
Nfia
Nfib
Foxg1
Nfix
Rarg
Notch1
Id3
Hey1
Six3
Nfe2l3
E2f8
Cbfa2t3
Nfatc2
Zfp641
Matb
Sox21
Runx1
Zeb1
Foxn4
Nr2f2
Heyl
C.
i OSN
Nfib
Nfix
Nfia
Mlxip
Cited2
Etv4
Neurod1
Nr4a3
Nr4a2
Nfe2l3
Six1
Foxg1
Rorc
Hey1
Tead2
St18
Bcl6
Otx1
Notch1
Dmrt3
Foxn3
E2f1
Mycn
Nr2f2
D.
mOSN
Nfix
Nfib
Nfia
Cebpa
Neurod1
Pax6
Dmrt3
Notch1
TPM
400
300
200
100
0
log2 fold change (zone5/zone1)
4
−6
E.
Nfia
Nfib
Nfix
TPM
300
200
100
0
GBC
INP
iOSN
OSN
zone 1
zone 2-3
zone 4-5
F.
iOSN ATAC-seq
104,509 peaks
zone 1
8,635 peaks
zone 4-5
5,534 peaks
EBF
(p-value 1e-1506)
NFI
(p-value 1e-1449)
b-HLH
(p-value 1e-527)
EBF
(p-value 1e-739)
Lhx2
(p-value 1e-493)
Lhx2
(p-value 1e-419)

Figure 6

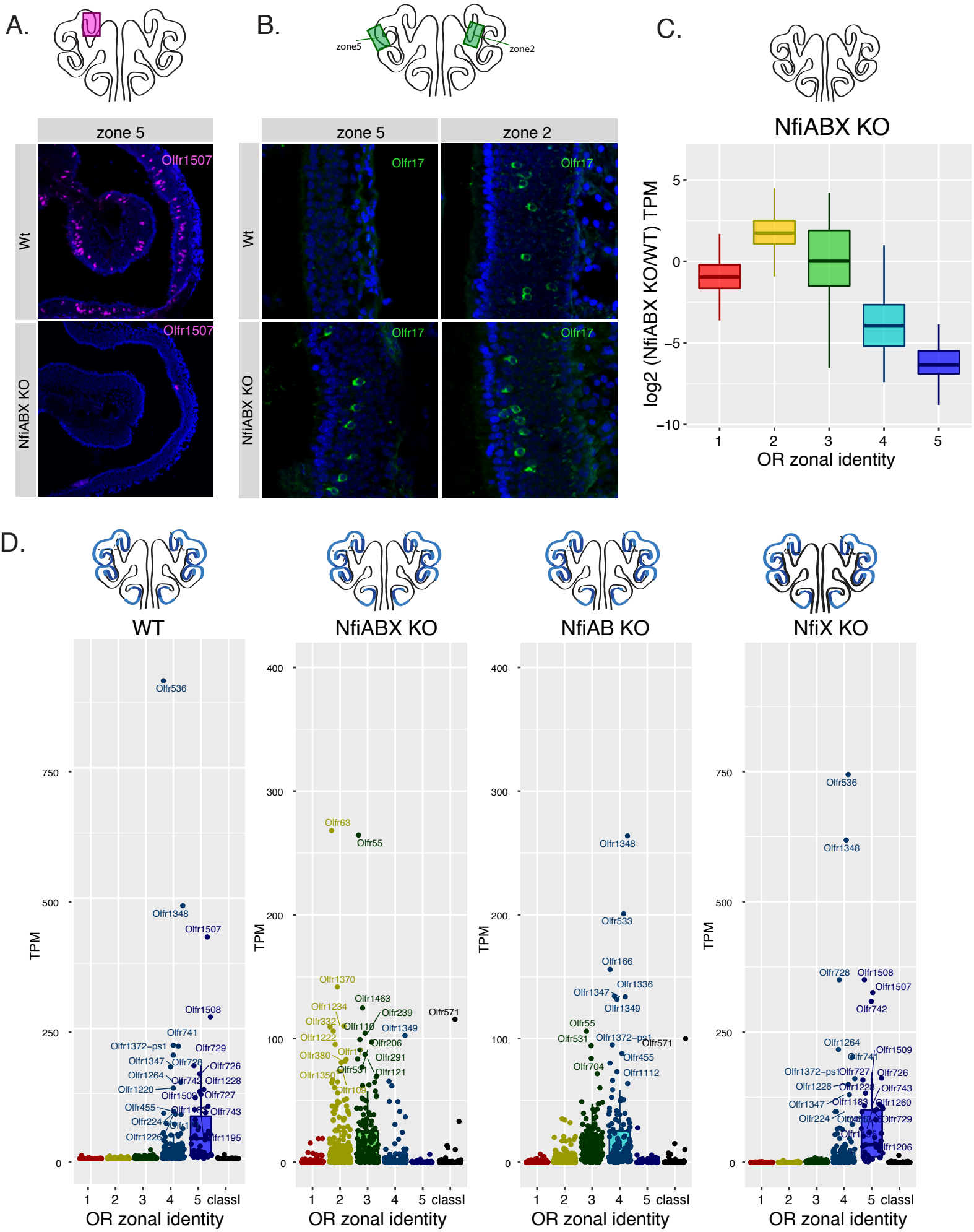
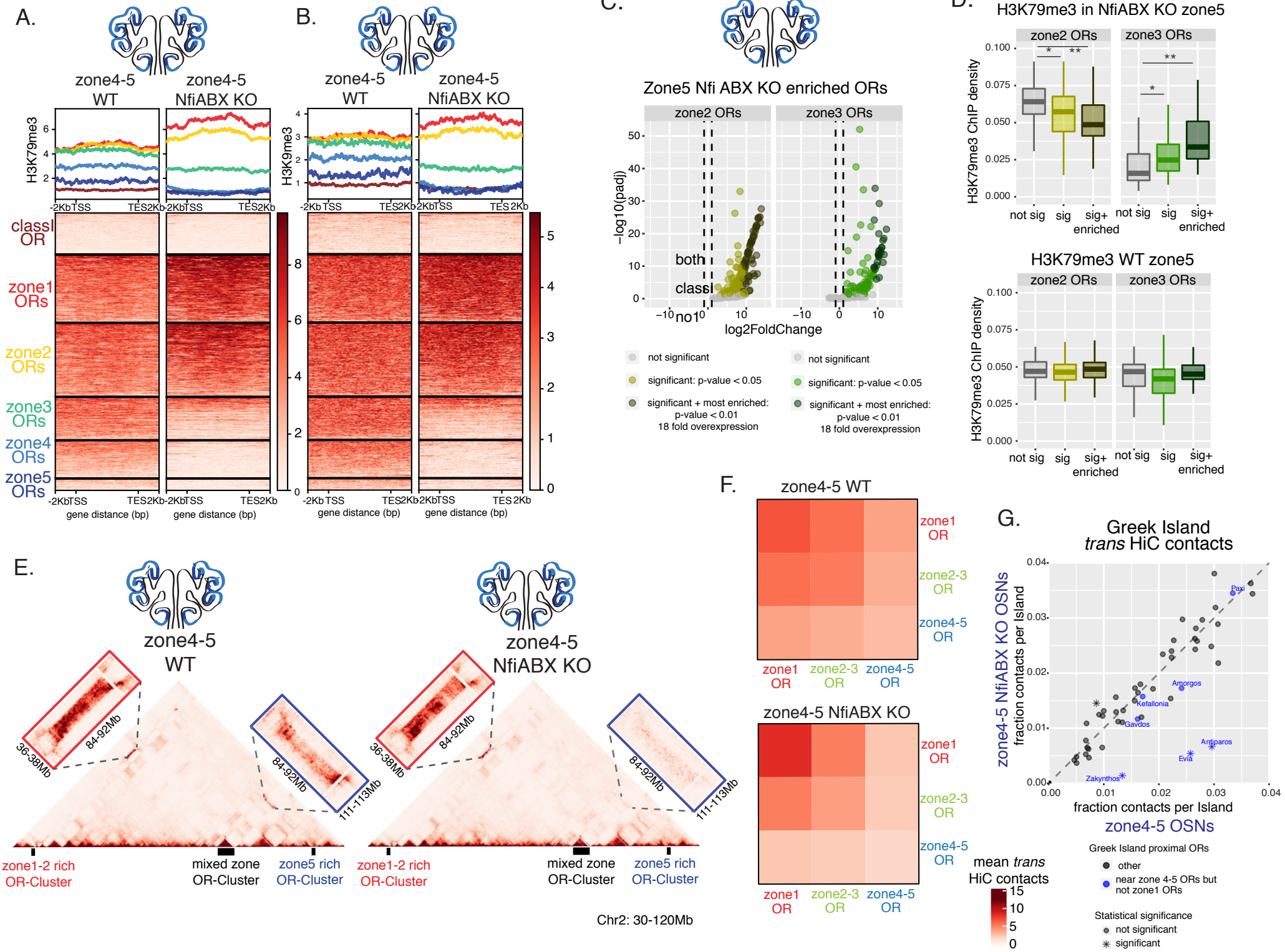

A.
zone 5
Wt
NfiABX KO
Olfr1507
B.
zone5
zone2
zone 5
zone 2
Olfr17
C.
NfiABX KO
log2 (NfiABX KO/WT) TPM
OR zonal identity
D.
WT
NfiABX KO
NfiAB KO
NfiX KO
TPM
OR zonal identity
classI

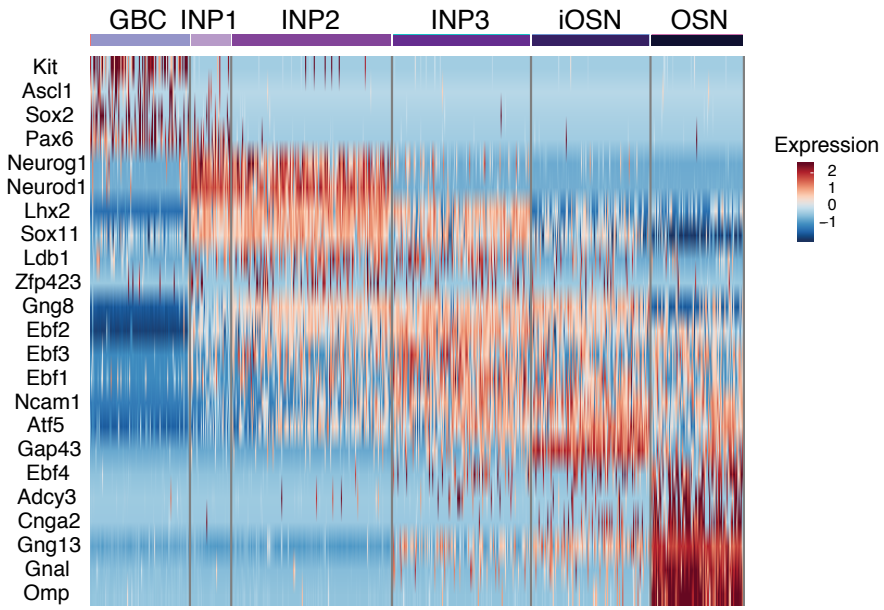
Figure 7


A.
zone4-5 WT
zone4-5 NfiABX KO
H3K79me3
classI OR
zone1 ORs
zone2 ORs
zone3 ORs
zone4 ORs
zone5 ORs
-2Kb TSS
TES 2Kb
gene distance (bp)
B.
H3K9me3
C.
Zone5 Nfi ABX KO enriched ORs
zone2 ORs
zone3 ORs
−log10(padj)
both
classI
no1
log2FoldChange
not significant
significant: p-value < 0.05
significant + most enriched: p-value < 0.01 18 fold overexpression
D.
H3K79me3 in NfiABX KO zone5
H3K79me3 ChIP density
not sig
sig
sig+ enriched
H3K79me3 WT zone5
E.
36-38Mb
84-92Mb
111-113Mb
zone1-2 rich OR-Cluster
mixed zone OR-Cluster
zone5 rich OR-Cluster
Chr2: 30-120Mb
F.
zone4-5 WT
zone4-5 NfiABX KO
zone1 OR
zone2-3 OR
zone4-5 OR
mean trans HiC contacts
G.
Greek Island trans HiC contacts
zone4-5 NfiABX KO OSNs fraction contacts per Island
zone4-5 OSNs fraction contacts per Island
Paxi
Amorgos
Kefallonia
Gavdos
Antiparos
Evia
Zakynthos
Greek Island proximal ORs
other
near zone 4-5 ORs but not zone1 ORs
Statistical significance
not significant
significant

Supplemental Figure S1

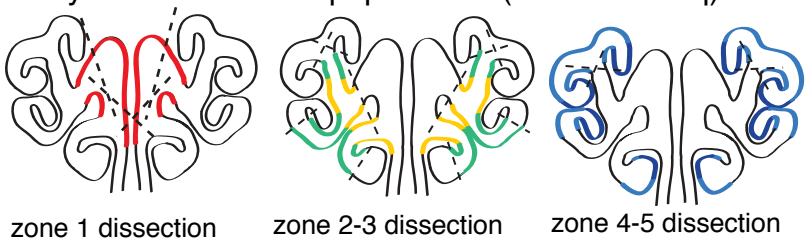
A.

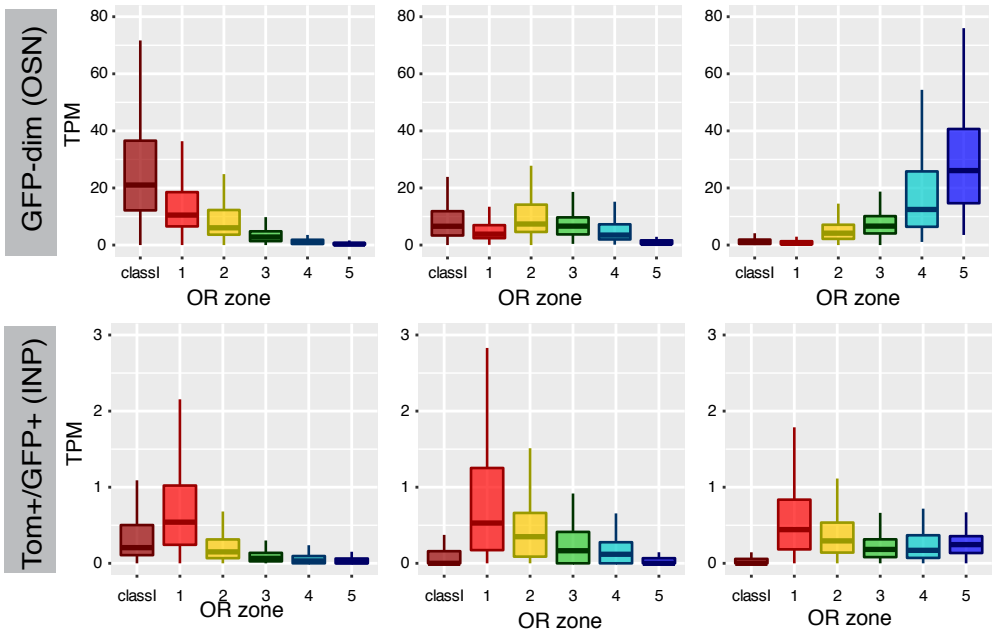
Olfactory marker expression (single cell RNA-seq populations)



B.

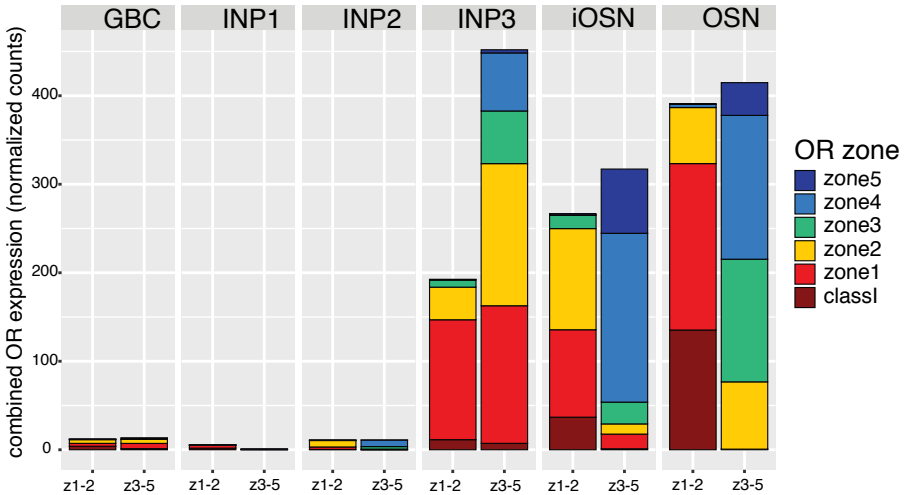
zonally dissected FACS populations (bulk RNA-seq)



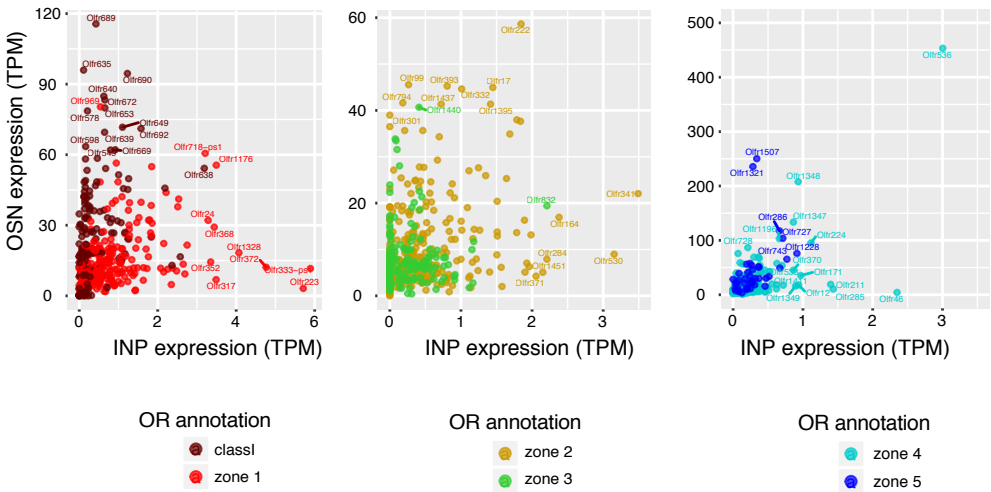


C.

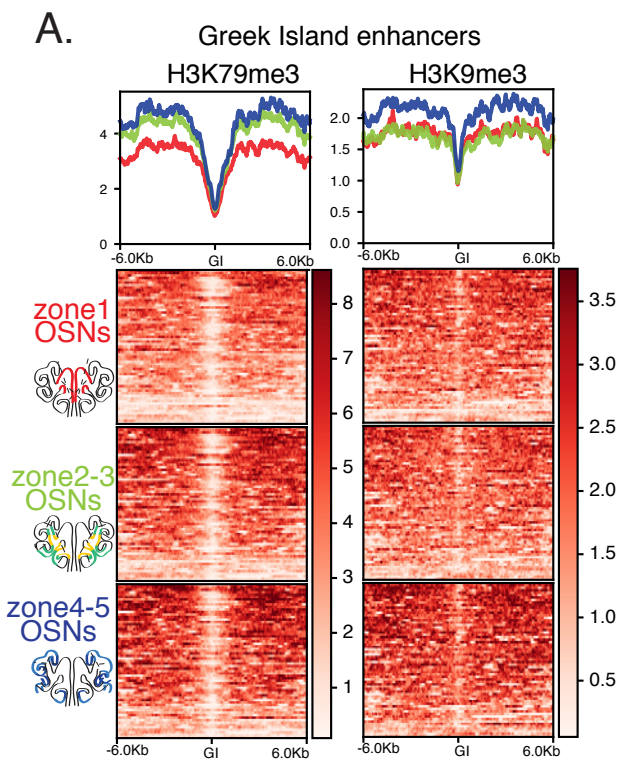
Olfactory receptor expression (single cell RNA-seq populations)

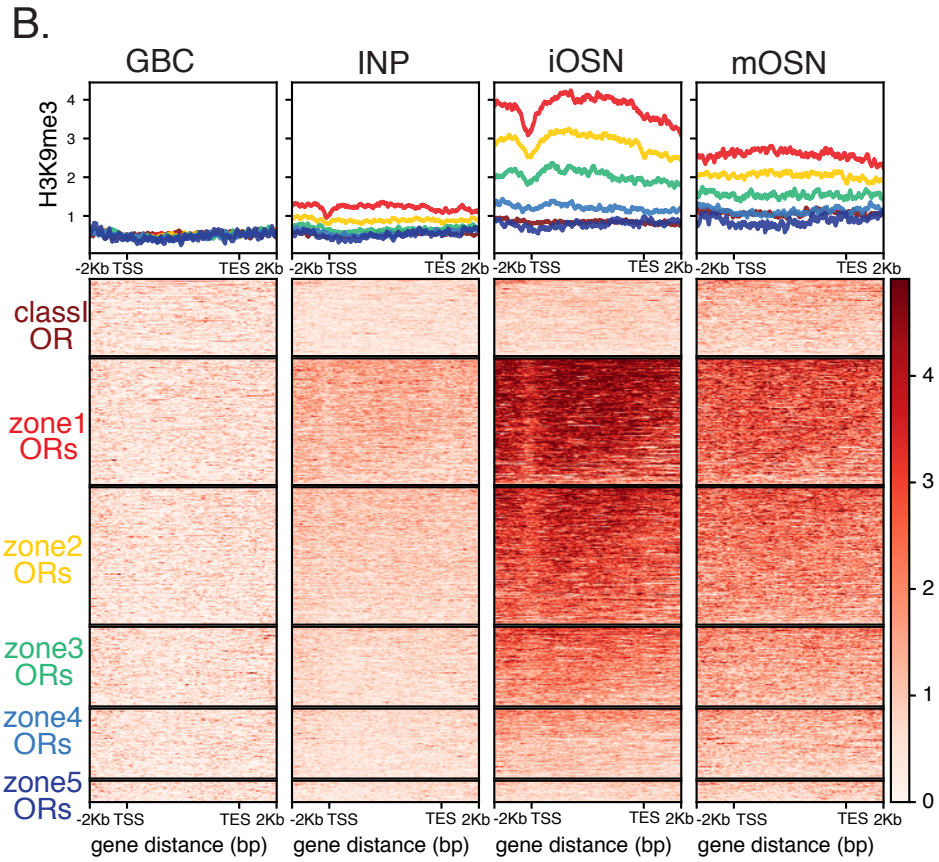


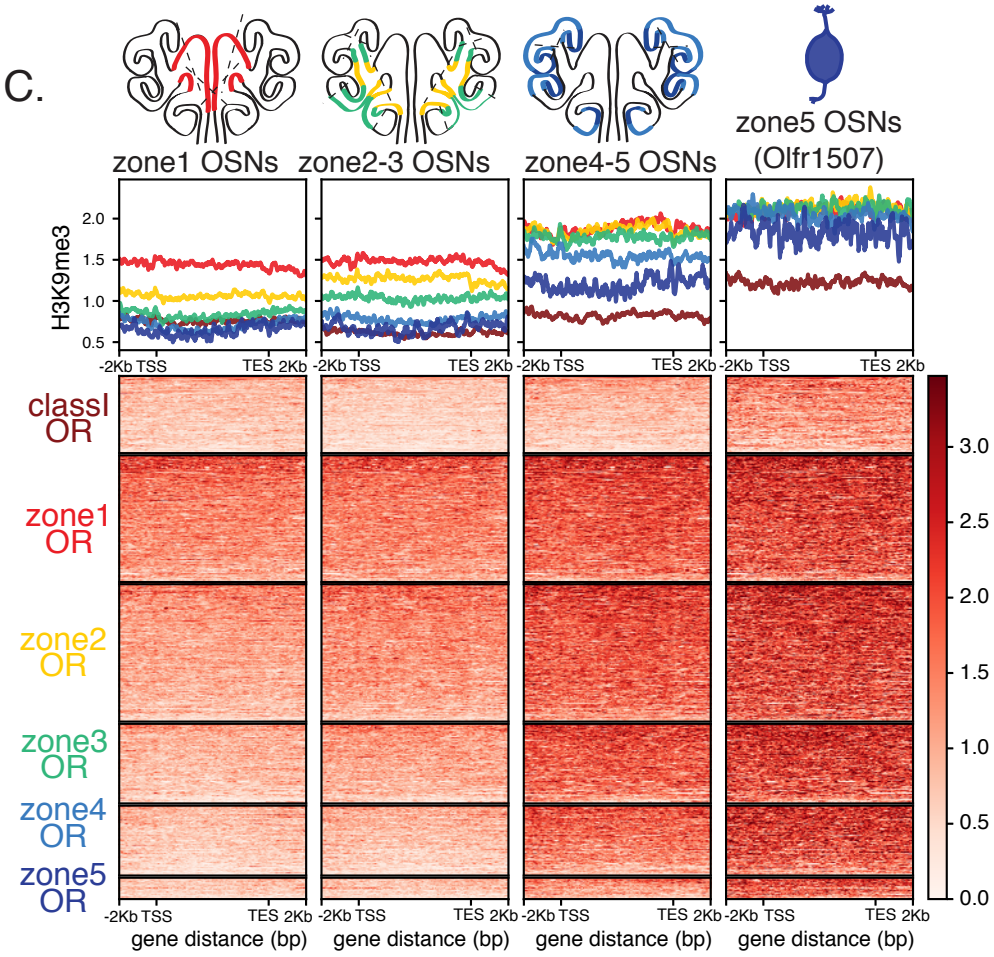
D.

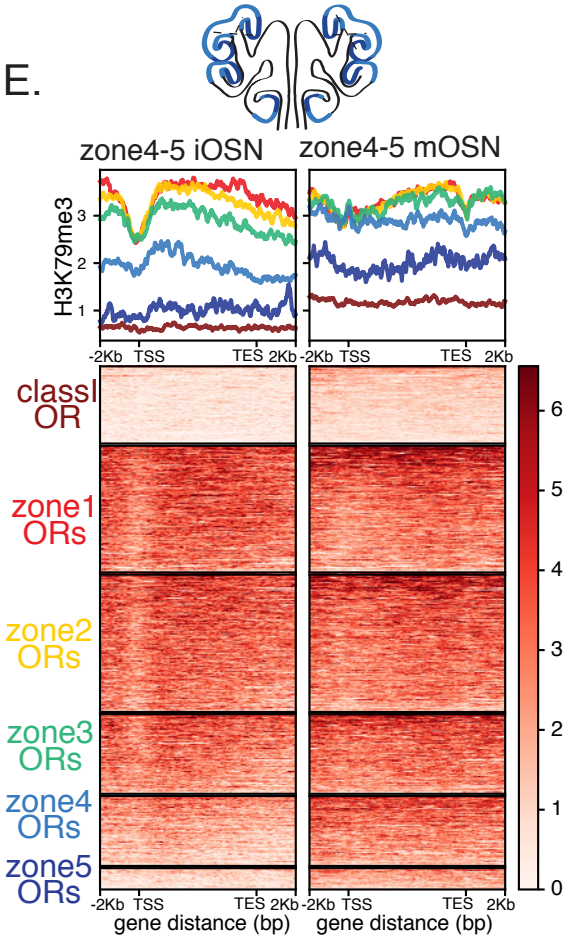


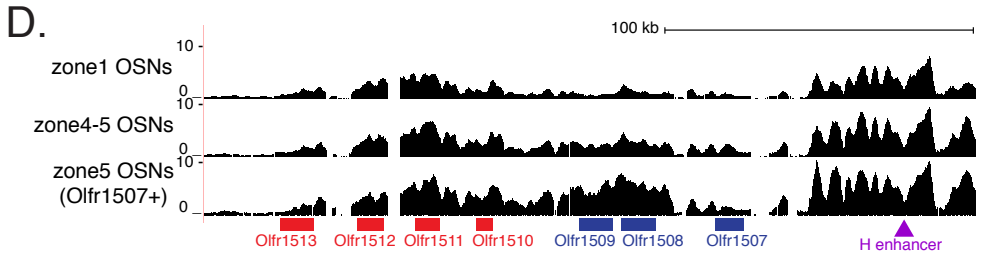
Supplemental Figure S2



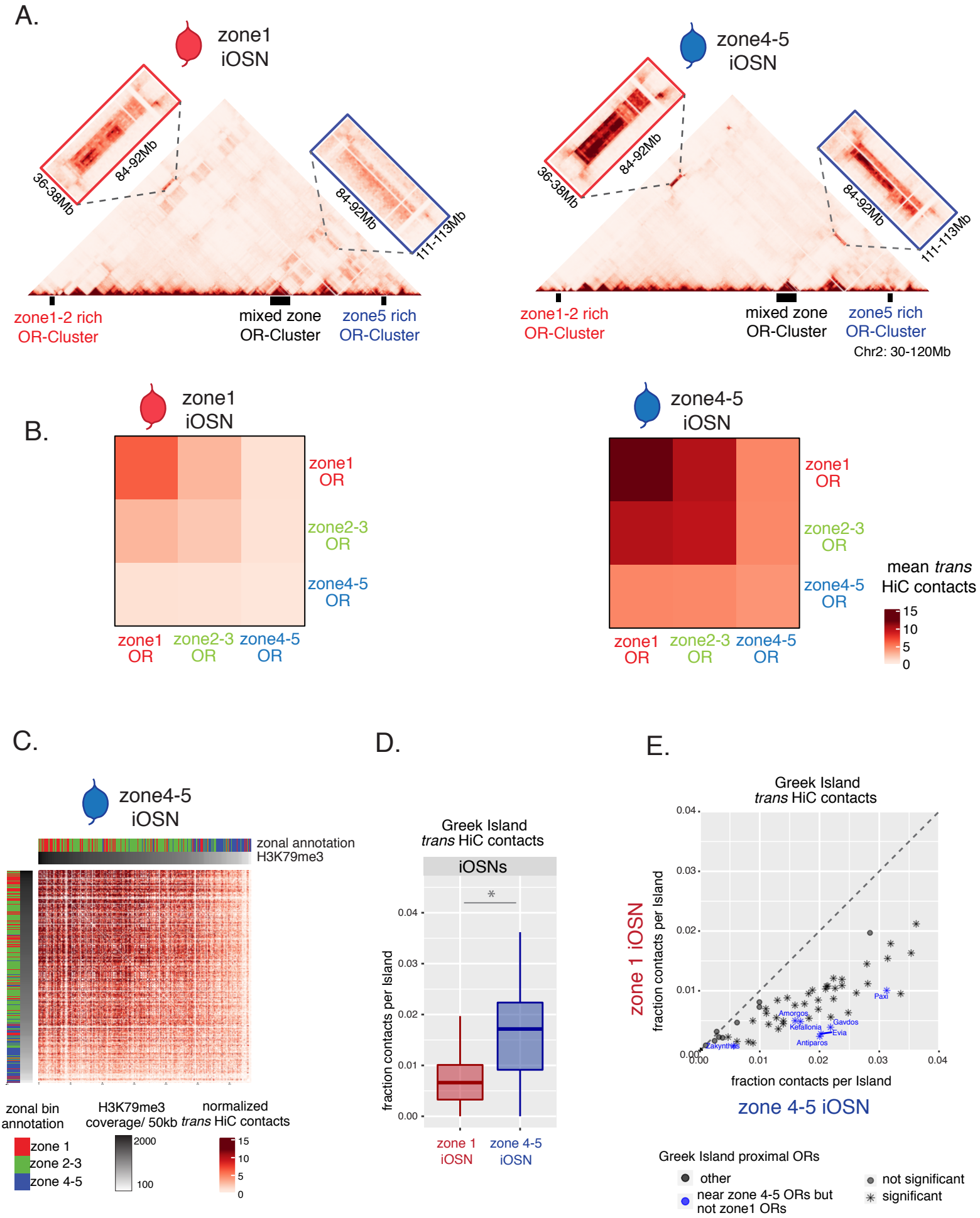






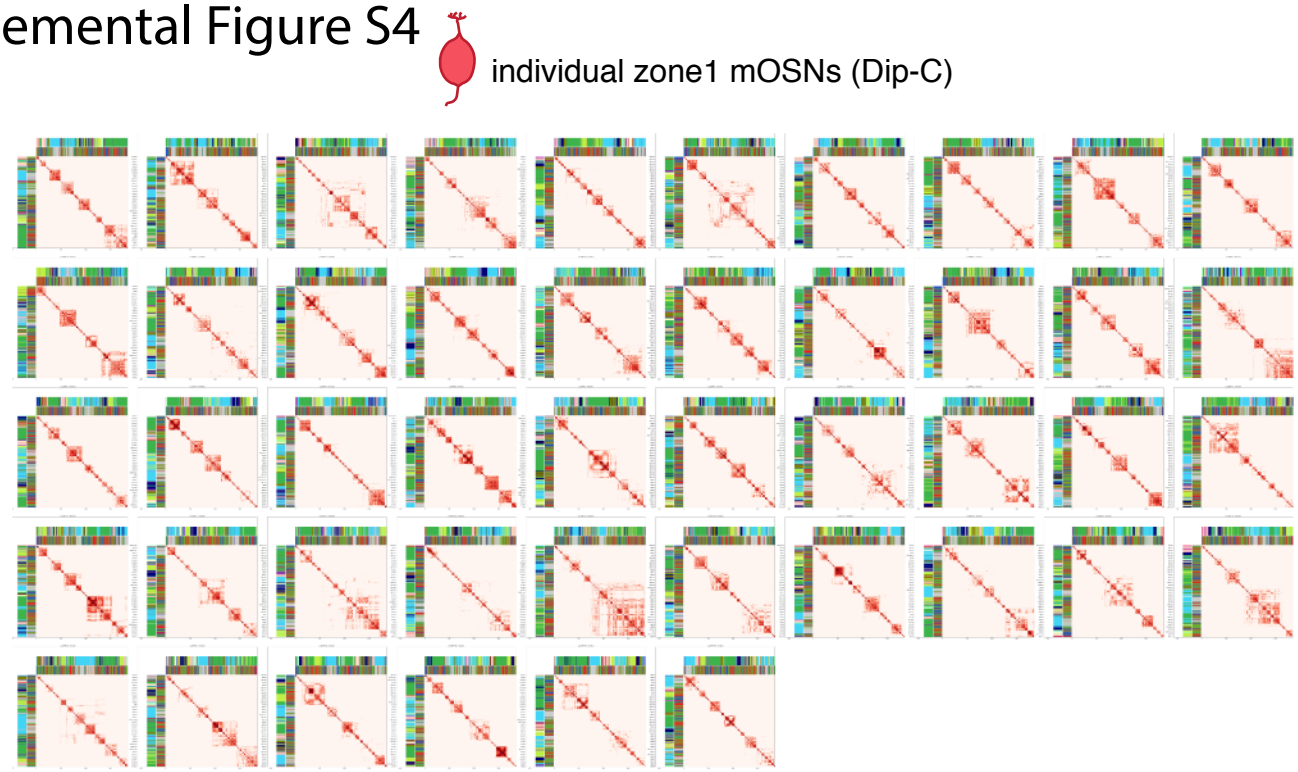


Supplemental Figure S3

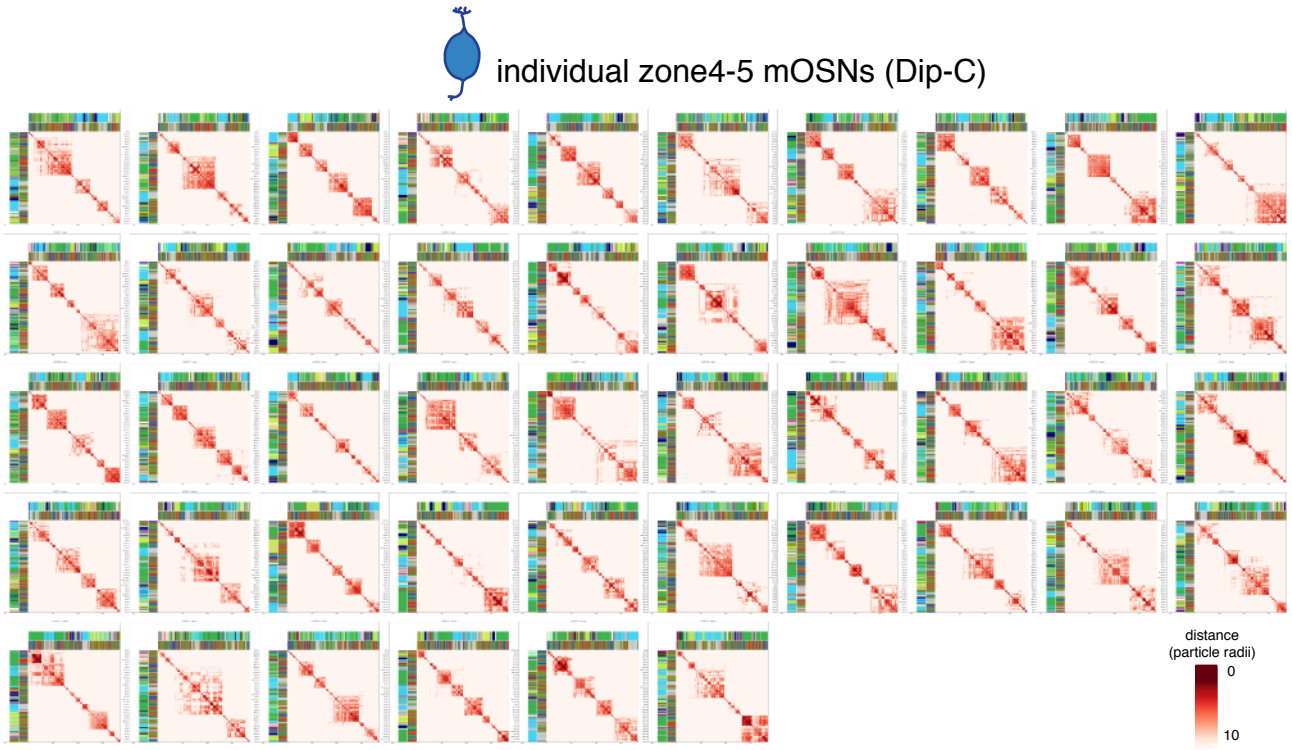


Supplemental Figure S4

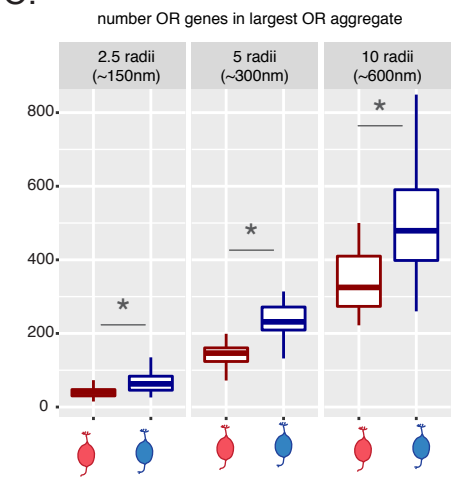
A.



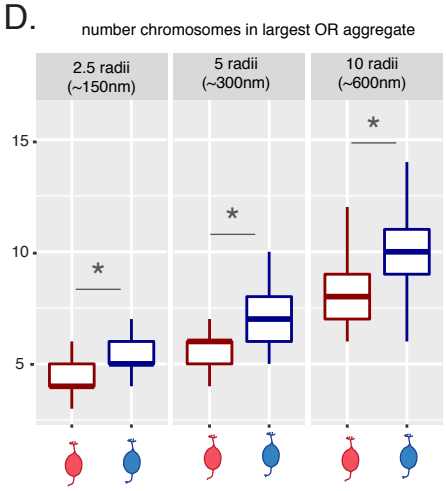
B.



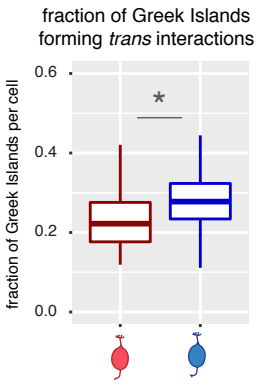
C.



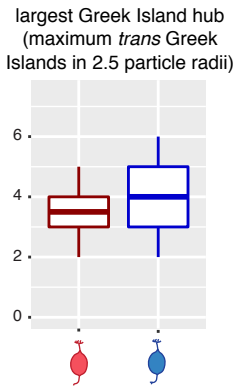
D.



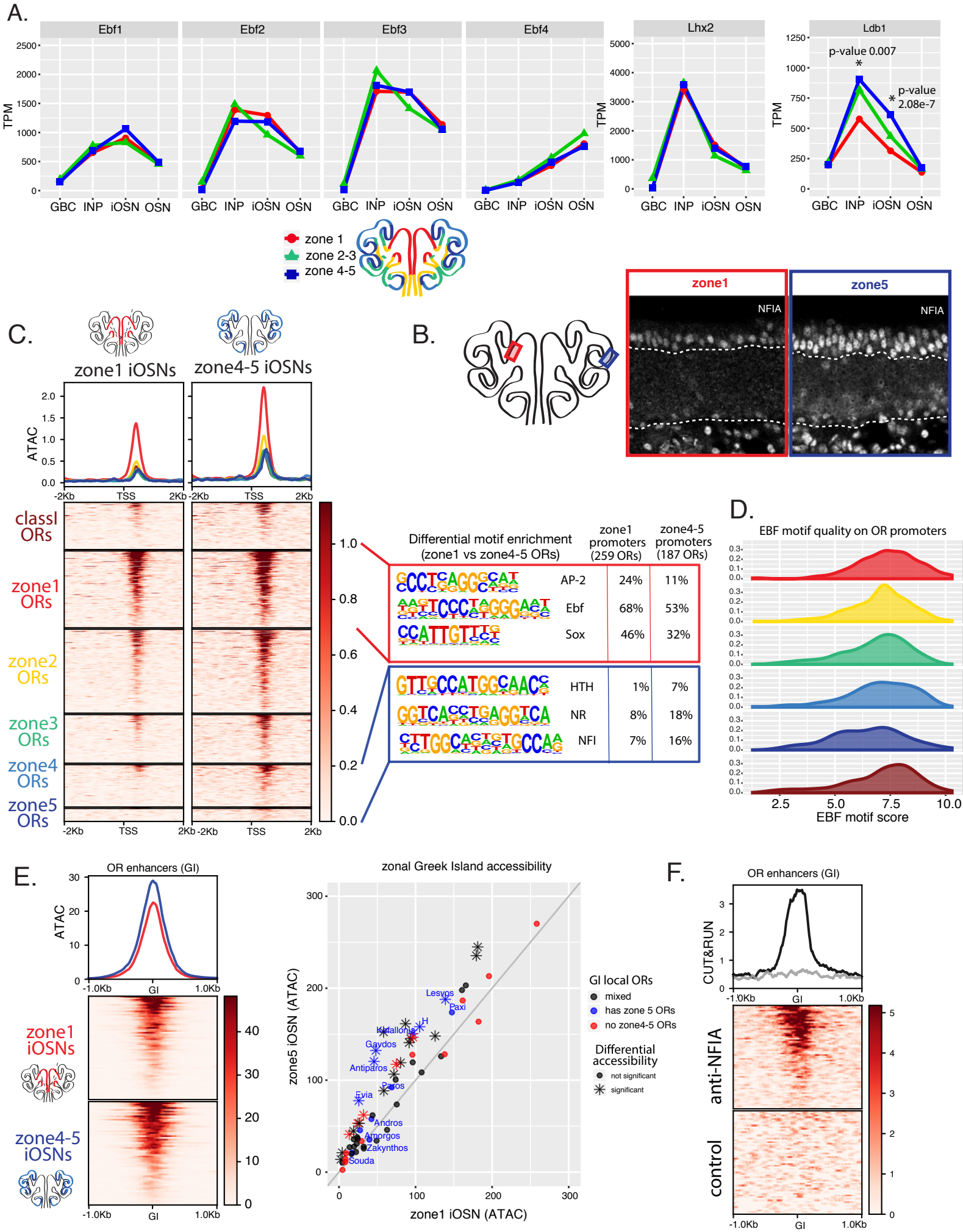
E.



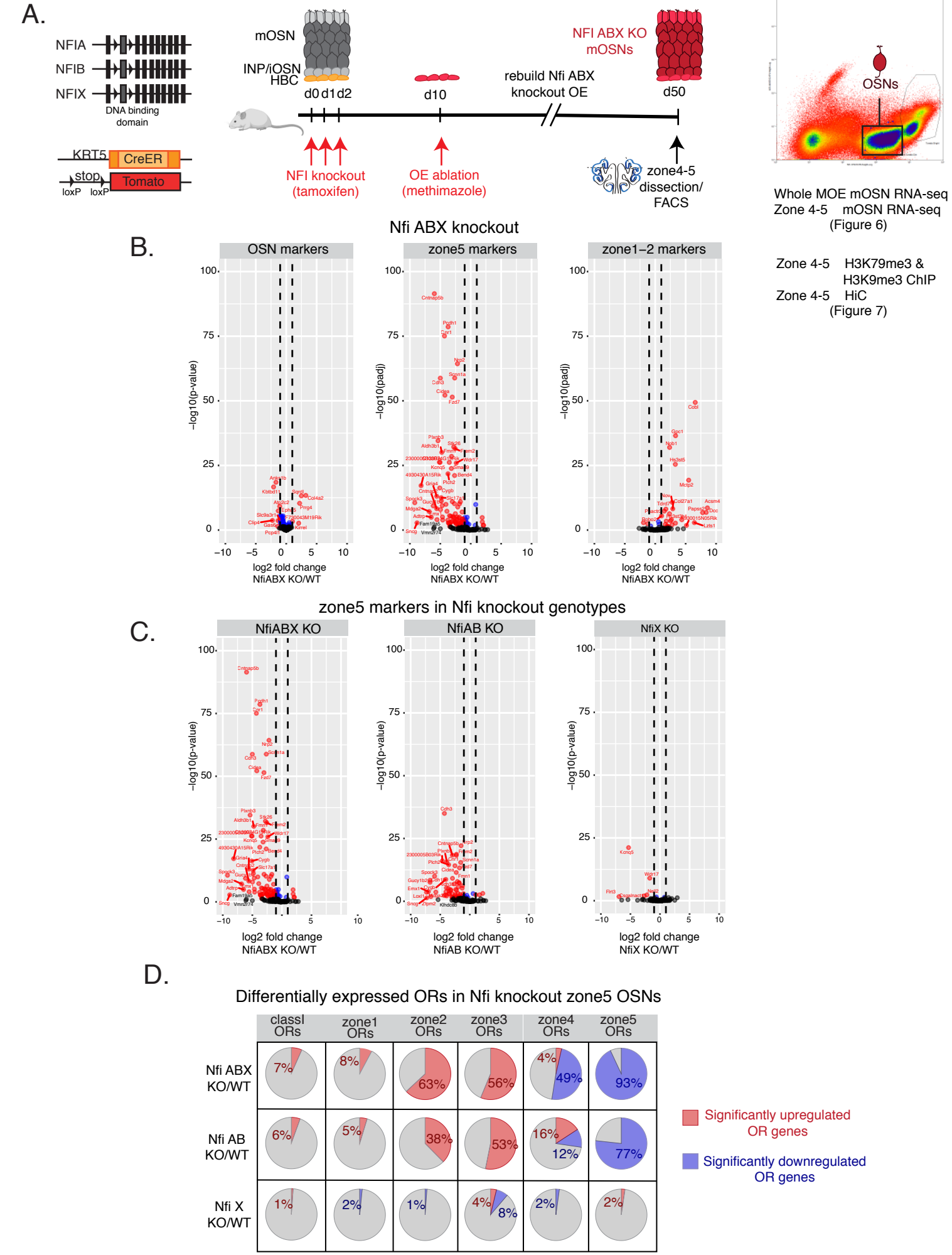
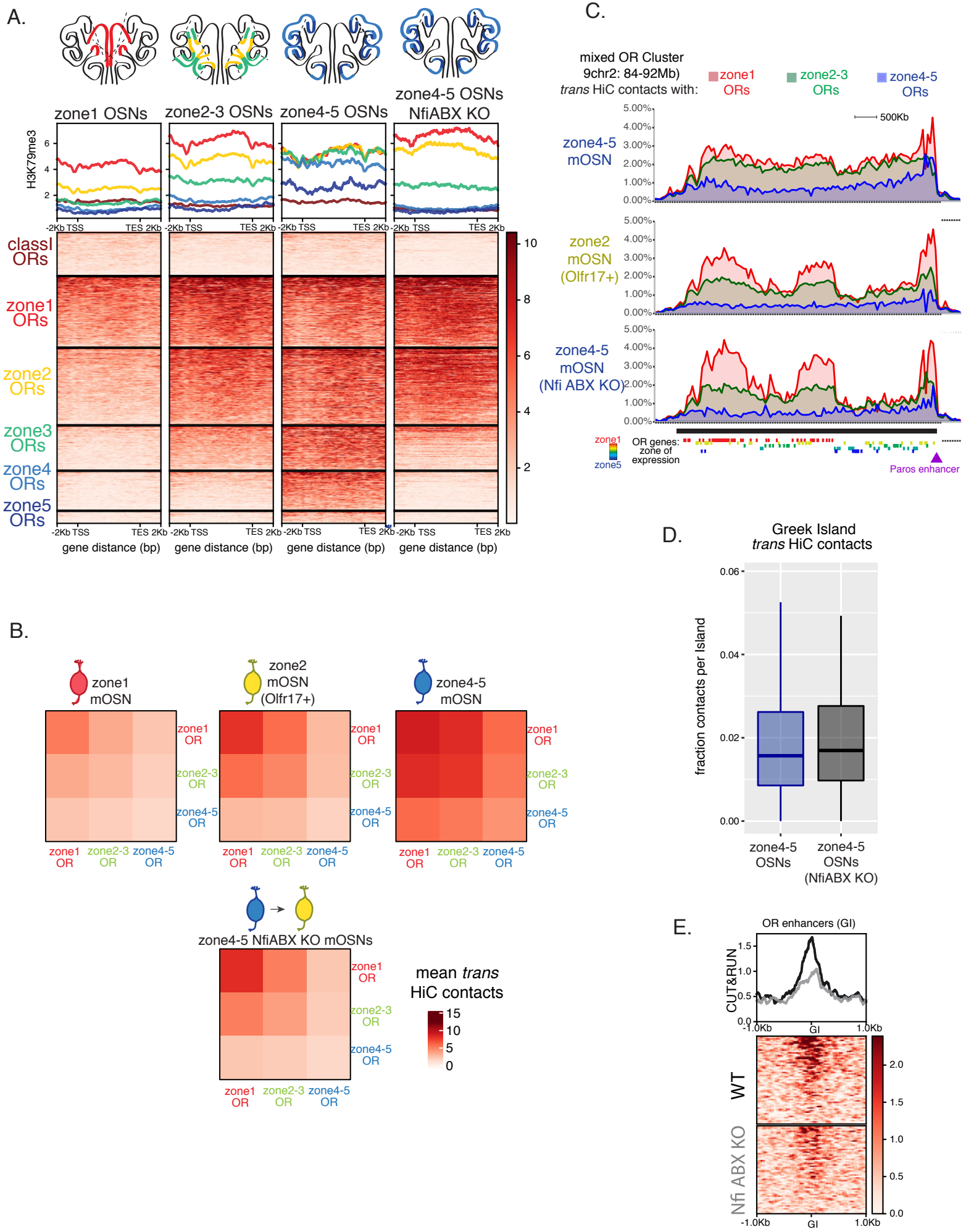
F.



Supplemental Figure S5



Supplemental Figure S6



Supplemental Figure S7