ANALYSIS OF POSITIVE AND NEGATIVE REGULATORY SITES AND THEIR INTERACTING TRANSCRIPTION FACTORS IN THE ZEBRAFISH OR101-1 GENE PROMOTER

by

Metin Özdemir

B.S., Molecular Biology and Genetics, Boğaziçi Unversity, 2014

Submitted to the Institute of Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree
Master of Science

Graduate Program in Molecular Biology and Genetics
Boğaziçi University
2018

To French Fries

ABSTRACT

ANALYSIS OF POSITIVE AND NEGATIVE REGULATORY SITES AND THEIR INTERACTING TRANSCRIPTION FACTORS IN THE ZEBRAFISH OR101-1 GENE PROMOTER

Each olfactory sensory neuron (OSN) of olfactory system expresses a single allele of an olfactory receptor (OR) gene from a large genomic repertoire. How an OSN selects and activates a single OR allele is not well understood yet. It was shown that epigenetic modifications are a key part of the process as well as proximal promoter region and long range cis-regulatory elements. Previous work on proximal promoter of OR101-1 gene has revealed both positive and negative regulatory sequences, deletion of which altered the expression efficiency in zebrafish embryo olfactory epithelium (OE) of transgenic vectors. It was shown that deletion of the intron from 1.2 kb upstream of OR101-1 transcription start site reduced the expression of transgenic construct, yet it was not clear whether this effect was caused by the absence of a putative O/E binding site within the intron site or by the absence of intron directly. To test this, three separate constructs were prepared. These constructs were injected in zebrafish embryos and expression efficiency was observed. While mutation of putative intronic O/E binding site resulted in 3% reduction, double mutation resulted in 7.5% reduction and swapping of intron with another resulted in 4% increase in expression, and after all, it was concluded that the expression reduction effect of intron deletion was due to the absence of the intron and not absence of the putative O/E binding site. A negative regulatory element with 562 bp length (i562) was also investigated in this study. Different length probes for each candidate site were prepared and electrophoretic mobility shift assay (EMSA) was applied to these probes with various protein sources and with anti-Zbtb7b antibody. Overall, it was seen that previously observed expression inhibition was not due to direct recognition of the candidate sites by any protein present in OE of zebrafish. The pattern of DNA-protein interaction in EMSA analyses suggests that there are cryptic sites within i562 participating in inhibitory function of the element.

ÖZET

ZEBRABALIĞINDA OR101-1 GENİNİN PROMOTÖRÜNDEKİ POZİTİF VE NEGATİF DÜZENLEYİCİ BÖLGELERİN VE BUNLARLA ETKİLEŞEN TRANSKRİPSİYON FAKTÖRLERİNİN ANALİZİ

Koku alma sisteminin her bir koku duyu nöronu geniş bir genomik repertuvardan bir koku almaç geninin tek bir alelini ifade eder. Bir koku duyu nöronunun, tek bir koku almaç alelini nasıl seçip etkinleştirdiği henüz yeterince iyi bilinmemektedir. Yakınsal promotör bölge ve uzun menzil cis-düzeneleyici unsurlarla birlikte epigenetik değişikliklerin bu süreçte anahtar bir rol oynadığı gösterilmiştir. Daha evvel OR101-1 geninin yakınsal promotörü üzerinde yapılan çalışmalar, delesyonuyla zebrabalığı embriyosu koku alma epitelindeki transgenik vektör ifadesini değiştiren pozitif ve negatif düzenleyici dizileri ortaya çıkardı. OR101-1 transkripsiyon başlangıç mevkiinin 1.2 kb yukarı bölgesinden intronun delesyonu, intron içinde yer alan varsayımsal O/E bağlanma mevkiinin eksikliğinden mi yoksa doğrudan intronun eksikliğinden mi kaynaklandığı bilinmeksizin, transgenik vektörün ifadesini azalttığı gösterilmiştir. Bu durumu test etmek amaçlı üç ayrı vektör hazırlanmıştır. Bu vektörler zebrabalığı embriyolarına enjekte edilmiş ve gen ifadesi verimini gözlemlenmiştir. İntronsal O/E bağlanma mevkiinin mutasyonu %3 ifade azalımına neden olurken, çifte mutasyon %7.5 ifade azalımına enden olmuş, başka bir intron ile yer değiştirim ise %4 ifade artımına neden olmuştur ve bunlara dayanarak intron delesyonunun ifade azalımındaki etkisinin varsayımsal O/E bağlanma mevkiinin yokluğundan ziyade intronun yokluğundan kaynaklandığı sonucuna varılmıştır. Bu çalışmada, 562 baz çifti uzunluğundaki negatif düzenleyici unsur (i562) da incelenmiştir. Her aday mevkii için farklı uzunluklarda problar hazırlanmış ve bu problara farklı protein kaynaklarıyla ve karşı-Zbtb7b antikoruyla elektroforetik hareketlilik öteleme tahlili (EMSA) uygulanmıştır. Nihayetinde, daha önce gözlenen ifade baskılanmasının, aday bağlanma mevkilerinin zebrabalığı koku alma epitelinde var olan proteinler tarafından doğrudan tanınmasından kaynaklamadığı gözlemlenmiştir. EMSA analizlerindeki DNA-protein etkileşim örüntüsü, i562'nin baskılayıcı işlevinde görev alan gizil mevkilerin varlığına işaret etmektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

5'SS	5' Splice Site
5' UTR	5' Untranslated Region
Adcy3	Adenynyl cyclase 3
AOB	Accessory Olfactory Bulb
ATF5	Activating Transcription Factor 5
bp	base pair
BTB ZF	Broad Complex, tramtrack, bric-a-brac Zinc Finger
cAMP	cyclic Adenine Monophospate
CDS	Coding Sequence
del	deletion
del100	100 bp probe with CNTCTGG deletion
del200	200 bp probe with CNTCTGG deletion
DNA	Deoxyribonucleic Acid
Dnmt3a	DNA methyltransferase 3a
eIF2a	eukaryotic Initiation Factor 2a
eEF4A	eukaryotic Elongation Factor 4A
eEF4G	eukaryotic Elongation Factor 4G
EJC	Exon-Exon Junction Complex
EMSA	Electrophoretic Mobility Shift Assay
eYFP	enhanced Yellow Fluorescent Protein
HDAC	Histone Deacetylase
H3K9me3	Histone 3 lysine 9 trimethylation

H3K20me3	Histone 3 lysine 20 trimethylation
i562	562 bp inhibitory region
i500	500 bp inhibitory region
kDa	kiloDalton
Lhx2	LIM-homeobox protein 2
LSD1	Lysine Specific Demthylase 1
mRNA	messenger ribonucleic acid
MOE	Main Olfactory Epithelium
MOR	Murine Olfactory Receptor
OB	Olfactory Bulb
OCAM	Olfactory Cell Adhesion Molecule
OE	Olfactory Epithelium
O/E	Olf1/Ebf1
OR	Olfactory Receptor
OSN	Olfactory Sensory Neuron
PCR	Polymerase Chain Reaction
Pol II	(RNA) Polymerase II
pre-mRNA	premature messenger Ribonucleic Acid
RNA	Ribonucleic Acid
RP58	Repressor Protein with a moleculer weight of 58
snRNP	small nuclear Ribonucleoprotein
TAAR	Trans-amine Associated Receptor
TFIIF	Transcription Factor IIF
TSS	Transcription Start Site

uaRNA	upstream anti-sense Ribonucleic Acid transcript
UPR	Unfolded Protein Response
VR	Vomeronasal Receptor
wt	wild type
wt100	wild type 100 bp probe
wt200	wild type 200 bp probe
Zbtb7b	BTB domain containing Zinc finger 7b
Zbtb18	BTB domain containing Zinc finger 18

1. INTRODUCTION

1.1. The Olfactory System

Chemosensation is one of the most ancient sensory mechanisms and describes an organism's ability to detect chemical cues from the environment. Chemosensation, in one form or another, is seen essentially throughout all branches of life from single-celled bacteria to multicellular organisms. Simple chemotaxis in bacteria evolved to become a 'binary' detection system in lower organisms, such as *C. elegans*, and a complex olfactory sensory system capable of distinguishing stereoisoforms of odorants in higher animals (Eisenbach and Bren, 2000; Bargmann *et al*., 1993; Reed *et al*., 1998). The system is based on sensing of chemicals released from conspecifics or the environment and provides an organism with a detailed view of food source, the availability of mates, or the presence of predators and other dangers (Prasad and Reed, 1999).

1.1.1. The Sense of Smell through Olfactory Receptors

The sense of smell is provided to the organism by the olfactory sensory neurons (OSNs) residing in the olfactory epithelium (OE) of the nasal cavity or related structures. OSNs express a specific subset of G protein-coupled receptors (GPCRs) called olfactory receptors (ORs), which comprise the largest gene family in many, if not all, organisms (Buck and Axel, 1991; Mombaerts; 2004; Shi and Zhang, 2007). The large heterogeneity of relevant chemical cues that are present in the environment is discussed as the possible driving force for the evolutionary expansion of the large OR superfamily. Accordingly, predictable variations in the number of functional OR genes are observed across species; humans possess slightly more than 400 functional OR genes in addition to a similar number of pseudogenes, while rodent genomes express up to 1.400 functional genes and possess a lower fraction of pseudogenes (Young *et al*., 2002; Niimura and Nei, 2007). The comparably small number of OR genes in some species, such as humans, does not necessarily reflect

poor sensory capabilities. On the contrary, primates including humans are shown to have a rather good sense of smell (Laska *et al*., 2000; Verbeurgt *et al*., 2014; Bushdid *et al*., 2014). This success of olfaction in spite of lower number of available OR genes is possibly related to the larger size of olfactory epithelium, which means a higher number of OSN per organism and related sensitivity, and the intense combinatorial nature of odorant detection by ORs, reflected by larger olfactory bulbs (OB; Olender *et al*., 2004; Shepherd, 2004; Maresh *et al*., 2008; Malnic *et al*., 2010).

1.1.2. The Anatomy and Function of the Vertebrate Olfactory System

In rodents, detection of odorants is performed by a specialized tissue called the main OE. The tissue resides in the nasal cavity of terrestrial mammals and therefore inhalation and olfactory sensation are linked allowing for the constant surveillance of the proximal environment for odorants (Song *et al*., 2008). The main OE in rodents house OSN which are ciliated neurons that express a single OR gene per OSN and project their axons to specific glomeruli in the OB. The nature of connectivity between the main OE and the OB results in a topographical one-to-one map of individual ORs to glomeruli in the OB (Vassar *et al*., 1994; Ressler *et al*., 1994; Mombaerts *et al*., 1996; Wang *et al*., 1998). In addition to the main OE, some organisms also possess an accessory olfactory sensory tissue called the vomeronasal organ (VNO), which detects pheromones and kairomones (Firestein, 2001). VNO neurons project their axons to a different region of the brain that is called the accessory olfactory bulb (AOB; Belluscio *et al*., 1999; Rodriguez *et al*., 1999). Different from the organization of the rodent olfactory system, the zebrafish has a single OE that houses different OSN subtypes and is structurally organized as a rosette-like tissue formed by several folding of lamella along a central midline raphe (Hansen and Zeiske, 1998).

The single zebrafish OE houses at least four different types of OSN, the ciliated, microvillous, crypt, and kappe neurons (Hansen and Zeiske, 1998; Hamdani and Doving, 2007; Ahuja *et al*., 2014). All OSN types innervate the same OB, yet functional, molecular, and anatomical separations similar to the distinction of the main and accessory olfactory system can be seen (Hansen and Zeiske, 1998; Ahuja *et al*., 2014; Sato *et al*., 2005).

Although morphologically different OSN subtypes occupy neighboring positions in the OE they segregate in the apical-basal axis of the tissue (Sato *et al*., 2005; Ahuja *et al*., 2013). By far the largest OSN populations in zebrafish OE are the ciliated and microvillous OSNs. Ciliated OSN have their soma close to basal portion of the tissue and project their long, ciliated dendrites to the apical surface of the tissue, while microvillous OSN have shorter dendrites harboring microvilli on and thus have a more apically positioned cell body (Thommesen *et al*., 1983; Yamamato and Ueda, 1978; Hamdani *et al*., 2001; Hansen *et al*., 2003; Sato *et al*., 2005; Oka *et al*., 2011). Crypt and kappe neurons have a globular shape, lack obvious dendrites and therefore reside in the most apical layers of the tissue (Ahuja *et al*., 2014).

The morphologically different cell types also express different groups of chemoreceptor proteins. While ciliated OSNs express ORs and trace amine-associated receptors (TAARs), microvillous OSN express V2R vomeronasal receptors (Mombaerts *et al*., 1996; Wagner *et al*., 2006). It was shown that crypt neurons as a whole express a single receptor protein of ORA4 (a V1R vomeronasal related receptor) and project their axon to a single glomerulus in the OB (Oka *et al*., 2011). Though it is not clear which receptor is expressed by kappe neurons, it is thought to be a single receptor since axons of kappe neurons also innervates a single glomerulus in OB similar to crypt neurons (Ahuja *et al*., 2014).

1.2. The Structure and Expression of Olfactory Receptors

1.2.1. Types and Structure of Olfactory Receptors

The detection of odors in OE is mediated specific ORs or related chemosensory proteins, which are a subfamily of rhodopsin-like type A GPCRs. The repertoire typically comprises a large number of genes, ranging from mere 179 in zebrafish to 800 in humans and 1.500 in mouse (Buck and Axel, 1991; Young, 2002; Zhang *et al*., 2004; Alioto and Ngai, 2005; Jacoby *et al*., 2006; Niimura and Nei, 2007). One interesting aspect of OR genes

is the fact that the coding sequence (CDS) resides within a single exon irrespective of the number of introns; a circumstance that allowed for rapid and efficient gene duplication during evolutionary expansion of the family (Young and Trask, 2002). Although typically the largest family of chemoreceptors present in vertebrate genomes, ORs and related vomeronasal receptors are not the only receptor class. TAARs are another group of GPCRs, which detect biological mono- and di-amines (Liberles and Buck, 2006; Hussain *et al*., 2009). Vomeronasal receptors (VRs) are proteins responsible for the recognition of pheromones and kairomones and comprise two subclasses: type 1 and type 2 vomeronasal receptors (V1Rs and V2Rs) (Dulac and Axel, 1995; Firestein, 2001). While V1Rs are intronless genes with a single uninterrupted CDS similar to ORs and TAARs, V2R genes contains multiple exons and extended N-terminal domains characteristic of Calcium-sensing receptors. Around 100 V1R and 200 V2R genes are expressed in the rodent VNO (Zhang *et al*., 2004; Mombaerts, 2004a; Young *et al*., 2005; Shi and Zhang, 2007).

OR genes generally reside in clusters in the genome and often neighboring genes share high sequence similarities and orientation, indicating a shared evolutionary past (Rouquier *et al*., 1998; Sullivan *et al*., 1996; Barth *et al*., 1997; Dugas and Ngai, 2001). While gene duplications and associated expansion of OR clusters results increased numbers of OR genes, other factors, such as a change in habitat may result in the decrease of functional OR genes and their pseudogenization. Hence, though the human genome contains more than 1000 OR genes, a significant number of 72% of these are pseudogenes (Zozulya *et al*., 2001), probably because bipedal lifestyle exposed humans to a different chemosensory environment. Yet only a relatively small portion of approximately 10% of OR genes are pseudogenes in mouse and zebrafish, underlining the dependence of these organisms on the sense of smell and their available OR repertoire (Zhang and Firestein, 2002; Alioto and Ngai, 2005; Niimura and Nei, 2005).

On a larger scale, sequence similarity analyses shows that ORs can be classified into two large groups; Class I type (fish-like) ORs are the predominant type in aquatic vertebrates such as teleosts and form a smaller repertoire of 10% of all OR genes in mammals. Class I ORs detect water-soluble odorants and have been shown to mediate innate odorant responses

(Kobayakawa *et al*., 2007). Class II type ORs are abundantly found in terrestrial animals and commonly detect volatile odorants (Alioto and Ngai, 2005; Freitag *et al*., 1998). Interestingly land animals did not lose their fish-like ORs as 10% of all ORs of mouse belong to the class I family. Mammalian class I ORs reside in a single large cluster contain less pseudogenes compared to Class II type ORs of mouse and are coordinately regulated by a cluster control element (Zhang and Firestein, 2002; Zhang and Firestein, 2009; Iwata *et al*., 2017). Class II type ORs are absent from the teleost genome except for *or101-1* of zebrafish (Alioto and Ngai, 2005; Niimura and Nei, 2005)

1.2.2. Monogenic and Monoallelic Expression of OR Genes

The typically large size of the OR repertoire allows an organism to detect a wide range of chemical compounds, yet does not directly explain how an organism is capable to discriminate between different odorants. This complicated task finds a simple solution in *C. elegans*, which only distinguishes attractive from repulsive odorants by expressing many different receptors that detect odorants of equal value in a single sensory neuron (Bargmann *et al*., 1993). However, in higher organisms odorants are discriminated by a combinatorial principle, similar to color vision (Friedrich and Korsching, 1997; Bozza *et al*., 2004). A prerequisite for this type of coding of odor information is that each OSN expresses only one or a specific subset of ORs from the available repertoire.

Supportive evidence for the 'one neuron – one receptor' rule was not easy to provide due to the large number of OR genes and their sequence similarity. Early evidence of monogenic OR expression came from in situ hybridization experiments performed with degenerate family specific and gene specific antisense riboprobes. The number of cells tagged by family specific degenerate probe was equal to the sum of cells tagged with gene specific probes (Kubick *et al*., 1997). Later, double in situ experiments done for three different OR gene on mouse did not showed co-localization, supporting monogenic expression (Tsuboi *et al*., 1999). Probably the best evidence came from reverse transcription polymerase chain reaction (RT-PCR) experiments on single OSNs for which only a single OR could be detected (Malnic *et al*., 1999). Recent studies, however, have shown that OSNs

may express multiple OR genes at least during early developmental stages (Hanchate *et al*., 2015)

It has also been shown that OSNs only express one of the two possible alleles of an OR (Chess *et al*., 1994). One of the early proofs of monoallelic expression comes from the RT-PCR experiments on hybrid offspring from two strains of mice, which are polymorphic for the I7 OR gene. When polymorphism specific primers were used in RT-PCR on OSN, it was seen that a single OSN expresses either the maternal or paternal allele of I7 OR gene but never both (Chess *et al*., 1994). In a more direct way tagging of each M71 allele with different fluorescent proteins showed no co-localization of the tags (Li *et al*., 2004). Either GFP or β-galactosidase tagged MOR28 ectopic genomic integrations were analyzed and no significant co-localization of either transgenes or the innate MOR28 gene expression could be observed, even though the regulatory and coding sequences were the same (Serizawa *et al*., 2000). The combination of monogenic and monoallelic OR expression if commonly referred to as singularity of expression.

1.2.3. Zonal Expression and OR Selection Frequencies

Expression of OR genes is spatially restricted in the OE (Vassar *et al*., 1993; Ressler *et al*., 1993; Miyamichi *et al*., 2005). Initially, four different topological expression zones were described (Ressler *et al*., 1993; Vassar *et al*., 1994). Later, however, it was shown that more than four zonal expression domains exist and that individual ORs are expressed in a restricted but overlapping manner along the dorso-ventral axis of the OE (Iwema and Schwob, 2003; Miyamichi *et al*., 2005; Tsuboi *et al*., 2006). An exception are the class I ORs, which are exclusively expressed in the dorsal OE (Zhang *et al*., 2004; Tsuboi *et al*., 2006). Zonal organization appears to be a general feature of OR expression that is conserved over evolution. However, even though OR expression in the zebrafish OE was also shown to be zonally organized, similar to the rodent OE (Weth *et al*, 1996), it was shown recently that the expression pattern is the outcome of OSN regeneration and migration and fundamentally different from the mouse (Bayramli *et al*., 2017).

While zone-specific OR expression certainly reflects an important aspect of selective OR gene choice, the molecular mechanisms underlying this pattern of expression are still elusive. Studies on transgenic animals show the importance of sequences within OR promoter regions for the zonal expression (Vassalli *et al*., 2002; Rothman *et al*., 2005) and comparison of expression patterns and promoter elements suggest that specific transcription factor binding sites may have a role in the regulation of OR gene expression in the zone specific manner (Hoppe *et al*., 2003; Hoppe *et al*., 2006). In addition to candidate transcriptional regulators, cell adhesion molecules, such as olfactory cell adhesion molecule (OCAM) and Neuropilin I, have been shown to be expressed in a zonal pattern (Yoshihara *et al*., 1997).

Other peculiar difference in the expression of different OR genes is that they are not expressed in the same number of OSNs. One particular example is the MOR28 gene expressed in the ventral OE of the mouse. Despite the fact that there are around 200 OR genes expressed in the same zone, MOR28 gene expressing OSN's correspond to 10% of the cell population in this zone (Shykind *et al*., 2004; Bressel *et al*., 2016), suggesting that additional factors and mechanisms are involved in the regulation of OR gene expression.

1.2.4. Negative Feedback Loop in OR Expression

Once expression of OR gene by OSN's is achieved, cells retain the expression of the selected OR gene throughout their lifetime and expression of other OR genes is prevented. Thus, a mechanism that established and maintains singular OR gene expression must be in place. In one study, frameshift mutations in the CDS of the MOR28 gene were analyzed and it was shown that a functional OR protein is needed to inhibit other OR genes from being expressed by the same OSN (Serizawa *et al*., 2003). Only in the presence of a functional OR protein, negative feedback to prevent on further expression of other OR genes was achieved.

Expressed pseudogenes are naturally occurring genes with frameshift mutations. It was shown that OSN that expresses an OR pseudogenes initially, and therefore cannot establish the feedback signal undergoes “OR switching” or a “second OR gene choice” (Shykind *et al*., 2004). As a consequence OSNs that express a random functional OR gene as a second choice innervate different glomeruli in the OB (Serizawa *et al*., 2003; Feinstein and Mombaerts, 2004). Interestingly, second OR gene choice is class restricted and OSNs that initially express a class I OR pseudogene always switch only to functional class I ORs (Bozza *et al*., 2009). Similar, class II expressing OSNs would make a second choice among class II genes.

Recently a mechanism for the negative feedback signaling was proposed to be mediated by the unfolded protein response (UPR; Dalton *et al*., 2014). The study proposes that OR gene translation triggers UPR in the ER, which is followed by Perk activation and eIF2a phosphorylation. eIF2a activates the activating transcription factor 5 (ATF5) expression, which consecutively enables Adcy3 transcription. Adcy3 then inhibits UPR and downregulates lysine-specific demethylase 1 (LSD1), a protein shown to take part in the OR transcriptional initiation (Lyons *et al*., 2013), resulting in the locked OR gene choice.

1.2.5. Gene Regulation in Olfactory Epithelium

Relatively random nature of OR gene choice and how it is regulated remains one of the crucial question in olfactory sensory system studies. Initially it was proposed that the OR expression happens through genomic DNA alterations as it happens in the immune system, yet the idea was disproven by clonal mouse experiments via nuclear transfer of mouse OSN expressing M71 receptor, which eventually showed not M71 OR restricted expression that would be the case if the DNA alteration was the mechanism by which OR selection happens (Tonegawa, 1983; Jung and Alt, 2004; Eggan *et al*., 2004; Li *et al*., 2004). Later a possible trans-regulatory element called H element present in chromosome 14 of mouse was shown to interact with OR gene promoters on other chromosomes which lead to the idea of H element driven trans OR gene activation (Lomvardas *et al*., 2006). Yet, deletion of H element resulted in expression abolishment of MOR28 cluster residing in the same chromosome with

H element, indicating that the element is involved in regulation of only the nearby MOR28 cluster genes (Fuss *et al*., 2007; Nishizumi *et al*., 2007). Recently, however, more elements similar to H element in relation to its epigenetic marks in MOE are found in other chromosomes linked to OR clusters (Markenscoff-Papadimitriou *et al*., 2014). When teste in zebrafish embryos via microinjection 16 of these elements were shown to recapitulate H element like enhancement in expression (Markenscoff-Papadimitriou *et al*., 2014).

This cluster based model of OR gene choice, however, does not explain every aspect of the process, such as zonal expression or the expression of individual OR genes that are not part of any cluster. For instance, a transgenic MOR23 gene with a minimal promoter that was inserted into the genome ectopically acquired the expression patterns of the endogenous gene indicating the significant role of proximal promoter sequences in OR gene choice (Vassalli *et al*., 2002). The recapitulation of OR expression with minimal promoter was also observed for other genes such as M4, M71 and mOR262-12 (Qasba and Reed, 1998; Rothman *et al*., 2005; Zhang *et al*., 2007). The analysis of the MOR23 and M71 minimal promoter showed the presence of homeodomain and O/E binding motifs whose mutations resulted in drastic expression change if not complete abolishment (Rothman *et al*., 2005; Vassali *et al*., 2006; Vassalli *et al*., 2011). Also, appending a nine times repeat of a 19 bp long homeodomain motif from P element to upstream of MOR23 minimal promoter lead to expression expansion within the MOE resulting in 10% coverage of the tissue by the receptor (Vassalli *et al*., 2011). Based on protein expression levels in MOE, it was thought that Lhx2 and Emx2 could be the proteins involved in homeodomain based OR gene choice and it was shown that Lhx2 is necessary for regular olfactory sensory system development and deletion of the gene resulted expression abolishment of Class II type OR genes (Hirota and Mombaerts, 2004; Hirota *et al*., 2007). However, Emx2 expression was shown to have a modulatory role rather than a regulatory role in MOE; while its deletion resulted in expression decrease of some OR genes, it had the opposite effect for some other OR genes (McIntyre *et al*., 2008).

Another motif that is present in minimal promoter of M71 is Olf1/Ebf1 (O/E-1) binding motif, which, when mutated, results in change in zonal expression or even complete

absence of expression from the minimal promoter of M71 (Rothman *et al*., 2005). An earlier study on olfactory marker protein (OMP) promoter showed that 0.3 kb upstream of the gene is capable of expression in a similar manner with the endogenous promoter and it contains an O/E binding motif (Kudrycki *et al*., 1993). Yet the mutation of the motif did not result in any significant expression change in MOE but altered the expression in other neuronal cell populations (Kudrycki *et al*., 1998). In addition, the deletion of O/E-1 gene did not alter the morphology of the olfactory epithelium and expression of OMP and Golf were still observed in the tissue (Lin and Grosschedl, 1995). Nonetheless, deletions of O/E-2 and O/E3 individually, which are homolog genes of O/E-1 protein and are abundantly expressed in MOE, disrupted proper axonal projection, though the expression of OMP or a cyclic nucleotide gated channel, CNGA2 were not affected.

The comparison analysis based on upstream region of 198 OR gene shed light on the enrichment of homeodomain and O/E-like binding motifs within the proximal OR promoters (Michaloski *et al*., 2006). While the homeodomains showed a widespread distribution over the OR promoters, O/E-like sites were enriched in 200 bp upstream of OR gene transcription start sites. An additional analysis including V1R upstream regions not only resulted in discovery of O/E-like sites but also inhibitory motifs resembling RP58 motif. This is the first case in which an inhibitory motif is shown to be present in OR promoters which might eventually partake in epigenetic silencing of OR genes (Michaloski *et al*., 2011).

1.3. Properties of *or101-1* Gene

As explained above, Class I type ORs are fish-like receptors and Class II type OR are predominant in terrestrial animals. Interestingly, the zebrafish genome contains only a single OR, *or101-1* that is closely related to the class II clade (Alioto and Ngai, 2005; Niimura and Nei, 2005; Tinaztepe, 2009).

The zebrafish *or101-1* gene resides on chromosome 21 and is part of larger cluster formed by the OR115 family. The cluster is surrounded by the nucleoporin gene nup98,

nuclear speckle splicing regulatory protein1 gene (nsrp1) involved in alternative splicing of pre-mRNA and slingshot protein phosphatase 2 gene (ssh2b) participating in actin polymerization on the telomeric side. The *or101-1* gene has two exons of 196 bp and 2.546 bp length and the entire *or101-1* CDS resides in second exon and specifies a 316 amino acid long protein.

Promoter bashing experiments performed on the genomic upstream sequence of *or101-1* using different lengths of promoter sequences showed that 800 bp are sufficient to drive expression. Interestingly, inclusion of a 562 bp sequence (i562) located 2kb upstream of the translation start site significantly reduced expression (Söğünmez, 2012; Sancer, 2015). Since rather shorter promoter constructs were shown to be expressed efficiently it can be concluded that the minimal *or101-1* promoter could be only 212 bp long (Kazcı, 2015). This is similar to the effective 395 bp that drive MOR23 expression in mouse transgenes (Vassalli *et al*., 2002). Further analysis of i562 revealed two candidate binding sites with CNTCTGG sequence, which were suggests to be binding motifs for BTB ZF (Broad Complex, tramtrack and bric-a-brac, Zinc Finger) family protein Zbtb7b (Söğünmez, 2012; Michaloski *et al*., 2011). Deletion of these sites from i562 resulted in elimination of inhibition (Sancer, 2015).

2. PURPOSE

The expression of OR genes by OSN depends on long-range and short-range cis-regulatory elements flanking OR gene loci. While OR enhancers can be located up to 100 kb away from the expressed gene, proximal promoter elements may control additional features of OR expression in the spatial, temporal, or quantitative domain . The genomic upstream region of the zebrafish *or101-1* gene has been extensively studied in our laboratory to understand proximal promoter effects on gene expression. Two, opposing effects on expression from transgenic promoter constructs have been described for different regions of the *or101-1* gene promoter: regions close to or around the transcription start site typically positively promoted expression of reporter transgenes, while more distal sequences reduced efficiency of expression. The overall purpose of this study was to gain further insight into the mechanistic nature of positive and negative regulation of OR genes and to identify DNA-binding proteins that mediate these effects.

As a first aim of this study I wanted to determine the role of two putative O/E binding motifs, one residing upstream of the transcription start site and another within the 5'-noncoding intron, in positive modulation of expression. The sites were studied by site-directed mutagenesis and intron swap experiments in a transgenic expression approach in zebrafish embryos.

In a second aim I set out to study transcription factor binding to the i562 negative regulator located further upstream of the *or101-1* gene. The sequence contains two candidate binding motifs that were tested for DNA-protein interactions by electrophoretic mobility shift and competition assays.

3. MATERIALS AND METHODS

3.1. Methods

3.1.1. Fish

The zebrafish (Danio rerio) used in this study was obtained from a local pet shop and this wild type strain was referred as PS-WT. If not obtained from outside, adults and embryos were raised at Boğaziçi University Life Sciences Center (Vivaryum).

3.1.2. Equipment and Supplies

The list of equipment, chemicals and consumables are provided in Appendix A and Appendix B.

3.1.3. Buffers and Solutions

Any buffer or solution, if not supplied by a manufacturer directly, were prepared according to Sambrook and Russell (1989). Solutions used specifically for zebrafish were prepared according to Westerfield (1997).

3.2. Materials

3.2.1. Zebrafish Maintenance

Zebrafish of AB/AB, AB/Tü and PS-WT strains were kept at constant 28ºC following a 14 hours light and 10 hours dark cycle. Maximum of five fish in 1 L tank, fifteen fish in 3 L tank and fifty fish in 10 L tank were kept in an automated zebrafish housing system with

five stage filtration, aeration, UV sterilization, temperature control and water circulation functions (Stand Alone System, Aquatic Habitats, FL). The fish were fed twice per day with flake food in the morning and live brine shrimp larvae (Artemia sp.) in the evening. Fresh water for the fish was prepared by supplying 100 L of reverse osmosis water with 2 g sea salt, 7.5 g sodium carbonate and 0.84 g calcium sulfate.

To obtain zebrafish embryo/fertilized egg, fish couples were placed in mating chambers that are composed of three pieces: one outer tank to keep fresh water, one smaller tank with perforated bottom to keep fish and eggs apart upon spawning and thus prevent predation by parents and a separator to keep female and male fish separated. Fish couples for mating were placed in mating chambers the day before the experiment either before or after the evening feeding and kept in static water separated by a separator overnight. An artificial plant was also placed per tank both to provide the female fish for a place to lay eggs and to provide buffer zone in case of aggression. The next morning by the beginning of light cycle, the separator was removed to initiate mating. Upon spawning, fish were placed in another tank and eggs, which naturally sink to the bottom of the tank, were collected by Pasteur pipette into Petri dish to be used later in the experiments. Injection experiments were performed within approximately 20 minutes after the spawning, which is the time interval to catch an embryo in single cell stage.

3.2.2. Microinjection into Zebrafish Oocytes

The night before the experiment, fish couples were placed in mating chambers with a separator in between per tank and in the morning, the separator was removed to allow the fish to breed. Upon spawning, the eggs were collected, cleared from any unfertilized egg or debris and placed on agarose injection mold. After placing on mold, eggs were positioned such that fertilized cell of each were easily visible. A glass capillary was filled with an injection mix having 50 ng/μL of each DNA construct (100 ng/μL in total), 10 mM potassium chloride and 0.01% Phenol Red, placed in the nozzle of FemtoJet® Express pressure injector (Eppendorf), and made ready for injection by snipping the tip from the farthest possible place. Each egg was injected approximately 2 nanoliters of injection mix

and removed into E3 medium. Dead embryos were removed and the medium was refreshed every day until the monitoring stage of the experiment.

3.2.3. Polymerase Chain Reaction (PCR)

Polymerase Chain Reactions (PCR) were performed using either homemade Taq, OneTaq® DNA polymerases or Q5® High-Fidelity DNA polymerase (NEB) according to manufacturer's instructions. Each PCR reaction contained 1-100 ng of DNA template, 0.5 µM of each primer, 0.2 mM of each deoxynucleotide, 1X reaction buffer. In case of Colony PCR, the reaction was also supplied with magnesium chloride to a final concentration of 1.5 mM since the Taq reaction buffer does not contain any. A typical PCR program begins with 3-5 minutes of initial denaturation step at 95 ºC followed by 25 to 36 cycles of denaturation at 95 ºC for 30 seconds, annealing at the melting temperature of primers (calculated using online tool provided by NEB, https://tmcalculator.neb.com) for 30 seconds and elongation at 72 ºC for 1 minute/kilobase and ends by final elongation at 72 ºC for 2-5 minutes. Upon completion of PCR, reactions were either kept at 4 ºC or -20 ºC depending on next step or directly run on agarose gel. Initial denaturation step for colony PCR was kept 10 minutes long to ensure bacterial lysis and annealing temperature was set to 52 ºC irrespective of the primers. For OneTaq® DNA polymerase elongation temperature was set as 72 ºC, though the suggested elongation temperature for the enzyme mix is 68 ºC. When PCR optimization was necessary, elongation at 68 ºC was also tested. For Q5® High-Fidelity DNA polymerase, denaturation and annealing times were set as 15 seconds and elongation time was set as 30 seconds/kilobase.

3.2.4. Restriction Endonuclease Digest of DNA

Restriction digest reaction were performed using the enzymes manufactured by New England Biolabs, Promega and Fermentas. The reaction mixtures contained 1X reaction buffer, 5-10 units of restriction enzyme per microgram of DNA and 1X BSA if not included in reaction buffer. Depending on DNA amount present in the mix, reactions were kept at 37

ºC for 1-4 hours and if plausible, reactions were terminated through high temperature (either 65 or 80 ºC depending on the enzyme) incubation for 20 minutes.

3.2.5. Agarose Gel Electrophoresis

Agarose gel was prepared as 1% of weight to volume ratio in 1X TAE buffer in an Erlenmeyer flask. Solution was boiled at a microwave until agarose was completely dissolved. 4 μL of ethidium bromide was added for a volume of 40 mL (0.5 μg/mL). Samples, mixed with 6x loading dye (NEB) to a final concentration of 1x, were loaded into wells and gel was run at 100 V until fragments were separated as desired. Detection was done by Bio-Rad GelDoc XR System under UV.

3.2.6. Gel Extraction and PCR Purification

Extraction of DNA fragments from agarose gels after separation or purification of amplified DNA fragments from PCR was performed using the NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) according to the manufacturer's protocol. Later, eluted DNA was quantified with NanoDrop® Spectrometer.

3.2.7. Plasmid Isolation

Plasmids isolated were always used either for cloning or microinjections. For the plasmid isolation purposes Plasmid MiniGeneJet Isolation kit (Thermo Scientific) was used according to the manufacturer's instructions. For injection purposes, after the isolation was done, ethanol precipitation was performed as to remove any buffer component from the plasmid DNA. 3 M sodium acetate (pH 5.2) with a 1/10 volume of starting material and 3 times the starting volume of ethanol (100%) was added onto the plasmid isolate and incubated at -80 ºC at least for an hour, preferable overnight. After the incubation, sample

was centrifuged maximum speed at 4 ºC for 30 minutes. A 70% ethanol aliquot was precooled in -20 ºC and upon the end of centrifugation, precipitated DNA was washed twice with cold ethanol and centrifuged for 10 minutes. At the last stage, as much ethanol as possible was removed from sample and let the DNA air-dry. As final stage, desired amount of ddH2O was added, and to help DNA dissolve completely, sample was kept at 65 ºC for 10 minutes.

3.2.8. Ligation of DNA Fragments and Vector

Before ligation, if unknown, DNA concentrations were determined by NanoDrop® and using concentration information and the sizes of fragments to be ligated, ligation reaction was set with DNA fragments (1:3 ratio of vector to insert) to maximum total of 100 ng, 2 µL of 10X ligation buffer, 1 µL of T4 ligase (NEB) and ddH2O to a final volume of 20 µL. The reaction was kept either at 25 ºC for 1 hour or 18 ºC overnight. Upon completion of reaction, transformation of the ligate was performed.

3.2.9. Transformation

Competent cells were thawed on ice for 5 minutes. Maximum 100ng of plasmid DNA or ligation product was added to the cells and mixed. The mixture was incubated on ice for 30 minutes and heat shocked in a thermomixer at 42ºC for 60 seconds and the mixture was immediately transferred to ice for minimum 2 minutes. For the recovery of heat shocked competent bacterial cells, 500-1000 µl of fresh LB without any antibiotics was added and the transformation mixture was incubated for 1 hour at 37 ºC on an orbital shaker. Finally, 200-600 µl of the transformation mixture was spread on appropriate selection plates. Plates were incubated at 37ºC overnight. Next day, the success and the efficiency of the cloning was analyzed by comparing the numbers of bacterial colonies on the ligation plates to the negative control plate.

3.2.10. Competent Cell Preparation: Rubidium Chloride

A single colony of the Top10 or DH5α bacteria strain was selected and inoculated overnight at 37ºC in 5 mL LB medium on an orbital shaker. 500 µl of the overnight culture was used to inoculate 500 mL of fresh LB medium. Then the prepared culture was incubated at 37ºC on the orbital shaker until the OD550nm of the culture has reached a reading of 0.6. Afterwards, the bacteria culture was incubated on ice for 15 minutes and centrifuged at 3000 rpm for 10 minutes at 4ºC. The supernatant was removed, and bacteria were resuspended gently in the remaining supernatant. 500µl of CT1 (30 mM potassium acetate, 10 mM CaCl2, 50 mM MnCl2, 100 mM RbCl and 15% glycerol) solution was added and incubated on ice for 30 minutes. Centrifugation step was repeated, and the supernatant removed. 20µl of CT2 solution (10 mM MOPS, 75 mM CaCl2, 10 mM RbCl and 15% glycerol) was added to resuspend the pellet. Finally, the bacteria suspension was prepared in 50 µl aliquots and immediately shock frozen in liquid nitrogen. Aliquots were stored in -80ºC until transformation.

3.2.11. Molding or Imaging of Embryo

The molding of embryos were performed with 2% low melting agarose that was constantly kept at 42 ºC in bottle warmer (Weewell). First a cover slide (18 x 18 mm) was stabilized perpendicular to the ground by dropping small amount of low melting agarose between the side wall of Petri dish and cover slide and letting it solidify. As the agarose between cover slide and side wall of Petri dish was solidifying, embryos were collected in a drop of E3 (10 embryos at most at a time) and by providing MS222 and constantly checking the reflex movement of embryos, animals were sedated. To lower the MS222 concentration to non-lethal levels for long-term exposure, additional E3 was provided. After stabilization of cover slide, embryos were aligned on Petri dish next to and perpendicular to cover slide as close as possible, excess E3 medium was removed and liquid low melting agarose was poured on top of embryos enough to cover them. If present, the disturbances in embryo positions during agarose addition were corrected before agarose began to solidify. After solidification, the cover slide was removed by initially pulling it away from the side wall of

Petri dish together with molded embryos and then moving gently towards one site that is perpendicular to the embryos. Then molded embryos were placed on a larger cover slide (22 x 50 mm) their body perpendicular to the cover slide and a small amount of low melting agarose was added to fix the mold onto the cover slide. Later using Knete® cover slide was fixed to a slide frame so as to keep the cove slide falling from microscope stage. Embryos were visualized under the confocal microscope.

3.2.12. Dot Blot Analysis

Previously biotinylated DNA fragments of probes were used to prepare a serial dilution of 1/10, 1/100 and 1/1000. After, each dilution was dropped on positively charged nylon membrane and let air dry. The membrane later exposed to 120.000 microjoules of 254 nm UV radiation with Stratalinker® UV Crosslinker (Stratagene) and developed with streptavidin conjugated HRP. Later stages of development was carried out according to detection protocol of LightShift Chemiluminescent EMSA kit and imaged.

3.2.13. Non-Denaturing Lysis and Protein Extraction

Collected tissue or cells were initially washed with 1X PBS at 4 °C and centrifuged at maximum speed for 5 minutes. Then, 1X PBS was removed and cell lysis buffer (10 mM HEPES, 10mM KCl, 0.1 mM EDTA, 1 mM DTT, 1 mM PMSF, 0.5% NP-40) and solubilized. The tissue sample were chopped into small pieces before addition of cell lysis buffer. After resuspending the sample in cell lysis buffer, it was kept on ice for 15 minutes and centrifuged afterwards full speed at 4 °C for 5 minutes. The supernatant was kept as cytoplasmic protein extract. The pellet was supplied with nuclear lysis buffer then (20 mM HEPES, 400 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1% NP-40, 1 mM PMSF, Protease inhibitor (one tablet per 5 mL solution(Roche))), resuspended and incubated on ice for 15 minutes. After incubation, sample was centrifuged and supernatant was removed into another container, supplied with glycerol to a final concentration of 10% and kept as nuclear protein extract. Later protein concentration was measured with NanoDrop® spectrometer.

3.2.14. Electrophoretic Mobility Shift Assay (EMSA)

Each reaction was set as 1X FastDigest Green buffer, 1 µL of protein extract (or protein extraction buffer as negative control), 1 µL of salmon sperm DNA (1 µg/ µL), 1 µL of probe DNA (approximately 0.1 pmol) and ddH2O up to the final volume of 20 µL. The samples were incubated at room temperature for 20 minutes and then run on a 5% non-denaturing polyacrylamide gel. Then the gel was blotted onto positively charged nylon membrane with Trans-Blot® SD Semi-Dry Transfer Cell according to the manufacturer's instructions. Upon blotting, the probes were fixed onto the membrane by exposure to 120.000 microjoules of 254 nm UV radiation within Stratalinker® UV Crosslinker (Stratagene). Later the membrane was developed according to the protocol provided with LightShift Chemiluminescent EMSA kit and imaged. In competition assays, desired amount of competitor was added by adjusting the ddH2O amount. In supershift assays, 0.25 µL of antibody (85 mg/mL) was added before or after probe-protein extract incubation at room temperature and after addition of antibody reactions were kept on ice for an additional 10 minutes.

3.2.15. Preparation of Protein Samples and SDS Gel Electrophoresis

Protein lysates were prepared by adding 4X laemmli loading dye and boiling the mix at 95 °C for 5 minutes. Those lysates were loaded into the wells of SDS-gel prepared at BioRad Mini-PROTEAN Tetra Cell where resolving gel (10%) polymerized at the bottom and stacking gel (6%) at the upper part (Table 3.1). The amount of lysates loaded were calculated according to the NanoDrop® results and the lowest concentrated protein lysate was the limiting factor for how much protein would be loaded into each well. Lysates were loaded into the wells in the vertical gel electrophoresis chamber, filled with 1X running buffer and the system was run at 100V until the loading dye gets into resolving gel, and then at 120V until the loading dye reached to the end of the gel.

Table 3.1. SDS polyacrylamide gel content

	Separating gel		Stacking gel
Stock	10% Gel	Stock	6% Gel
Tris (pH 8.8)	2250 µL	Tris (pH 6.8)	375 µL
Acrylamide	2000 µL	Acrylamide	397.5 µL
ddH2O	1633 µL	ddH20	2175 µL
SDS(10%)	60 µL	SDS(10%)	30 µL
APS (10%)	60 µL	APS (10%)	30 µL
TEMED	3 µL	TEMED	7.5 µL

3.2.16. Blotting - Transfer from SDS Gel to PVDF Membrane

Blotting and PVDF membrane, cut in the size of the blotting pad, were prepared. PVDF membrane was activated in methanol for 10 seconds, then washed in dH_2O for 1 minute and kept in 1X transfer buffer until it was placed in the blotting system. Meanwhile, glasses of running apparatus were separated from each other, by leaving the SDS gel on one of the glasses. Under distilled water the stacking part was removed, and the gel was placed on the blotting paper. Under the blotting paper, there was one of the blotting cassettes and one blotting pad. Membrane was placed on the gel by paying attention not to leave any bubbles between blotting paper and gel, and between gel and membrane for an efficient transfer. For that purpose, those steps were performed in a container partially filled with transfer buffer, so bubbles could be removed easily. Another blotting paper and blotting pad was placed onto the membrane. After other part of the cassette was placed on the pad, clips were closed to compress the system. Transfer cassette was placed into the transfer chamber in the right orientation. Membrane should be at positive, and gel should be at the side of negative pole because proteins are negatively charged due to SDS. Transfer was performed

at 250 mAh for 1hour and 45min. Membranes were washed with 1X TBS-T (0.1% Tween 20) for several times to remove methanol.

3.2.17. Staining and Antibody Incubation

Prior to staining with antibodies, the membrane was blocked for 1 hour at room temperature with 5% BSA (w/v) prepared in TBS-T and then incubated overnight at 4^{o}C, in certain dilution of the desired antibody, also prepared in blocking solution (5% BSA). Membrane was washed 3 times with TBS-T with five-minute intervals before incubation with HRP linked secondary antibody, diluted in TBS-T. Secondary antibody incubation was performed at room temperature for 2 hours. Membranes were washed 3 times with TBS-T in five-minute intervals. Blocking, washing and antibody incubations were performed on a benchtop orbital shaker.

3.2.18. Coomassie Staining

After the proteins were run on polyacrylamide gel, the gel was removed from the glasses and placed in a container, washed once with ddH2O for a brief moment to remove possible remnants. Then, staining solution (0.1% Coomassie Brilliant Blue R-250, 50% methanol, 10% glacial acetic acid) was poured into the container and incubated up to 30 minutes. Afterwards, the staining solution was decanted and destaining solution (40% methanol, 10% glacial acetic acid) was poured into the container and membrane was incubated until clear patterning was observed. Finally, destaining solution was decanted, membrane was washed with ddH2O to remove remaining destaining solution and imaged.

4. RESULTS

In general, gene expression can be influenced by positive and negative regulation through specific transcription factors that bind cis-regulatory sequences within gene promoters or enhancers. Typically, the specific combination of these regulators determines the time and the place at which a given gene is expressed. In the olfactory system, however, genes coding for chemosensory receptors are expressed in a mutually exclusive fashion in otherwise identical OSNs (Sato *et al*., 2007). Therefore, the detailed molecular events that result in the expression of individual OR genes has been referred to as 'OR gene choice' (Vassalli *et al*., 2002; Fuss and Ray, 2009). Despite this fundamental difference, cis-regulatory sequences have been shown to influence expression of individual OR genes in a spatial and temporal fashion (Qasba and Reed, 1998; Vassalli *et al*., 2002; Vassalli *et al*., 2011). In the study presented in this thesis I use the zebrafish *or101-1* olfactory receptor gene locus as a model to characterize and functionally test candidate positive and negative regulatory sites within proximal OR promoter sequences.

In previous experiments, it was shown that a 3.5 kb genomic sequence upstream of the *or101-1* coding sequence is sufficient to drive expression of an eYFP-containing transgenic construct with high efficiency and cell type specificity (Söğünmez, 2012). Subsequent promoter bashing experiments using increasingly shorter genomic sequences with sizes of 2.5kb, 2.0 kb and 1.2kb in the same expression vector (these and similar constructs will be referred to as p3.5-eYFP, p2.5-eYFP, p2.0-eYFP, and p1.2-eYFP) resulted in the identification of a 562 bp region (referred to as i562) located between 1.981 and 1.419 bp upstream of the transcription start site (TSS), which negatively influences expression. When the i562 sequence was omitted from transgenic constructs construct, efficiency of transgene expression increased regardless of the overall length of the construct tested (Söğünmez, 2012). In addition to this negative regulatory site, deeper investigation of *or101-1* promoter-sequences revealed the presence of a homeodomain (HD) binding and two Olf1/Ebf1 (O/E) binding sites, which may be candidate positive regulators of or101-1 expression by analogy to their function in mouse OR promoters (Vassalli *et al*., 2002; Rothman *et al*., 2005; Vassalli *et al*., 2011; Plessy *et al*., 2012). However, shortening of the genomic upstream sequence to

0.8 kb, which excludes the candidate HD-binding site or deletion of the upstream O/E-binding site, which resides upstream of the or101-1 TSS) had only mild effects on expression of transgenic promoter constructs. A more profound and negative effect on transgene expression could be observed when the 5' intron was removed from the otherwise efficiently expressed p1.2-eYFP construct (Kazci, 2015).

4.1. Candidate Positive Regulatory Sites within the Proximal OR101-1 Promoter

or101-1 gene promoter revealed the presence of two candidate O/E-binding sites (Figure 4.1). Considering Since one of these two candidate regulatory sites is located within the 5' intron of *or101-1*, the observed decrease in expression of the intronless p1.2-eYFP transgenic construct may be due to the loss of function of this site. To further test the significance of the intronic O/E site for *or101-1* expression, p1.2-eYFP-derived transgenic constructs with mutations in the intronic and upstream O/E site were examined. An alternative explanation for the observed effect may be that the loss of other, so far uncharacterized, intronic sequences, or the absence of intron splicing *per se* negatively affect transgene expression. This possibility was further tested by swapping the *or101-1* gene intron with the first intron of the zebrafish *myod1* gene.

4.1.1. Expression of the Basic Construct p1.2-eYFP

Since the function of the *or101-1* gene promoter has previously been investigated by three different experimenters (Söğunmez, 2012; Kazci, 2015; Sancer, 2015), it was necessary to establish a common experimental baseline by comparing expression of the basic construct p1.2-eYFP (Figure 4.1). For this purpose, 5 individual injection experiments were performed by injecting a mixture containing equal amounts (50 ng/µl, each construct) of the test construct p1.2-eYFP and an efficiently expressed reference construct, pomp-mCherry, which contains the olfactory marker protein (OMP) gene promoter (Celik *et al*., 2002) and which can be expressed by all ciliated OSNs. The use of the pomp-mCherry reference constructs serves as an internal standard and corrects for inter-experimental variation in

injection efficiency and oocyte quality. In total 667 embryos (5 individual injection experiments: 68, 110, 274, 69, 146 embryos, respectively) have been injected, 257 of which survived until the 3rd day post fertilization (3 dpf). When the surviving embryos were screened for transgene expression, 147 embryos (7, 21, 49, 20, 50, respectively) were observed to be positive for the pomp-mCherry construct, 99 out of which (4,13,28,15 and 39 respectively) were double-positive for eYFP-expression from the p1.2-eYFP construct. Thus, the average frequency of double-positive expression for embryos was calculated to be 65.8 ± 4.5% (mean ± SEM). The double positive embryos were further examined for the number of cells expressing the p1.2-eYFP construct per olfactory epithelium (OE) relative to the number of cells expressing pomp-mCherry. This parameter has previously been termed penetrance of expression of the tested construct and has been shown to be a more sensitive indicator of efficiency of OSN-specific transgenic constructs (e.g. Kazci, 2015). A total of 174 OEs were analyzed for cell counts, and 60.3 ± 2.0% (mean ± SEM) of cells counted were positive for expression of both constructs.

Thus, the efficiency of transgene expression in injected zebrafish embryo at 3 dpf for the basic p1.2-eYFP construct of 66% in my injection experiments is similar to the frequency of 61 to 83% that were observed in previous studies (Söğunmez, 2012; Kazci, 2015; Sancer 2015). Therefore, results obtained for the modified constructs described below can be interpreted in the framework of existing data.

4.1.2. Expression of a Modified p1.2-eYFP Construct Lacking the Candidate Intronic O/E Site

After establishing expression frequency for wild type p1.2-eYFP construct, a series of 10 independent injection experiments was performed to test a modified version of the p1.2-eYFP construct that lacks a candidate positive regulatory site (Figure 4.2). Motif scans for known regulatory sites on the first 1.2 kb of sequence upstream of the *or101-1* translation start site revealed the presence of two candidate O/E binding sites, one positioned upstream of the TSS and second site positioned in the single 5' intron of the *or101-1* gene. The possible role of the upstream O/E site on *or101-1* regulation was studied previously (Kazci, 2015)

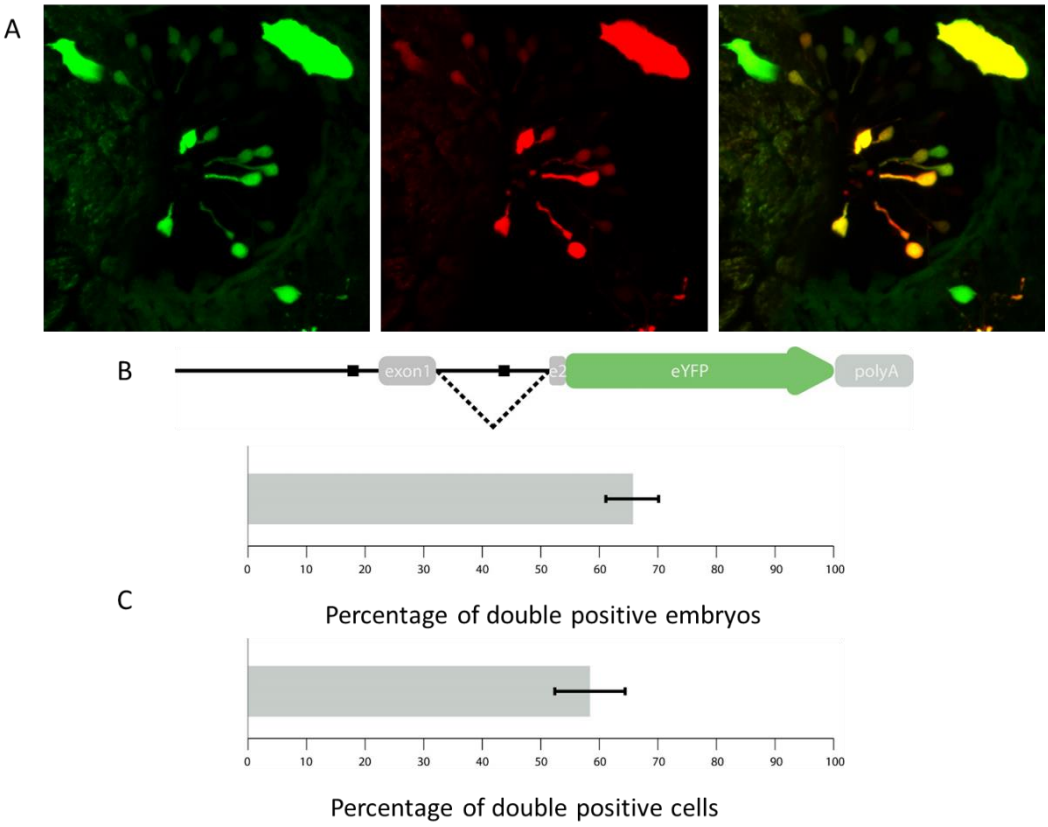

Figure 4.1. Transgene expression of the p1.2-eYFP construct.

and the site was shown to play a positive, yet minor role in transgene expression. Thus, the second intronic O/E motif may either be more important for *or101-1* expression or serve redundant function. To examine if the intronic O/E site has any significant regulatory function, 1.284 embryos (10 individual injection experiments: 130, 51, 26, 117, 127, 314, 87, 224, 131, 77 embryos, respectively) were injected with a mut-p1.2-eYFP, which lacks the intronic site, and pomp-mCherry constructs in equal concentrations. Out of those, 387 embryos survived until 3 dpf. Screening of surviving embryos showed that 194 embryos (22, 6, 3, 18, 16, 64, 21, 17, 6 and 21, respectively) were positive for pomp-mCherry and 104 of those (5, 4, 2, 13, 6, 32, 16, 8, 5 and 13 respectively) were double positive for mut-p1.2-eYFP. Thus, the average frequency of double expression was 58.4 ± 5.9% (mean ± SEM). As above, double positive embryos were analyzed for penetrance of expression by counting individual transgene-expressing cells per OE. In total, 173 OEs were analyzed, resulting in 57.2 ± 1.6% (mean ± SEM).

Thus, removal of the intronic O/E-like sequence motif reduced the frequency of expression by 7.4% at the level of transgene-expressing embryos and by 3% for cell counts. Yet, expression of the modified construct was still high, suggesting that additional unidentified regulatory elements might play a positive role in controlling expression. The difference in expression was not statistically significant ($p = 0.42$ for embryos, $p = 0.066$ for individual cells; Student's t-test, two-tailed, independent samples).

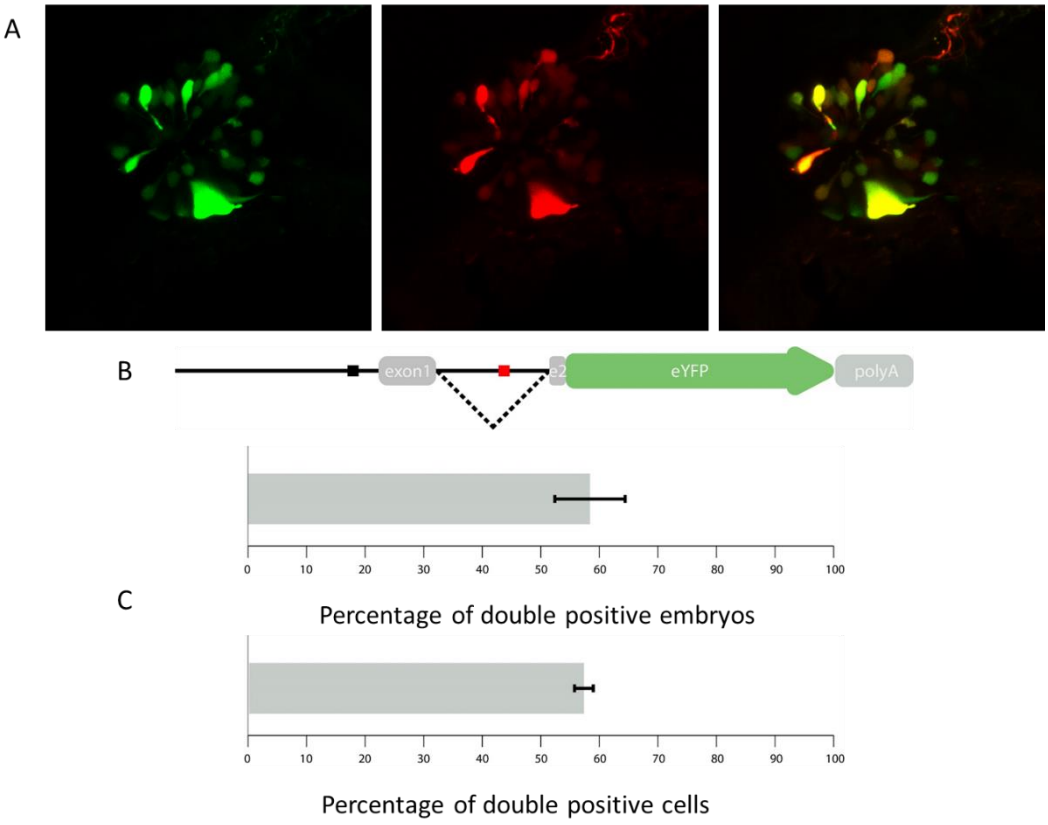

Figure 4.2. Transgene expression of the mut-p1.2-eYFP construct.

4.1.3. Expression of a Modified p1.2-eYFP Construct Lacking Two Candidate O/E Site

In addition to the single mutant mut-p1.2-eYFP construct described above, a double mutant version of the construct was examined to investigate the possible redundant function of upstream and intronic O/E sites (Figure 4.3). For this purpose, a total of 1.025 embryos

(6 individual injection experiments: 165, 230, 200, 186, 150, and 94 embryos, respectively) were injected with ddp1.2-eYFP, in which the upstream and intronic O/E sites were removed, and pomp-mCherry constructs. Out of those, 374 embryos survived until 3 dpf and were further analyzed for expression. Among the screened embryos, 285 (78, 32, 46, 43, 49 and 10, respectively) were found to be positive for pomp-mCherry and 146 of those (43, 17, 33, 22, 24 and 7 respectively) were double-positive for ddp1.2-eYFP. Thus, the average rate of expression of this construct was 58.3 ± 4.1% (mean ± SEM). Penetrance of expression was determined for 268 OEs, resulting in an average double expression rate of 52.3 ± 1.3% (mean ± SEM).

Thus, simultaneous removal of two candidate O/E binding sites reduced expression of the transgenic construct by 7.5% at the level of transgene-expressing embryos and by 8% at the level of individual cells. However, the difference was only statistically significant for the penetrance of expression but not at the level of embryos ($p = 0.25$ for embryos, $p = 0.0005$ for individual cells; Student's t-test, two-tailed, unpaired). The double-mutant construct was not statistically different from the single mutant construct ($p = 0.99$ for embryos, $p = 0.20$ for individual cells), despite the small decrease in cell number of 3%. Nevertheless, the (small) differences in expression between p1.2-eYFP and ddp1.2-eYFP and the overall (high) similarity between ddp1.2-eYFP and mut-p1.2-eYFP suggests that the intronic O/E site has a more prominent role in regulating *or101-1* expression than the O/E site contained within the genomic upstream region of the gene.

4.1.4. Expression of a Modified p1.2-eYFP Construct with a Swapped Intron

Previous experiments have shown that removal of the entire 5' intronic sequence from the p1.2-eYFP transgenic construct reduced the frequency of expression to 45% (Kazci, 2015). This effect could be due to the presence of regulatory sites within the intron, such as the O/E motif described above, or the fact that intron splicing has a general positive effect on gene expression (reviewed in Shaul, 2017). Studies in various eukaryotic systems *in vivo* and *in vitro* have shown that the presence of an intron can increase the expression of a protein several ten- to hundred-fold (Lu and Cullen, 2003). To further test this idea, an intron swap

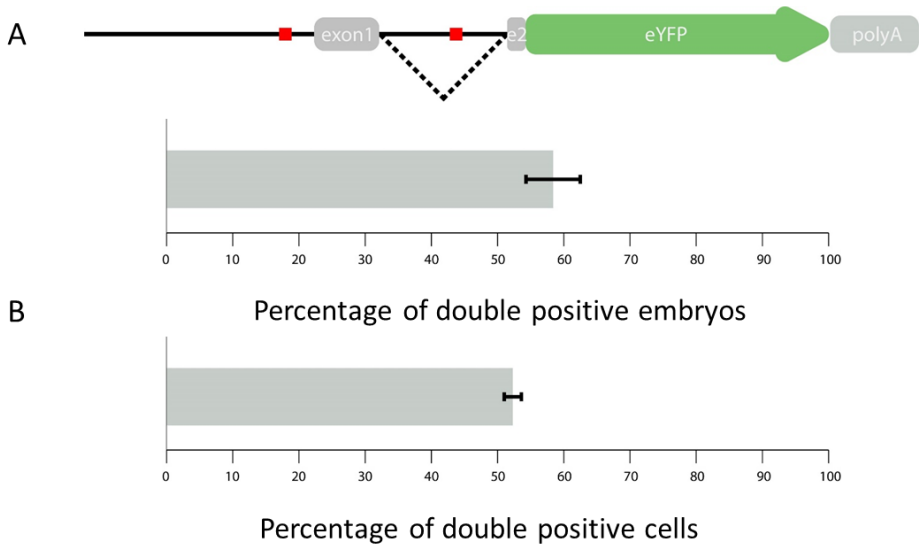

Figure 4.3. Transgene expression of the ddp1.2-eYFP construct.

of the native *or101-1* intron with the first intron of the myod1 gene (Zerbino *et al*., 2018) was generated in the context of the p1.2-eYFP construct (Figure 4.4). The rational is that the myoD1 intron may be spliced properly but would not contain regulatory sites that are essential for or influence *or101-1* expression. Using this construct along with pomp-mCherry, 858 embryos were injected in seven individual experiments (114, 165, 138, 47, 156, 115, and 123 embryos, respectively). Among the 199 embryos that survived until 3 dpf, 157 (13, 23, 26, 22, 57, 8 and 8 respectively) were positive for pomp-mCherry, and 105 of those (10, 16, 18, 14, 35, 5 and 7 respectively) were positive for the test construct p1.2-myod1-eYFP. Thus, the overall rate of expression was 70.1 ± 3.5% (mean ± SEM) at the level of transgene-expressing embryos. When double positive embryos were monitored for their penetrance of expression at the level of individual OSNs, it was found that 64.8 ± 1.5% (mean ± SEM) of cells expressing the pomp-mCherry were also expressing the p1.2-myod1-eYFP construct.

Surprisingly, replacement of the native *or101-1* intron with an unrelated intron slightly increased the expression compared to p1.2-eYFP by 4%, suggesting that the intron-splicing event itself may play an important role for effective expression of the gene. Yet, the

difference was not statistically significant at the level of embryos or cells (p = 0.47 for embryos, p = 0.96 for individual cells; Student's t-test, two-tailed, unpaired).

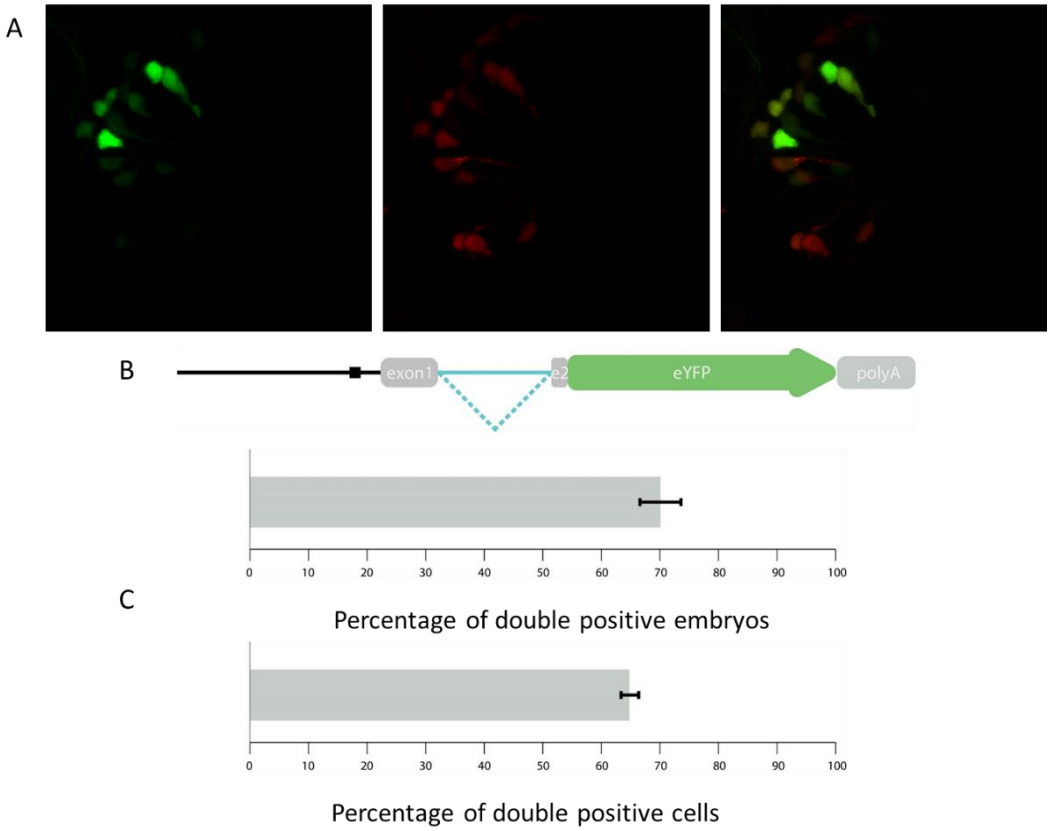

Figure 4.4. Transgene expression of the p1.2-myod1-eYFP construct.

4.1.5. Motif Scan on Myod1 Intron for Human and Mouse O/E Binding Sites

While choosing for a suitable intron for the experiments described in section 4.1.2, two constraints were set; first, the intron to be swapped into the construct should not be much longer than the native *or101-1* intron to keep possible distance effects to a minimum and second, the selected gene should not be expressed in OE tissue. Since myod1 is a myogenic factor involved in fibroblast to muscle cell differentiation (Tapscotr *et al*., 1988) it was expected not to contain an active regulatory element that drives expression in OSNs. In addition, its first intron has a size of 288 bp, comparable to the 345 bp-sized intron of *or101-*

1. Based on the observation that intron swapping enhanced the expression of p1.2-eYFP construct, even though insignificantly, it was suspected that the *myod1* intron might contain regulatory sequences that bear similarity to O/E binding sites. To assess the presence of candidate O/E sites within the *myod1* intron, a motif scan was performed using the JASPAR database and scanning tools (Stormo, 2013) for the MA0154.1 (mouse) and MA0154.2 (human) EBF1 motifs. The scan revealed the presence of sequences within the *myod1* intron with slight similarity to EBF1 binding sites, albeit with a lower score than the O/E binding sites within the native *or101-1* intron (Table 4.1).

Table 4.1. JASPAR motif scan for EBF1 binding sites in myod1 and or101-1 introns.

myod1 intron							
Matrix ID	Name	Score	Relative score	Start	End	Strand	Predicted sequence
MA0154.1	EBF1	799.218	0.840513214102	219	228	-	acacaagaga
MA0154.2	EBF1	-0.587437	0.801101224861	218	228	+	atctcttgtgt
OR101-1 intron							
Matrix ID	Name	Score	Relative score	Start	End	Strand	Predicted sequence
MA0154.1	EBF1	921.618	0.873951009225	199	208	+	cctttaggga
MA0154.2	EBF1	524.966	0.867482078981	199	209	-	ttccctaaagg
MA0154.2	EBF1	188.782	0.829250374361	198	208	+	acctttaggga

4.1.6. Statistical Comparison of Native p1.2-eYFP with Modified Constructs

To understand whether the observed expression differences between all p1.2-eYFP constructs were significant or not, Student's t-test was applied. When the t-test was applied to embryo counts, it was seen that none of the samples were significantly different from the others. Yet at the cell count level, some of the observed differences were statistically

significant while some other were not. The expression difference of 7.4% between p1.2-eYFP and mut-p1.2-eYFP has a p value of 0.221, which is statistically insignificant. Yet, when t-test was applied for p1.2-eYFP and ddp1.2-eYFP constructs, a p value of 0.004 was calculated meaning observed 7.5% expression difference is statistically significant. 4.5% expression difference between p1.2-eYFP and p1.2-myod1-eYFP again did not reach statistical significance (p=0.067).

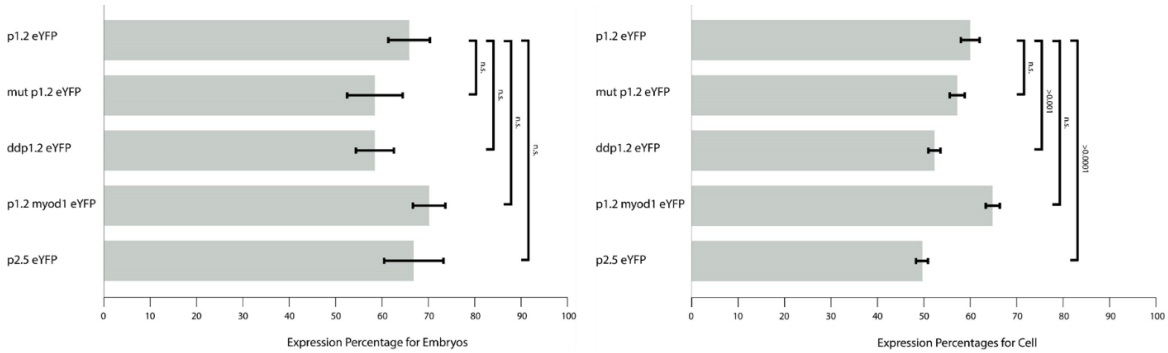

Figure 4.5. Summary and statistical analysis of p1.2-derived transgenic constructs.

4.2. Candidate Negative Regulatory Sites within the Distal OR101-1 Promoter

The experiments detailed above concern the contribution of candidate regulatory sites that have been shown to positively affect transgene expression from *or101-1*-derived promoter constructs. In addition a strong negative modulation of expression has been demonstrated for the i562 sequence (Söğünmez, 2012; Sancer, 2015) that is located further upstream of the *or101-1* gene and which is not included in the p1.2-eYFP and its derived constructs. Central to the inhibitory effect of i562 may be two binding motifs for zinc-finger broad tramtrack bric-a-brac (ZBTB) family of transcription factors (Michaloski *et al*., 2011) that are contained with i562. Removal of either one or both of the CNTCTGG motifs from a larger p2.5-eYFP (Söğünmez, 2012) or i562-p1.2-eYFP (Sancer, 2015) constructs resulted in recovery of expression. The *or101-1* gene is the last gene at the boundary of the combined OR115/OR101 cluster (Figure 4.7), suggesting that the presence of i562 may not be arbitrary. While the exact function of i562 is speculative at this point, it may have barrier function between OR regulatory elements of OR115/OR101 genes and non-olfactory genes

located adjacent to the OR cluster. While transgene expression suggests that i562 and the CNTCTGG motifs have a negative regulatory role, the binding of regulatory proteins to the sequence has not been demonstrated and specific transcriptional repressor proteins have not been identified. In this section, I will describe biochemical experiments aiming at consolidating the role of the recognized candidate CNTCTGG motifs in controlling *or101-1* gene expression by DNA-protein interaction assay based only electrophoretic mobility shifts (EMSA).

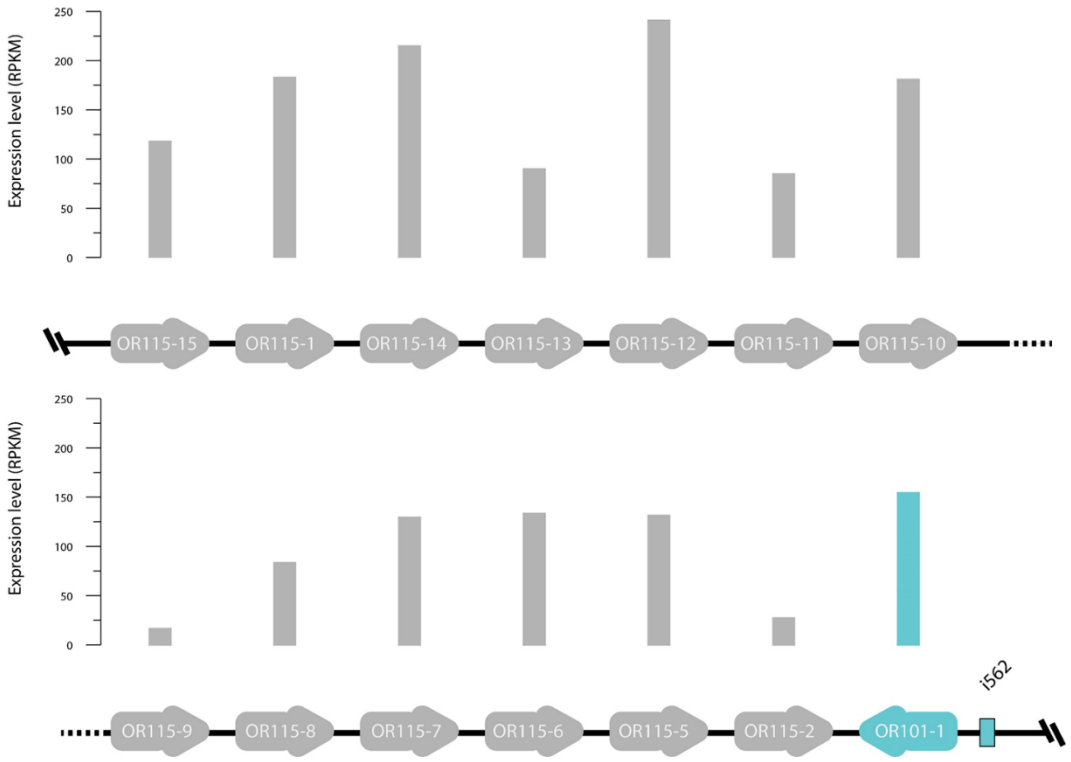

Figure 4.6. Schematic genomic organization of the OR115/OR101 gene cluster.

4.2.1. Expression of a p2.5-eYFP Transgenic Construct

In previous experiments it has been shown that adding the i562 sequence upstream of the highly expressed p1.2-eYFP construct dramatically reduced expression in injected embryos (Sancer, 2015). To confirm the inhibitory effect of the i562 sequence, a construct containing the uninterrupted 2.5 kb upstream of the *or101-1* translation start site (p2.5-

eYFP) was injected into fertilized zebrafish oocytes and scored for transgene expression. A total of 476 embryos were injected with p2.5-eYFP and pomp-mCherry constructs in six individual experiments (122, 60, 63, 56, 76, and 95 embryos, respectively). 208 of the injected embryos survived until 3 dpf and when screened 155 (45, 5, 14, 30, 26, and 35 embryos, respectively) of those were positive for pomp-mCherry expression, 98 of which (22, 3, 13, 23, 17 and 20 respectively) were also positive for p2.5-eYFP. Thus, the overall rate of double expression was 66.8 ± 6.4% (mean ± SEM). Further screening of double positive embryos (corresponding to 187 OEs) for cell count resulted in an average expression frequency of 49.7 ± 1.3%.

Thus, the presence of the i562 sequence reduced the expression by 10% from 60.3 to 49.7% at the cell count level. This difference reached statistical significance at a high level ($p << 0.0001$; Student's t-test, two-tailed, unpaired). However, the effect was not seen at the level of embryos and was highly different from previous reports in which a decrease from 61.3% to 12.7% (Söğünmez, 2012) or from 82.6% to 58.6% for i562-p1.2-eYFP (Sancer, 2015) were reported. The later study also observed a reduction in penetrance from 59% to 27%. The reason for the quantitative differences are not known but the results of this study, in principle, confirm the overall negative effect of i562 on transgene expression.

4.2.2. Biotinylation Efficiency of EMSA Probes Generated for i562

In order to test the recognition of candidate binding sites within i562 by OE-derived protein extracts, either pairs of complementary oligonucleotides spanning the conserved 5' and 3' CNTCTGG motifs over a length of 30 bp were annealed or larger regions spanning 100 bp or 200 bp over the candidate binding sites were amplified by PCR and gel-purified (Figure 4.8). In addition, for the longer 100 bp and 200 bp probes mutant versions (del100, del200) were generated in which the CNTCTGG motif was excluded from the sequence. Thus, a total of 10 different probes, 5 for each of the 5' and 3' motifs were generated. Each probe fragment was biotinylated at equimolar concentration using the Biotin 3’ End Labelling Kit (Pierce) and the biotinylation efficiency was determined (Figure 4.9). For this purpose, serial dilutions (0.1x, 0.01x and 0.001x) were prepared for each probe, blotted on

positively charged nylon membrane (Roche), developed by LightShift Chemiluminescent EMSA kit (Thermo Fisher) and imaged for chemiluminescence. As seen in Figure 4.9, the efficiency of biotinylation varied for each reaction and among probes. Overall, labeling efficiency was higher for the candidate 3'-binding site probes, with the exception of the labeling reaction for the 30 bp oligo-derived sequence, for which labeling was at least 10-times less efficient than for the 100 bp and 200 bp probes. Labeling of the 5'-probes was more uniform among the set but on average 2-times lower than for 3'-probes with the exception of the 200 bp wild-type probe, which was similar in labeling efficiency to 3'-probes. Nevertheless, all probes showed good biotinylation and could be used to proceed further with EMSA analysis.

4.2.3. Protein Extraction under Non-denaturing Conditions

Different protein sources were used as input in electrophoretic mobility shift assays (EMSAs), reflecting the different abundance of total protein or enrichment for suspected candidate binding partners. In general, because it was important to protect the three-dimensional structure of proteins for EMSA experiments, all protein extractions were performed with non-denaturing buffer containing protease inhibitor cocktail (Roche). Two different extraction protocols were used based on the protein source. As a rich source of ZBTB proteins the RLM11 cell line was used in addition to mouse and zebrafish tissues. For RLM11 cell line, mouse OE, brain and thymus tissues, a two-stage extraction protocol was applied, which separates cytoplasmic proteins from nuclear proteins. For the RLM11 cell line, one 100 mm plate containing 10% fetal bovine serum (FBS) in RPMI medium was inoculated and incubated at 37ºC and 5% carbon dioxide until it reached near-confluency and cell were collected and processed further for protein extraction. For protein extraction from mice, two animals of the C57BL/6 strain were sacrificed and their OE, brain, and thymus tissues were harvested for nuclear protein extraction. Since the zebrafish OE is relatively small, when compared to other sources (average diameter of 500 μm), 13 animals were sacrificed, the OEs were collected and a whole protein extraction protocol was applied. Protein concentrations were measured using Colibri Microvolume Spectrometer (Titertek Berthold). Concentrations were 5.19 mg/mL for RLM11 nuclear extract, 1.6 mg/mL for

mouse OE nuclear extract, 4.7 mg/mL for mouse brain nuclear extract, 3.76 mg/mL for mouse thymus nuclear extract and 2 mg/mL for the zebrafish whole OE extract.

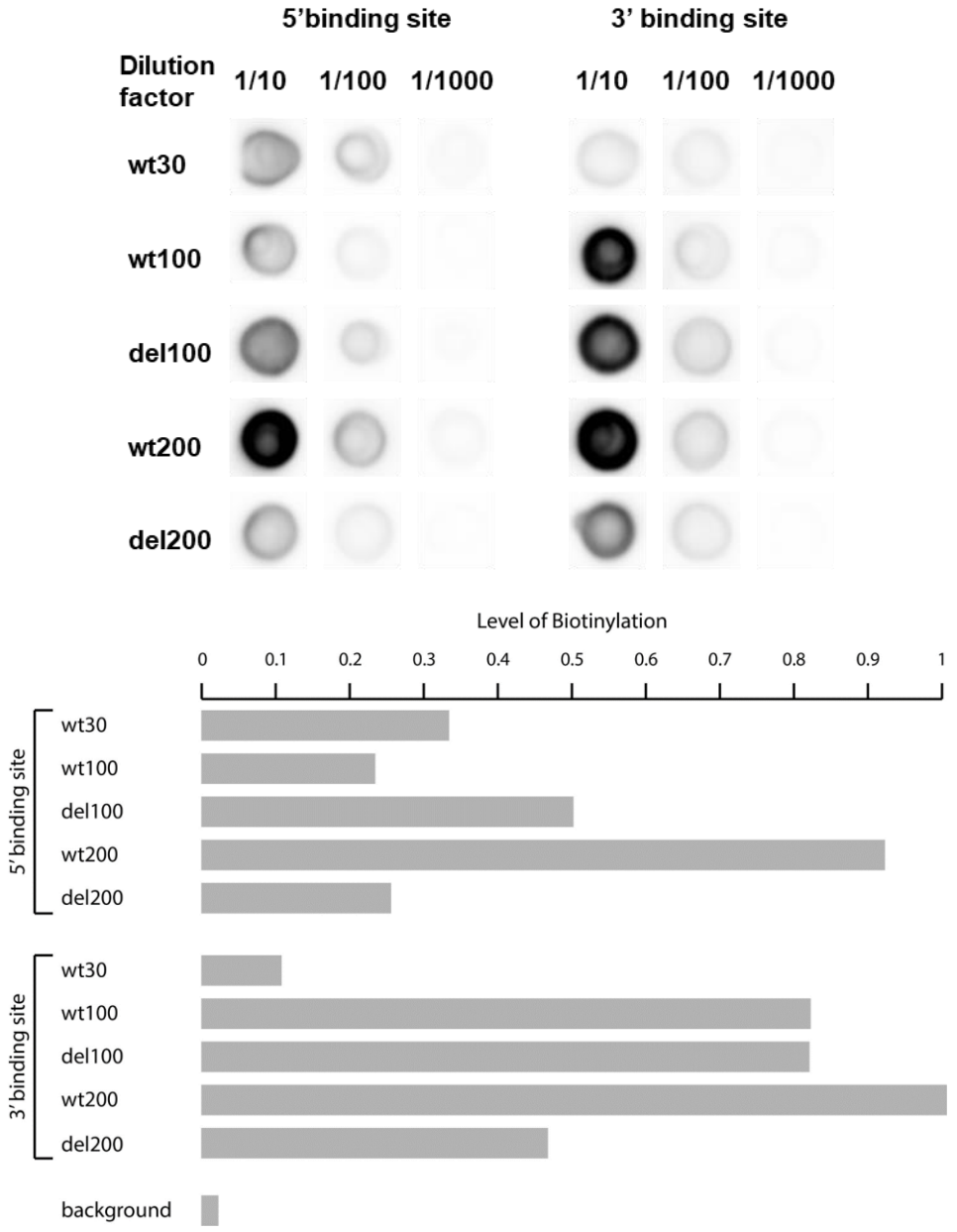

Figure 4.7. Biotinylation efficiency of EMSA probes. Raw chemiluminescence (top) of serial dilutions (1x, 0.1x, 0.01x) for 5' (left) and 3' (right) probe sets and quantification of biotinylation efficiency (bottom).

4.2.4. Western Blot Analysis to Verify the Expression of Zbtb7b in Various Tissues

Michaloski *et al*. (2011) proposed that the CNTCTGG motif is overrepresented upstream sequences of OR and vomeronasal receptor genes and that it may be a binding motif for the BTB-ZF protein Zbtb7b (also known as cKrox, thPOK and ZFP67). Thus, a hypothesis that was tested in this study is that i562 might exert its inhibitory function through Zbtb7b binding to the two copies of the motif, which are present in i562 and which were shown to be functional in transgenic assays.

Inspection of RNAseq expression data revealed that the zebrafish paralogues of *zbtb7b* is expressed in the zebrafish OE and brain. However, a Zbtb7b-specific antibody raised against the mouse protein did not label OSNs in the zebrafish OE or in Western blots of zebrafish OE-derived protein extract (data not shown), most likely due to lack of cross-reactivity of the antibody with the zebrafish protein. While EMSA assays would show the binding of a protein to a DNA sequence, it would not provide any information as to which specific protein interacts with the sequence. Yet, supershift of EMSA bands with protein-specific antibodies could provide this answer. For this reason, and because the zebrafish OE is rather small and the amount of protein harvested from zebrafish OE is limited, alternative sources for Zbtb7b, such as mouse OE and brain tissue and mouse thymus-derived RLM11 cell line were investigated by Western blotting. RLM11, a $CD4^+/CD8^-$ cancer cell line (Landry *et al*., 1993; Ehlers *et al*., 2003), is expected to be rich source for the Zbtb7b protein, because it is the determinant factor in maturation of double-positive $CD4^+/CD8^+$ precursor T cell to single-positive $CD4^+/CD8^-$ T cells (Kappes, *et al*., 2006; Egawa and Littman, 2008).

To examine whether protein extracts contained Zbtb7b protein, Western blot analysis was performed on RLM11 mouse OE, brain and thymus nuclear extracts (Figure 4.8). The expected molecular weight for the longest splice variant of the mouse Zbtb7b protein is 58.9 kDa and the remaining splice variants have a molecular weight either equal to or lower than 12.8 kDa (Zerbino *et al*., 2017). RLM11 nuclear extracts showed a strong positive band between 75 and 50 kDa bands, while labeled bands from mouse OE, brain and thymus nuclear extracts appeared to have a lower molecular weight than 50 kDa, suggesting that

RLM11 and OE/brain express different isoforms of the protein (possibly unregistered to the database). Even though the antibody did not appear to be specific, as additional bands could be detected on the Western blots, the pattern for recognitions over three different protein sources were consistent with each other. Thus, mouse or RLM11-derived protein extracts may be used in supershift assays to gain insight into the interacting proteins, albeit with different isoforms of the protein.

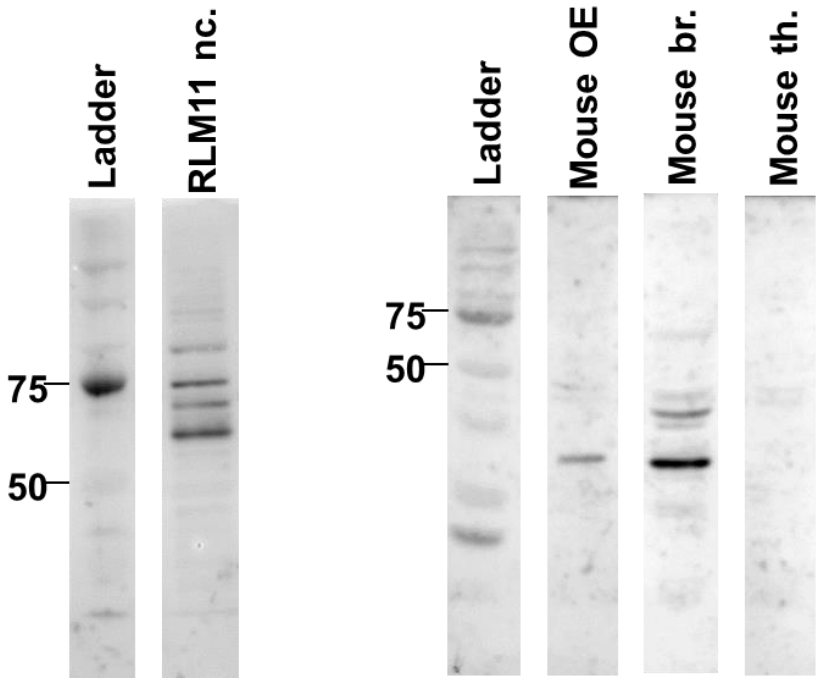

Figure 4.8. Western blot analysis for Zbtb7b protein.

4.2.5. DNA-Protein Binding of the 5' Binding Site with RLM11 Nuclear Extract

Motif scans on i562 revealed the presence of two candidate ZBTB-binding sites that were located 5' (5bs) and 3' (3bs) relative to their position within i562. For each binding site five probes were prepared: a 30 bp probe that covers the 7 bp motif (wt30) and two probe sets of around 100 and 200 bp (one, which contains the candidate wild-type motif (wt100, wt200) and a probe in which the CNTCTGG motif was deleted (del100, del200). As a first test, to see whether RLM11-derived protein extracts would be able to shift the probe bands,

the same amount of RLM11 nuclear extract was incubated with the five different 5bs probes (Figure 4.9). For all probes, a clear shift of the bands could be observed, indicating that the probe sequences were recognized by protein(s) contained within the RLM11 extract. Possibly due to the difference in size and complexity of the various probes, different banding patterns were observed for each probe. Five individual bands could be seen for the wt30 probe, while two bands and possibly a group of bands that are too close to each other to give separable signals were detected for wt100. For the wt200 probe, three clear bands could be detected, even though the exact number of bands cannot be determined due to smearing of the bands at higher molecular weight. With the exception of the wt30 probe, free unbound probes disappeared completely for all reactions that were incubated with RLM11 protein extracts, indicating that the concentration of interacting protein is sufficiently high in RLM11 extracts. However, a band of remaining free probes could be detected for wt30. This result may indicate that the probe occupied all of the available proteins within the reaction and that wt30 probe was in excess. This difference may again be due to the difference in the complexity of the probes, as the longer 100 and 200 bp probes may contain additional sequences that can act as targets for additional DNA-binding proteins from the RLM11 extract.

Importantly, rather similar patterns and efficiency of band shifts could be observed for wild type probes and probes that contained a deletion of the investigated binding motif. However, it appears that the lower molecular weight bands observed for wt100 and wt200 were not present in del100 and del200 lanes. Thus, proteins present in complex RLM11 nuclear extract recognize the i562-derived probes spanning the 5bs and the absence of bands/low molecular weight smear in del100 and del200 suggests that the 5' CNTCTGG motif may indeed be recognized by specific proteins.

4.2.6. DNA-Protein Binding of the 3’ Binding Site with RLM11 Nuclear Extract

A similar EMSA was performed for 3bs using RLM11 nuclear extract as protein source (Figure 4.10). As seen for the 5' probe set, all five 3' probes were efficiently shifted to higher molecular weights upon incubation with protein extract. However, even though

protein(s) present within the extract recognized the probes, it was not possible to discriminate individual bands in this assay. The gel of this EMSA was overrun and the free probe for wt30 ran off the gel, while the available probes were all occupied by proteins and

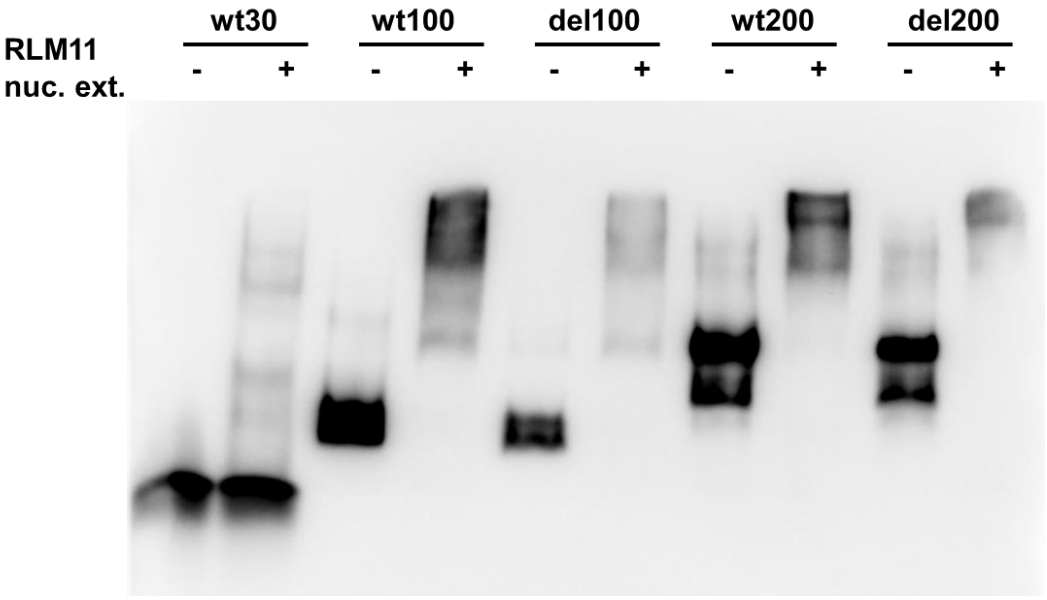

Figure 4.9. EMSA for 5' binding site with RLM11 protein extracts.

shifted completely in remaining reactions. The pattern of band shifts between wt100 and del100 reactions were rather similar and were not indicative of DNA-protein interaction that were specific to the presence of the CNTCTGG motif. Yet, due to the observed smear, it may be difficult to detect the presence/absence of specific interactions. A loss of lower molecular smear was detectable between wt200 and del200 probes. However, the labeling efficiency for del200 was lower than for wt200 and the absence of low molecular weight smear may be due to difference in the concentration of effectively labeled probe molecules. To express the differences in probe binding, a signal intensity analysis was performed on the image using FIJI image processing tools. Intensities of shifts and free probes of the control reactions were measured, background signal was extracted from the measured intensity values and free probe values were used to normalize the shift signal intensities. The intensity of wt100 calculated to be 62.8% that of del100 while intensity of del200 was calculated to be 24.3% that of wt200. These differences observed are possible caused by differential transfer efficiency over the membrane.

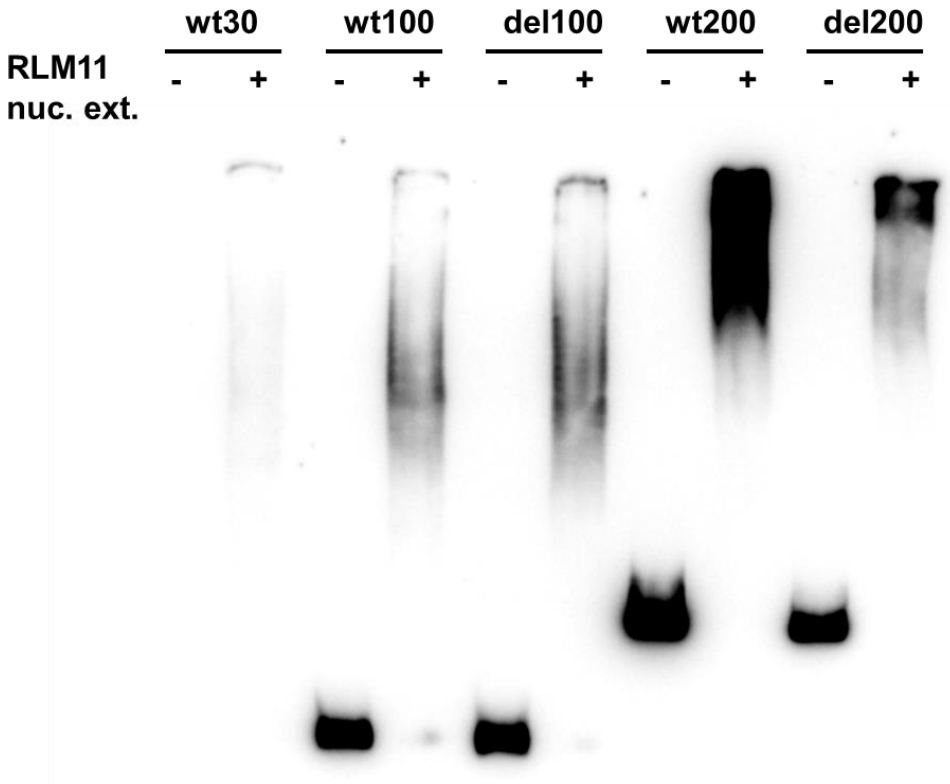

Figure 4.10. EMSA for 3' binding site with RLM11 protein extracts.

4.2.7. DNA-Protein Binding for 5’ and 3’ Binding Sites with Nuclear Extracts from Mouse Olfactory Epithelium, Brain and Thymus Tissues

As a next step, nuclear protein extracts from OE, brain, and thymus tissues from the mouse were incubated with the wt30 probes for either 5bs or 3bs and resolved by non-denaturing PAGE (Figure 4.11). Binding of protein(s) occurred for all probe-nuclear extract combinations investigated (OE, brain, thymus for 5bs and OE and brain for 3bs). Two separate bands for 5bs probe with mouse OE nuclear extract, a minimum of three bands for 5bs with brain nuclear extract, and two separate bands for 5bs with thymus nuclear extract could be observed. Differences in the pattern of OE and thymus nuclear extracts binding with the 5bs were observed for high molecular weight band position. The binding pattern of 3bs probe was fairly similar to the 5bs probe for the same protein extract, yet, there were two significant differences. First, an extra band was visible in the 3bs/brain nuclear extract sample and second, the intensity of these bands in total appears to be higher than for the 5bs

counterpart. In relation to previous EMSA experiment with RLM11 nuclear extract, band patterns for shifts are different from the combination of 5bs wt30 and RLM11 nuclear extract, suggesting that the different protein sources contain different proteins that may interact with the probe.

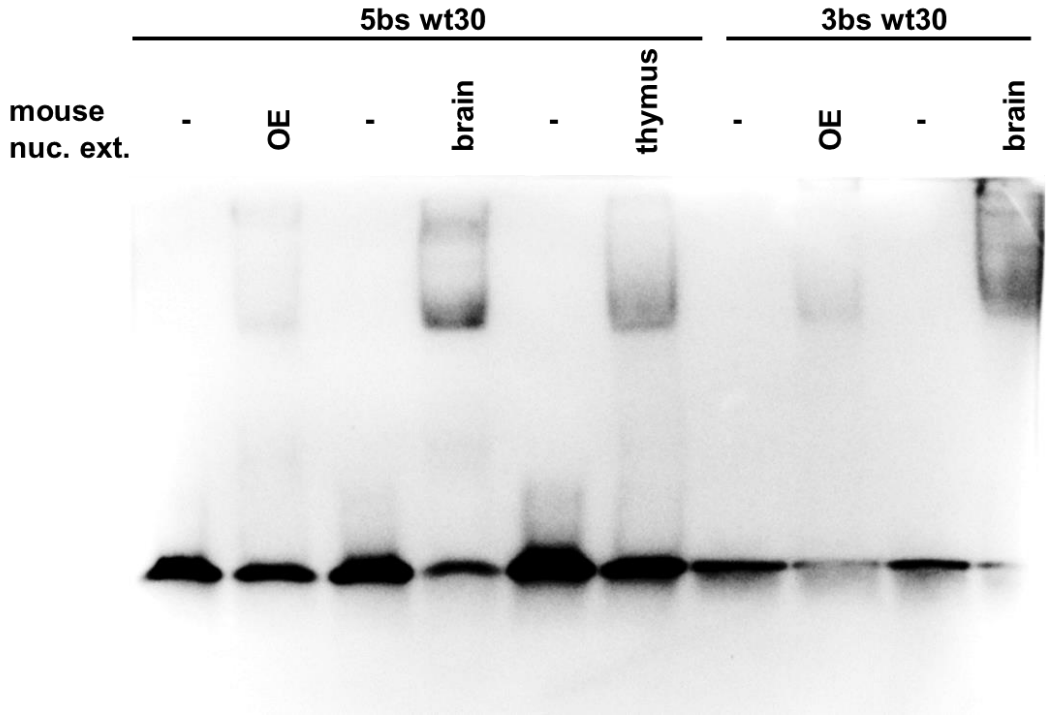

Figure 4.11. EMSA for 5' and 3' binding site with protein extracts from mouse tissues.

In sum, the tested nuclear protein extracts from three different mouse tissues that are suspected to express Zbtb7b and/or related BTB-ZF proteins include protein(s) capable of recognizing the probes containing the CNTCTGG motifs of i562. The pattern differences between nuclear extracts from different tissues suggest that proteins involved in probe recognition are of various molecular weight, consistent with the differences observed in Western blot analysis with an anti-zbtb7b antibody on brain and thymus extracts (Figure 4.10). Thus, either different proteins, or different isoforms of the same protein are able to bind the probes. Alternatively, the primary DNA-binding protein may be capable of forming different higher molecular weight complexes with additional protein that are differently expressed in brain, thymus and OE.

4.2.8. DNA-Protein Binding in the Presence Of Zbtb7b Antibody – Supershift Assay

The experiments detailed so far suggest that proteins bind to sequences flanking the 5bs and 3bs of i562 and that specific interactions may occur directly at the respective CNTCTGG sequences. A suspected target protein that has been implicated in regulating OR and vomeronasal gene expression is Zbtb7b (Michaloski *et al*., 2011). Thus, incubation of the DNA-protein binding reactions with a specific antibody against Zbtb7b is expected to shift bands derived from Zbtb7b-probe interaction to a higher molecular weight, at least for the mouse protein extracts in which specific bands could be detected by Western blot analysis.

To test this hypothesis, EMSA experiment with 5bs and 3bs and mouse OE or brain nuclear extracts were performed in the absence and presence of an antibody against Zbtb7b (Figure 4.12). Two different antibody incubation conditions were used in which either the protein and the antibody were incubated on ice together before addition of the probe, or probe and protein extract were incubated together at room temperature before the antibody was added to the reaction and incubated on ice. Interestingly, intensity differences were observed for shifted probe when OE and brain nuclear extracts were used. This suggests that proteins capable of binding to the wt30 probe were more abundant in brain extracts. However, for reactions in which the 5bs was used, no obvious or specific differences could be detected between reactions that did or did not contain the anti-Zbtb7b antibody. This suggests, that the observed band shifts are not due to interaction of Zbtb7b with the 5bs or the portion of shift that is due to Zbtb7b is low and is overshadowed by other binding events. In addition, no distinct pattern difference between pre and post antibody incubations were seen, suggesting that the order in which the antibody is added to the reaction does not alter binding. It should also be noted, however, that for the 3bs additional complexes were retained within or close to the loading wells. However, the observed band does not allow the conclusion that a specific supershift has occurred.

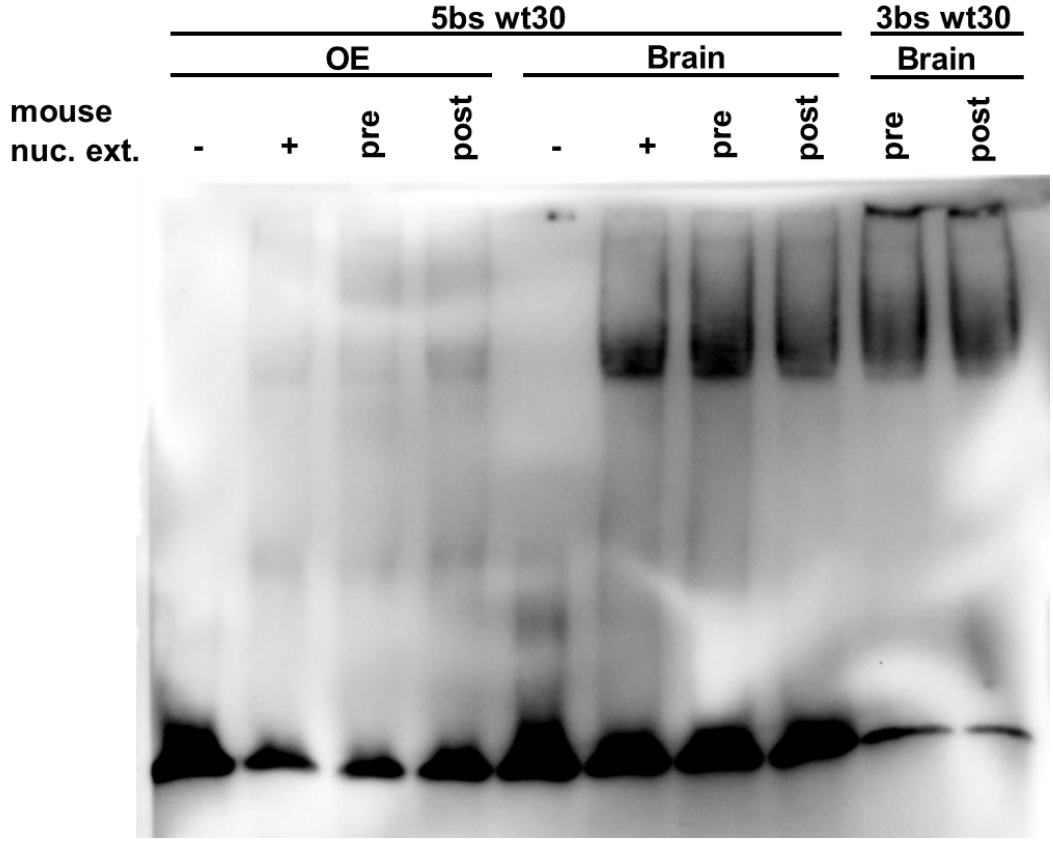

Figure 4.12. Supershift assay for 5' and 3' binding site with OE and brain protein extracts.

4.2.9. DNA-Protein Binding with Mouse Brain Extract in the Presence of Zbtb7b Antibody – Supershift Assay

As a partial repeat of previous supershift assay and to better observe the binding dynamics of 3bs in the presence of Zbtb7b antibody, another EMSA experiment was performed (Figure 4.13). The mouse brain nuclear extract was used because it resulted in stronger shifts in previous assays (Figure 4.12). Incubation of the probes with protein extract was performed at room temperature followed by antibody incubation on ice. Each reaction showed strong band shifts, while addition of the antibody resulted in an overall decrease in shift intensity. However, a distinct supershifted band was not visible, either because it was too large to enter the gel during electrophoresis, or because it has a low signal intensity and remained invisible within the smeared signal. To see whether the difference in intensity for bands incubated with antibody was real, the intensities were analyzed quantitatively with

FIJI imaging software. Since each binding reaction contained same amount of protein and probe, it was assumed that the ratio of free probes might be used as a correction factor in case where there are intensity differences between the lanes. To correct for uneven background across the blot intensity measurement were recorded for each measured region and adjacent positions. Another problem that was present in the membrane was the lack of free probes for 5bs wt100 probe. But, calculation showed that there were close to minimal intensity difference between the lanes and hence each shift intensity could directly be used for comparison. Overall, the intensity of shifts for the wt30 probe for 5bs was 27.7% in the presence of the antibody when compared to control conditions. Similarly, the intensity of shifts for the wt30 for 3bs was 58.9% in the presence of the antibody compared to control lanes. A similar observation was made for the longer 5' bs (wt100) for which the intensity decreased to 43.8% in the presence of the antibody. Thus, even though not discernable as a distinct band, addition of antibody consistently reduced the intensity of shifted bands, suggesting that the mobility of the hybrids may have changed.

4.2.10. DNA-Protein Binding of 5bs by Mouse OE Nuclear Extracts

Binding activity of mouse OE, brain and thymus nuclear protein extracts were tested before for wt30, while their DNA binding behavior for longer probes were not. Therefore, the various 5bs probes were incubated with mouse OE nuclear extract. Protein(s) within the extract bound all five probes, yet differences in pattern could be observed for the different probes. Interestingly, a clear difference between binding patterns of the wt100 and del100 probes could be observed, yet the difference was largely quantitative as no distinct bands could be recognized. Therefore, it is difficult to conclude whether the observed difference is due to the absence of specific interaction of nuclear proteins with del100 or if the difference is due to slight signal intensity difference between samples. However, a similar difference was observed between wt200 and del200 in which smear of intermediate molecular weight disappeared.

To obtain a better estimate of the strength of interaction and to compensate for differences in loading and labeling efficiencies, a quantitative analysis of band densities was

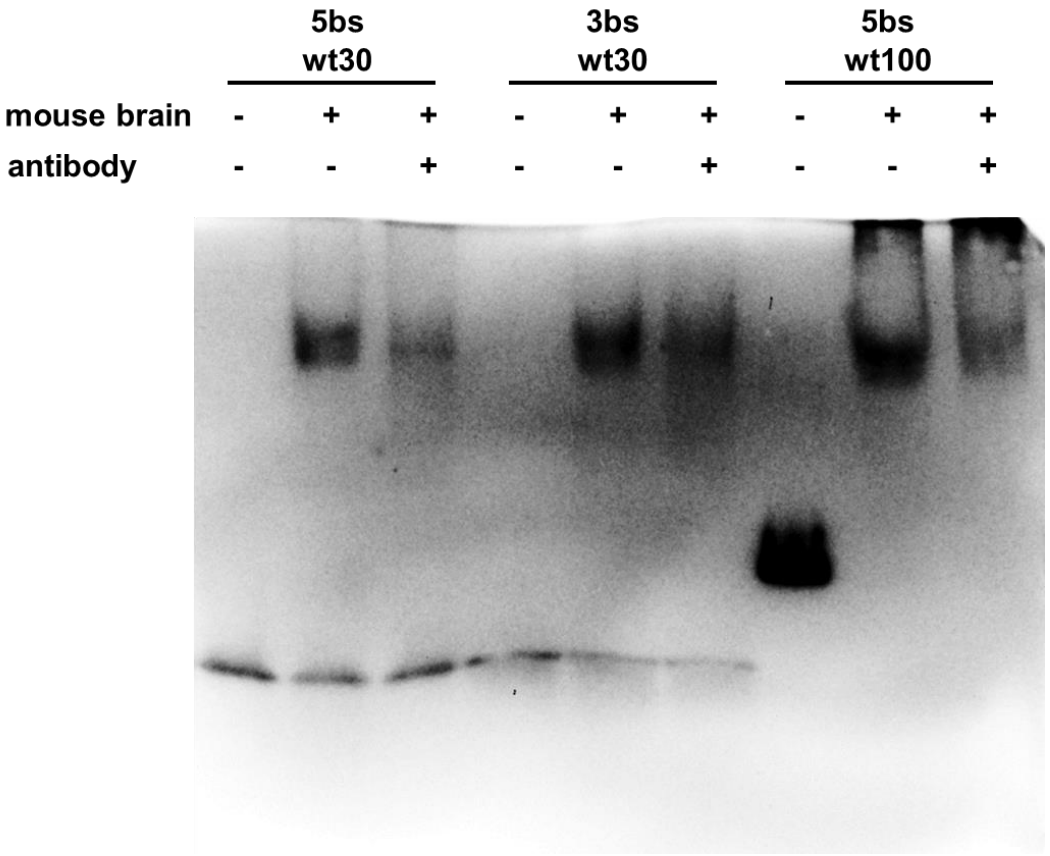

Figure 4.13. Supershift assay for 5' and 3' binding sites with brain protein extracts.

performed. To do so, integrated intensity of suspected molecular weight bands and free probes of negative control were measured, while the background intensity was extracted for each value. The ratio of free probes was used as correction factor and integrated intensities of suspected molecular weight bands were compared. The analysis for wt100 and del100 showed no significant difference suggesting the observed difference was due to the labeling efficiencies of probes, while the signal intensity of del200 was only 54.2% of that of wt200.

Thus, the probes are recognized by the protein or proteins present in mouse OE nuclear extract and the candidate site might be recognized by a protein that is present both in RLM11 and mouse OE nuclear extracts, which strengthens the idea that 5bs candidate site has a role in inhibitory function of i562. In principle, indication of the relevance of the 5bs and 3bs for this interaction could be obtained since specific molecular weight interactions were absent in del200 probe that did not contain the motif.

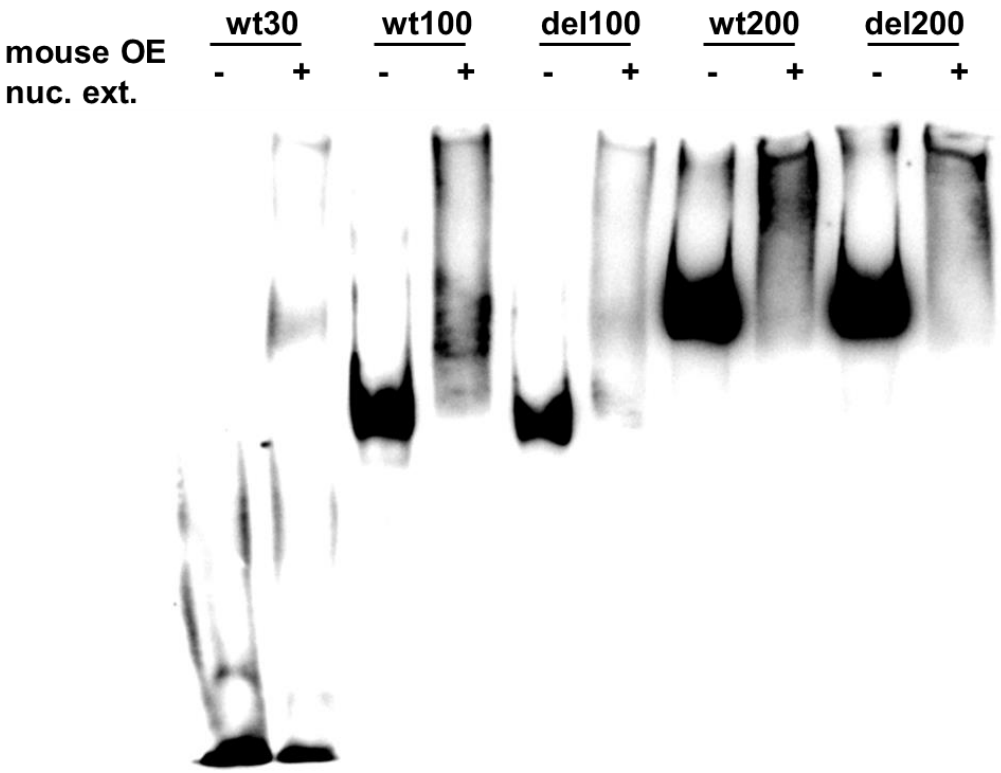

Figure 4.14. EMSA analysis of 5' probes with mouse OE protein extracts.

4.2.11. DNA-Protein Binding for 5bs and 3bs with Zebrafish Whole OE Protein Extract

Ultimately, it was necessary to show that binding also occurs between the various DNA probes and protein extracts from zebrafish OE because the i562 is active in this tissue. Therefore, wt100 and wt200 probes for the 5bs and 3bs were incubated with zebrafish whole OE protein extract (Figure 5.15). Protein(s) present in the zebrafish OE extract recognized all probes and clear band shifts could be observed. However, it should be noted that a random sequence oligo probe that was included as a negative control showed a similar gel shift pattern. Control reaction for 3bs wt100 probe lacks the probe, hence the empty lane and mistakenly wt100 probe was included in control reaction of 3bs wt200 hence the extra bands. In conclusion, protein extracted from the zebrafish OE may also interact with i562 as

expected. However, the shifts observed for the negative control oligo cast some uncertainty onto this finding.

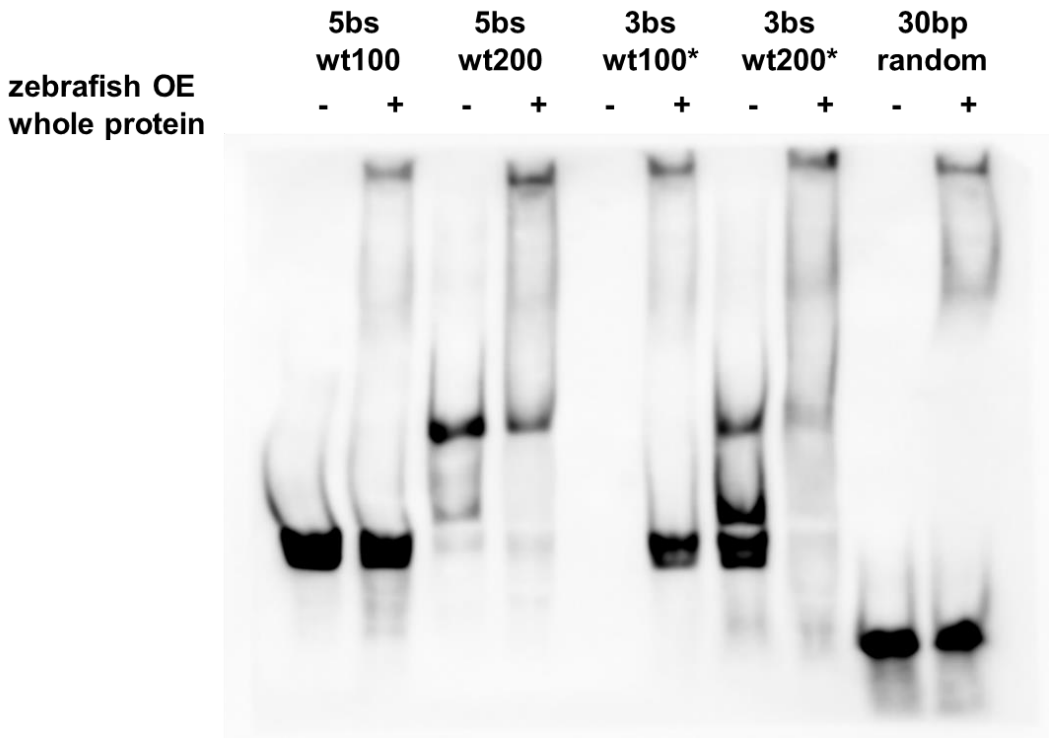

Figure 4.15. EMSA analysis of 5' and 3' probes with zebrafish nose whole protein extract.

4.2.12. Motif Scan on Randomly Generated Oligo Control Probe

As shown above (Figure 5.15), a random oligonucleotide probe was generated but is subject to similar band shifts as probes that contain CNTCTGG motifs. To gain more insight into the nature of this possible interaction, a motif scan on the random oligo sequence was performed on JASPAR (Storno, 2013), which revealed multiple binding sites present, some of which sharing more than 90% similarity with the motif detected. Hits obtained with motif scan were crosschecked in available transcriptome data for their expression; while very few showed no significant expression in zebrafish OE, some others showed significant levels of expression (lower but similar to *ompb* expression level) and most showed medium expression level (Table 4.2).

4.2.13. DNA-Protein Binding for 5bs and 3bs with Zebrafish OE Protein Extract in the Presence of Competitor

After binding between zebrafish OE-derived protein and large probes (wt100 and wt200) of each candidate site was shown, a competition assay was performed to assess the role of candidate binding sites in recognition of probes (Figure 4.16). To do so, wt100 and wt200 probes of 5bs and 3bs were incubated with 200-fold excess of a 30 bp unbiotinylated oligo. Unlike previous EMSAs with non-zebrafish protein sources, binding reactions with protein extract contained free probes indicating the full saturation of available DNA binding proteins by the probes. For each probe, clear band shifts could be detected, yet the overall binding ratio was somewhat less than previous cases. Since the protein source used here was whole protein rather than the nuclear fraction, the ratio of effective proteins was expected to be lower in comparison to nuclear extracts. It is worth mentioning here that increasing the amount of protein per reaction would also mean increasing the salt concentration, which might negatively alter the binding ability of proteins. Presence of protein also triggered accumulation in the wells blocking the entrance of probe into the gel, which might mean that binding of proteins onto probes eventually leads to formation of complexes and this is expected since predicted proteins (either Zbtb7b or some other BTB ZF family protein) are able to dimerize and/or interact with other proteins that may participate in epigenetic silencing. Nonetheless, addition of competitor to either 5bs or 3bs wt200 reactions did not significantly alter the binding efficiency. No obvious differences could be observed between lanes that did and lanes that did not contain the competitor. It might be that the level of competitors used was too low for effective competition or that other proteins were present within the extract that were able to recognize regions outside the competed portions of the probe with a better ratio.

4.2.14. Effect of Competitor Concentration

In the previous competition assay, the 30 bp competitor oligonucleotides did not show any significant change in the shift for wt200 probes. One possible reason might be that necessary competitor concentration was not reached to see any effect. To examine the effect

Table 4.2. Expression of transcription factors in zebrafish OE that might have a binding motif on random oligo probe.

		RPKM
ENSDARG00000009094	gata2b	340
ENSDARG00000042796	yy1a	2166
ENSDARG00000055398	foxc1b	203
ENSDARG00000077467	sox10	113
ENSDARG00000052383	tcf3b	2259
ENSDARG00000026053	hltf	704
ENSDARG00000056130	neurog1	70
ENSDARG00000016854	neurod2	1
ENSDARG00000012078	meis1	1624
ENSDARG00000035735	gsx1	0
ENSDARG00000061836	nfixb	1837
ENSDARG00000053586	creb1a	535
ENSDARG00000055398	foxc1b	203
ENSDARG00000058999	cdx1b	12
ENSDARG00000001784	hoxa5a	0
ENSDARG00000036074	cebpa	523
ENSDARG00000043322	gsx2	0
ENSDARG00000043210	nfic	1248
ENSDARG00000030402	twist1a	168
ENSDARG00000068976	bsx	0
ENSDARG00000077120	arid5a	224
ENSDARG00000013125	dlx1a	0
ENSDARG00000055294	atoh1a	8
ENSDARG00000042291	dlx6a	400
ENSDARG00000030703	otx1a	238
ENSDARG00000011235	otx2	24
ENSDARG00000000175	hoxb2a	0
ENSDARG00000005150	tbx20	3
ENSDARG00000088347	sp1	1197
ENSDARG00000075713	shox2	4
ENSDARG00000044075	nkx6.2	0
ENSDARG00000071560	dlx4b	419
ENSDARG00000023527	smad4	683
ENSDARG00000032380	ompb	5515

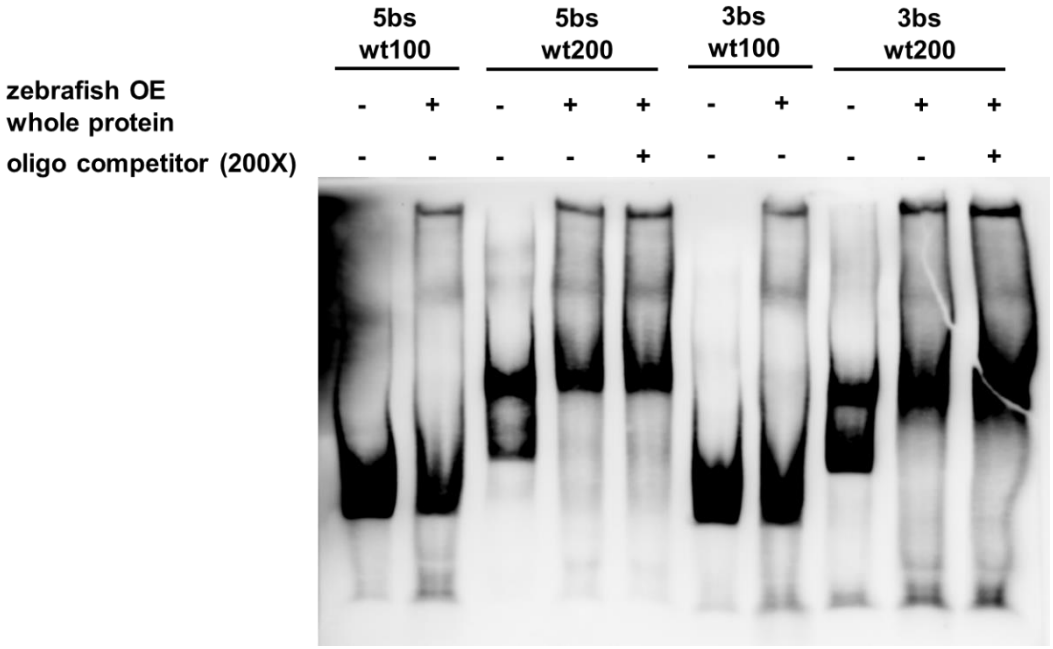

Figure 4.16. EMSA analysis of 5’ and 3’ probes in the presence of competitor with zebrafish nose whole protein extract.

of competitor concentration assays were performed with increasing concentrations of competitor (Figure 5.17). For both 5' and 3' wt200 probes 250x, 500x, and 750x concentrations of the respective competitor were tested. Reaction mixture were prepared with all components included and incubated at room temperature. Addition of the competitor, irrespective of the concentration, altered the mobility of free probes within the gel. However, irrespective of the competitor concentration, the overall gel shift pattern did not change. If anything, addition of the competitor increased binding of protein to the probe.

Since uneven background of the membrane obscured analysis for the 3bs, the experiment was repeated with exact same conditions (Figure 5.18). In this experiment, the amount of free probe increased slightly with increasing competitor concentration and a clear three banded shift pattern was observed. For the 750x excess of competitor a reduction in intensity was also noticeable. However, this may have originated from inefficient transfer because the intensity of the free probe was also reduced.

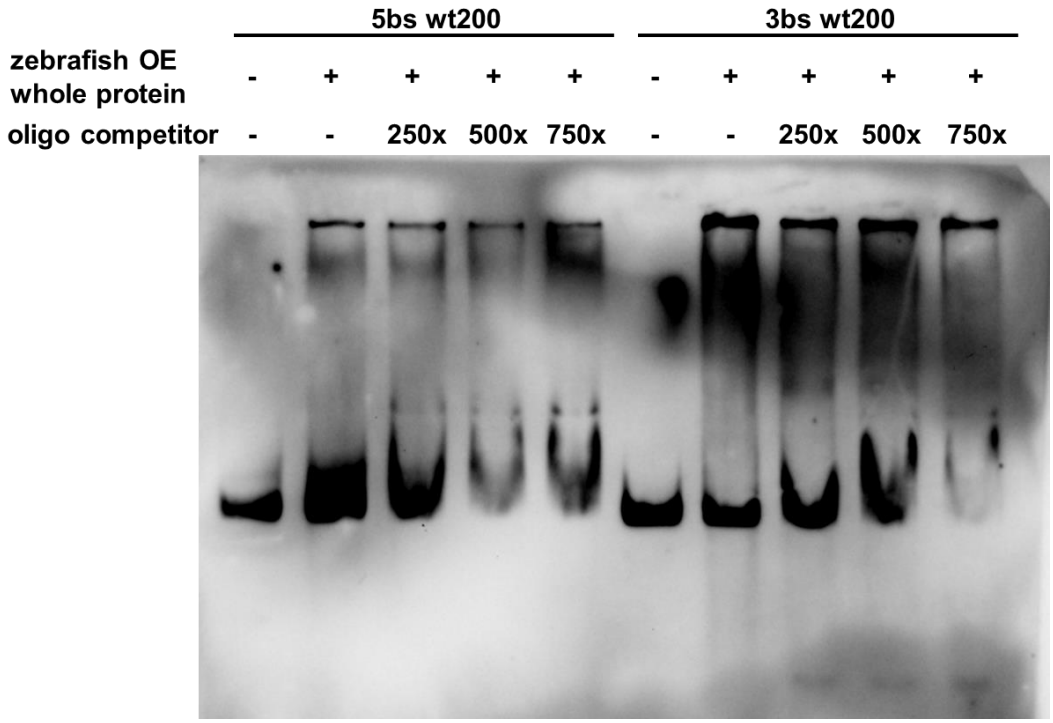

Figure 4.17. EMSA analysis of 5’ and 3’ probes with different competitor concentrations and with zebrafish nose whole protein extract.

4.2.15. Final Summary and Conclusion.

Overall, probes of each candidate binding site were recognized by proteins of RLM11 cell line, various mouse tissue nuclear extracts and zebrafish nose whole protein extract. Yet, the deletion of candidate sites from the probes or competition of large probes with core 30 bp oligonucleotides resulted in no significant shift change suggesting the previously observed inhibitory effect of i562 does not directly originate from these candidate motifs and their interactions with proteins. Also, experiments for supershift with anti-Zbtb7b antibody did not provided reliable support for the hypothesis that i562 performs its inhibitory function by providing Zbtb7b docking site(s).

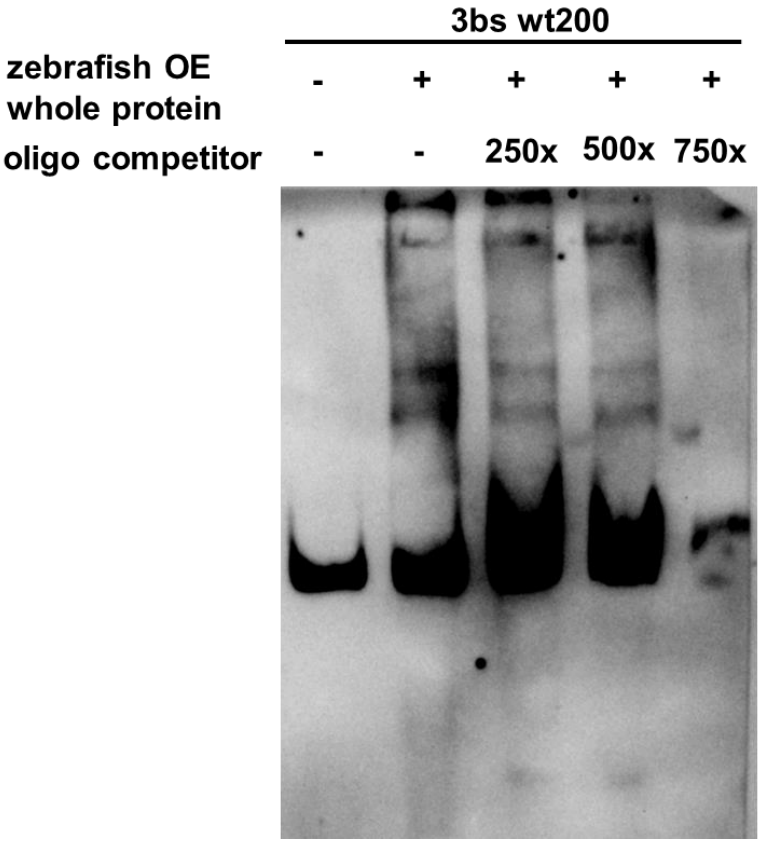

Figure 4.18. EMSA analysis 3’ probes with different competitor concentrations and with zebrafish nose whole protein extract.

5. DISCUSSION

One of the key characteristics of OSNs is that they express only a single OR gene from a large genomic repertoire (Malnic *et al*., 1999); a feature that has been termed the 'one neuron – one receptor rule' (Mombaerts, 2004). In addition, OR gene expression is also monoallelic (Chess *et al*., 1994; Ishii *et al*., 2001), which, in combination with monogenic expression, is commonly referred to as expression 'singularity' (Vassalli *et al*., 2002). Reaching expression singularity is a critical step during OSN maturation and fundamental to the proper function of OSN. OSNs that accidentally select more than one OR gene are typically eliminated from the tissue (Tian and Ma, 2008; Abdus-Saboor *et al*., 2016). In case a pseudogene is selected, pre-mature OSNs either reenter the OR selection process or undergo apoptosis (Lewcock and Reed, 2004; Serizawa *et al*., 2003; Shykind *et al*., 2004; Fuss *et al*., 2013). The OR genes are typically arranged in clusters in the genome (Mombaerts, 2004) that are coordinately controlled for OR selection and expression (Markenskoff-Papadimitriou *et al*, 2014; Iwata *et al*., 2017). Thus, OR gene choice and expression are tightly controlled processes that include multiple layers of regulation, which act at the level of individual OR gene loci, larger clusters of OR genes, and the physiology and cell biology of OSNs as a whole (Fuss and Ray, 2009).

In this study, I was interested in two distinct aspects of OR gene regulation imposed by genomic sequence flanking the zebrafish *or101-1* gene: positive regulation of OR gene choice from proximal promoter elements and negative modulation of expression by a largely uncharacterized distal element described here as i562. Positive regulators have been shown and are expected to exist in proximal OR gene promoters (Rothman *et al*., 2005; Michaloski *et al*., 2006; Michaloski *et al*., 2011; Clowney *et al*., 2012; Plessy *et al*., 2012) to attract the intricate OR gene choice machinery to individual OR gene loci (Dalton and Lomvardas, 2015). Thus, proximal promoter elements are mechanistically required for expression of a given OR gene. Two candidate binding sites for Olf1/Ebf1 transcription factors, one located upstream of the TSS and a second one located in an intron were tested for redundant function using transient expression of reporter constructs in zebrafish embryos. The intronic site appears to have a more prominent role, however, neither of the two sites was strictly required

for expression. Interestingly, the splicing of an intronic sequence located in the 5'-untranslated sequence of the *or101-1* gene appears to be essential for proper expression of the gene. The *or101-1* gene is part of a larger OR gene cluster on zebrafish chromosome 21 and *or101-1* is located at the far end of the cluster in an outward-oriented fashion. Previous studies have shown that a sequence upstream of the gene, thus located at the boundary of the cluster negatively affects transgene expression. The *or101-1* gene promoter, or the larger cluster may be sensitive to interference with outside regulators and the i562 sequence may either be required to tune *or101-1* expression or it may define a boundary element to isolate the OR cluster from flanking sequences (Ali *et al*., 2016). Two candidate binding sites for members of the zinc finger broad tramtrak bric-a-brac (ZBTB) family of transcription factors have previously been shown to contribute to the overall repressive activity of i562. Here, I could demonstrate that i562 interact with proteins expressed in the OE, while specific binding to the suspected sites could not be shown conclusively.

5.1. The Molecular Mechanisms of OR Gene Choice

The principle of relaying external sensory input in a topographically-mapped manner is a common scheme in sensory systems and is found in the visual, auditory, and even in tactile sense (Patel *et al*., 2014; Talavage *et al*., 2000; Killakey *et al*., 1995;). A related phenomenon of topographical mapping is found in the olfactory system through the projection of axons of OSNs that express the same OR to a specific glomerulus in the olfactory bulb (Mombaerts *et al*., 1996; Wang *et al*., 1998). A prerequisite for this anatomical organization is the singularity of OR expression seen in OSNs. The observation that a single OR gene is selected from a vast number of possible OR genes (1400 ORs in mouse, more than 400 ORs in human and 175 ORs in zebrafish), poses the question of how expression singularity is achieved.

Sequences flanking OR genes in the main OE of the mouse are enriched in repressive epigenetic marks of H3K9me3 and H3K20me3, hence most genes in an OSN are transcriptionally silenced (Magklara *et al*., 2011). It has been shown that transient expression of the lysine-specific demethylase LSD1 desilences OR genes randomly by erasing di- and

tri-methylations from histones (Lyons *et al*., 2013). Removal of the silencing marks followed by induction of H3K4me3 modifications results in OR activation. Expression of a functional OR protein then triggers the expression of adenylyl cyclase type III, *Adcy3*, which normally converts sensory input coming from GPCRs (in this case ORs) into the secondary messenger molecule cAMP. Interestingly *Adcy3* expression also results in downregulation of LSD1 expression and thereby stops the OR gene choice process. Expression of an OR pseudogene somehow does not result in expression of *Adcy3* and in the absence of *Adcy3* activity, LSD1 continues to activate OR genes until a functional gene is released from epigenetic silencing. This interplay between OR gene translation and LSD1 bridged by *Adcy3* establishes the singularity of OR expression that is required for the proper function of the olfactory system.

The transient expression of LSD1 is necessary for OR selection but is not required for the maintenance of activated OR throughout the life of OSN. Once activated, a feedback loop based on Adcy3 ensures the singular expression by simply keeping the rest of the OR genes at their epigenetically silenced state (Dalton *et al*., 2013). However, important questions as to how OR expression is achieved or continued throughout the life of OSN that have not conceptually been integrated with the above described OR gene choice mechanism remain. For instance, OR gene choice is not uniform across the repertoire and some OR genes are more frequently expressed than others (Ressler *et al*., 1993). In the mouse, OR genes are also expressed in specific zones of the OE (Miyamichi *et al*., 2005) suggesting position-specific bias of OR gene choice. OR genes are distributed over ~50 clusters in the mouse genome and it remains unanswered how OR choice is coordinated across clusters. Evidence is emerging that each functional OR gene clusters are controlled by enhancer elements that are similar to locus control regions (Serizawa *et al*., 2003; Fuss *et al*., 2007; Khan *et al*., 2011, Iwata *et al*., 2017). Enhancers from different clusters and chromosomes appear come together in the nucleus to form expression hubs for OR genes (Marcenskoff-Papadimitriou, 2014). Thus, it may occur in these enhanceosome hubs that bias towards expression of specific OR genes is generated.

5.2. The Role of Sequence-specific Transcription Factors in OR Expression

Expression studies in the mouse OE have pointed towards homeodomain and O/E binding sites within the proximal promoters of OR genes in relation to expression and frequency of selection (Vassalli *et al*., 2002; Rothman *et al*., 2005; Michaloski *et al*., 2006; Clowney *et al*., 2012; Vassalli *et al*., 2011; Plessy *et al*., 2012). Homeodomains and O/E binding sites are relative enriched in OR proximal promoters and enhancers (Hirota *et al*., 2004) and it was shown that deletion of both sites from the minimal promoter of MOR23 in transgenic mouse line resulted in complete abolishment of expression (Vassalli *et al*., 2002). Later in another study, it was shown that a LIM-homedomain protein Lhx2 partakes in OR expression and conditional ablation of Lhx2 from MOE resulted in Class II type OR expression abolishment but had no significant effect on Class I type OR expression (Hirota *et al*., 2007).

Similar to ORs in the mouse MOE, the frequency of expression varies for individual ORs in the zebrafish OE (Kaya, 2016). Putative homeodomain and O/E binding sites were observed also within the proximal promoter of zebrafish ORs, pointing to a possible evolutionary conservation of regulatory mechanism in olfactory systems across species. As mentioned above, deletion of Lhx2 in the mouse resulted in selective loss of Class II OR expression while Class I ORs were not affected (Hirota *et al*., 2007). Considering the fact that fish possess almost only Class I receptors (Niimura and Nei, 2005; Alioto and Ngai, 2005; Tinaztepe, 2009), OR choice in fish might utilize different transcriptional proteins or combination thereof. Thus, subtle but key differences between fish and mammalians olfactory systems may exist. Here, I utilized the presumed regulatory sequences of the zebrafish *or101-1* gene. Interestingly, *or101-1* is the only fish OR that has close sequence similarity to mammalian Class II genes (Niimura and Nei, 2005; Tinaztepe, 2009). Therefore, it remains to be seen in as much observations made for the *or101-1* gene can be generalized across the zebrafish OR gene repertoire.

5.3. Expression Control of the *or101-1* Gene

Previous promoter bashing work on proximal promoter of *or101-1* showed that sequence upstream of the translational start codon ranging in length from 2.0 kb to 0.8 kb were able to drive the expression of marker gene at similar high levels (Söğünmez, 2012; Kazci, 2015). The deletion of a candidate O/E binding site within the minimal promoter (212 bp in the p0.8-eYFP construct) did not show any significant difference (Kazci, 2015). At this time it was speculated that the presence of a second candidate O/E binding site within the intron might be compensating for the loss of the upstream site. To test this idea, an intronic mutant version of p1.2-eYFP was prepared and injected. The results of this experiment were largely similar to removing the O/E binding site in the upstream promoter, supporting the idea of redundancy. Similar effects of redundancy have been observed for the mouse M71 OR gene, in which removal of an O/E site from a transgene had a larger effect than removal of the same site in the genome (Rothman *et al*., 2005). Thus, it appears that the larger context affects expression and that redundancy between site, even at a considerable distance of several 100 bps exists. An additional construct in which both sites were mutated showed 7.5% reduction in embryonic counts and 8% reduction in cell counts when injected. However, the double mutation did not abolish the expression of marker gene completely, indicating that other regulatory elements may be present within the proximal promoter. However, regardless of the size of the effect, O/E binding sites are involved in *or101-1* regulation.

5.4. Long- and Short-range Control of OR Expression

As mentioned earlier, the selection of OR genes in OE depends on LSD1 mediated activation by epigenetic de-repression (Lyons *et al*., 2013). Even though this mechanism may explain singular expression, it cannot directly explain differences in frequency of expression observed for individual OR genes. LSD1 is a lysine-specific demethylase that removes methyl groups from histone tails irrespective of their position on the protein. Thus, LSD1 can both erase methylation marks that are specific for heterochromatin and euchromatin and thereby act as an activator and silencer of transcription. It is conceivable

that LSD1-mediated gene activation requires specific recruitment elements that guide or bias the process towards certain OR genes. Markenscoff-Papadimitriou *et al*. (2014) proposed that cluster enhancers compete for LSD1 and differences in competitive strength may result in a non-uniform expression frequency for ORs. In contrast, however, an ectopically integrated construct containing the minimal MOR23 promoter could drive MOR23 expression similar to that of the original gene (Vassalli *et al*., 2002). Though long-range and short-range cis-acting elements might have a role in OR selection.

For a functional OSN, maintenance of OR expression is as crucial as initial selection of an OR gene. Transgenic analysis in mouse showed that proximal promoter could be enough to drive expression of an OR with a similar frequency as the original gene (Vassalli *et al*., 2002). This effect occurs even when the location of the transgene is not within the original or any other OR cluster, thus devoid of OR-specific long-range enhancer elements. However, this is not the case for all OR genes and expression of the majority of ORs may strictly depend on enhancers (Serizawa *et al*., 2002; Fuss *et al*., 2007; Nishizumi *et al*., 2007). Once activated, expression maintenance of the OR is crucial for both the function of the OSN and the olfactory system as such. Another question that arises is then, how the continuation of expression is achieved throughout the rest of the life of mature OSN. Both, removal of homeodomain and O/E binding sites from OR promoters and deletion of Lhx2 (Hirota *et al*., 2007) and Emx2 (McIntyre *et al*., 2008), which bind to homeodomain sites in mouse MOE resulted in loss of specific OR subsets. Ablation of the O/E-binding factors O/E2 and O/E3, however, resulted in dysregulation of axon guidance but did not alter the expression of mature OSN marker gene OMP (Wang *et al*., 2004). Thus, it appears that Lhx2/Emx2 and O/E proteins are involved in both OR selection and maintenance

5.5. The Role of O/E Sites in p1.2-eYFP Expression

While single mutations of putative O/E binding sites had no significant effect on expression of a p1.2-eYFP construct, double mutation of these sites resulted in statistically significant expression reduction. Yet, expression was not completely abolished even when both sites were removed, suggesting that tested O/E binding sites are not the only regulatory

elements within the proximal promoter of *or101-1*. Another candidate regulatory site previously identified on p1.2-eYFP is a sequence with similarity to a homeodomain site and is located 209 bp upstream of *or101-1* TSS. A previously tested p0.8-eYFP construct only comprises half of this motif and most likely renders the site non-functional but again the effect on expression of the transgene were only moderate (Kazcı, 2015). Depending on the stringency of the settings, motif scans on the 1.2 kb upstream of *or101-1* CDS can identify 16 additional homeodomain- and 2 additional O/E binding sites. These sites obtain different scores with a threshold of 80% similarity to MA0700.1 (human, Lhx2), MA0154.1 (mouse, Ebf1) and MA0154.2 (human, Ebf1) motifs (data not shown) and different scenarios can be invoked to explain the overall small impact of the deletion of homeodomain and O/E sites from the *or101-1* promoter sequence. The two sites tested here are the sites that received the highest scores, yet the underlying positional weight matrices for homeodomain and O/E sites may not accurately enough describe efficient fish motifs. Alternatively, multiple weak binding sites may compensate for the loss of strong sites, at least in the assay performed. Overall, however, the tested O/E binding sites appear to play a partial role in expression of p1.2-eYFP construct, though possibly in a redundant manner.

The quantitative effect of O/E and homeodoamin site removal in the fish constructs, together with the observations made for MOR23 and M71 transgenes in the mouse suggest a scenario in which the frequency and zonality of expression is somehow determined, or at least altered, by the activities of homeodomain binding and O/E proteins. Under the light of these, it is not far-fetching to propose that frequency and zonal information for OR genes is encoded, partially for some and possibly completely for others, by these motifs. Peculiarities, such as abundance, spacing, and conservation (i.e. strength of binding) may all affect OR expression in the spatial and quantitative domain. I suspect that the positional information is extracted from the promoter by differential binding dynamics of O/E proteins and the promoter. It is interesting to note that at least the known enhancers of OR gene expression also contain identical motifs and that these motifs may be necessary for efficient or regulated promoter-enhancer interactions.

5.6. Expression Variation due to Intron Splicing

As an alternative way of testing whether intronic sites, including the putative O/E site are relevant for *or101-1* expression an intron swap of the native *or101-1* intron with the first intron of the zebrafish *myod1* gene was performed. It was previously shown that removal of the entire intron resulted in severe loss of expression from otherwise highly expressed constructs (Kazcı, 2015). This effect could be due to the removal of relevant regulatory sites contained within the intron or the known fact that intron splicing can enhance expression of a gene up to several hundred folds (Shaul, 2017 and the references therein). Interestingly, when the intron was swapped a modest increase in expression of 4.5% at the level of transgene-expressing embryos or transgene-expressing individual OSNs was observed. There are several possible causes for the observed increase in expression, such as nucleotide variations in exon-intron junction or the unexpected or inadvertent presence of regulatory elements within the intronic sequence that was swapped in.

Enhancement of expression due to splicing is largely due to the involvement of U1 small nuclear ribonucleoprotein (snRNP) during transcription, whereas decoration with exon-junction complexes facilitates nuclear export and translation (Le Hir *et al*., 2001). The U1 snRNP interacts with the Cyclin H component of TFIIF, leading to phosphorylation of the C-Terminal Domain of RNA polymerase Pol II, which can result re-initiation of transcription or stimulation of promoter escape (Guiro and O'Reilly, 2015; O'Gorman *et al*., 2005). This interaction between TFIIF and U1 snRNP increases the recruitment of TFIIF and Pol II typically when a gene has a short 5'-UTR and the first intron is in close proximity to the promoter (Damgaard *et al*., 2008; Kwek *et al*., 2002). The binding affinity of U1 to 5'-splice sites (5'SS) also alters the expression of the gene, in the sense that mutations within 5'SS might increase or decrease the expression without disrupting pre-mRNA splicing *per se* (Furger *et al*., 2002). Eukaryotic promoters, in general, lack an innate directionality for the expression and thus a significant portion of transcription happens in the antisense direction generating upstream antisense RNA transcripts (uaRNA) which overall decrease expression. However, the presence of an intron shifts the balance in favor of sense mRNA through interaction of U1 snRNP with 5'SS of intron and the presence of additional U1-

binding sites within the exon and intron (Almada *et al*., 2013). Not only this interaction increase the fidelity of transcription, it also prevents premature cleavage and polyadenylation of mRNA thus increasing the functional mRNA levels (Almada *et al*., 2013; Berg *et al*., 2012; Kaida *et al*.,2010). Another mode in which U1 snRNP increases the number of mRNA molecules per cell is through facilitating gene looping and re-initiation by recycling transcription factors, preferentially in the sense direction (Agarwal and Ansari, 2016). Exon-junction complexes on mature and spliced mRNA facilitates transport of mRNA into the cytoplasm (Valencia *et al*., 2008; Le Hir *et al*., 2001) and stimulates binding of eEF4A and subsequently eEF4G, which is necessary for initiation of ribosome scanning

The significance of U1 for eukaryotic expression suggests that differences between the 5'SS between the native and the swapped intron may have effects on expression. The exon-intron junction sequence of AAGGT/CAGT of p1.2-eYFP was converted to AAGGT/GAGA in the p1.2-myod1-eYFP construct. Thus, two nucleotide changes occurred, a C to G transition in position 6 and a T to A transition in position 9. These changes are increasing and decreasing the similarity between the junction and the conserved U1 binding motif (5′-(C/A)AGGU(A/G)AGU-3), respectively (Piasecka *et al*., 2015). Since base pairing strength of G and C is greater than that of A and T, overall nucleotide changes are in favor of increasing the similarity between the junction sequence and conserved motif. This transition argues for a better recognition of the hybrid *or101-1*/*myod1* intron-exon junction in the transgenic construct by U1, which may lead to the observed increase in expression.

The constrains used to choose a suitable intron to replace the *or101-1* intron in the swap experiment were the length of the intron and the tissue in which the gene that contains the intron is normally expressed. The gene should not normally be expressed in the OE to minimize the risk of transferring a functional regulatory element. As a general conclusion then, the similarity in expression between p1.2-myod1-eYFP and p1.2-eYFP might mean that the expression reduction seen in intronless p1.2-eYFP injections was due to the loss of intron. In order to rule out the possibility of the *myod1* intron having regulatory function in OE, a motif scan was performed on the sequence, which unexpectedly returned motif hit for

mouse and human Ebf1 binding motifs, although with a lower conservation than the tested motifs. This raises another question: how likely is it to find Ebf1 motifs in the introns of other genes? Therefore, Ebf1 motif scans were performed on the first introns of five genes and six randomly generated DNA sequences with varying lengths between 100 bp to 1000 bp. With the exception of the *jag1a* intron, all the remaining sequences contained at least one Ebf1 binding motif (tested introns are from beta actin (*actb1a*), snaptotagmin, MTHFR, GAPDH and *jag1a* genes; data not shown). Thus, reasonable O/E binding motifs may be very common in the genome, in line with the redundant function of flanking genomic sequences of the M71 and MOR23 genes (Rothman *et al*., 2005).

5.7. The Role of Zbtb7b in Inhibitory Function of i562

In previous works, it was shown that deletion of either the 5'- or the 3'-half of i562 partially abolished its inhibitory effect (Söğünmez, 2012). When further scrutinized, a repetition of the sequence CNTCTGG was found in each half. The same sequence motif was shown to be enriched in mouse OR and VR promoters (Michaloski *et al*., 2011). As a critical test, removal of the motifs from transgenic constructs reduced the the inhibitory effect of i562 (Sancer, 2015). After showing that these motifs are somehow involved in inhibitory function of i562, the next step was to determine if any protein recognizes the sites and if soto identify the interacting protein. In their study, Michaloski *et al*. (2011) suggested that Zbtb7b, a known transcriptional repressor, might be the protein binding to the CNTCTGG sequence based on expression studies of ZBTB proteins in the mouse OE. A similar high expression of zbtb7b is seen in the zebrafish based on RNAseq data (Fuss, unpublished).

Zbtb7b is a key transcriptional inhibitor protein crucial in differentiation of $CD4^+/CD8^+$ double-positive immature T cells into $CD4^+/CD8^-$ single positive T helper cells. The default fate of thymocytes is to become single positive $CD4^-/CD8^+$ cytotoxic T cells. During the positive selection stage of maturation, Zbtb7b expression drives the silencing of Runx3 gene through epigenetic modifications and thus prevents development into cytotoxic cells (Egawa and Littman, 2008).

An antibody raised against mouse Zbtb7b was used in EMSA to investigate if Zbtb7b indeed binds to sequences of i562. In the assay in which wt30 probes of each site and wt100 probe of 5bs were incubated with mouse brain nuclear extract in the presence of antibody, a decrease in intensity of the shifted bands was observed despite the otherwise identical conditions. Not directly a clear supershift result, as no single band could be detected, the observation could mean that the fraction of probes that interacted with zbtb7b was affected in its mobility in the gel. It should be noted that the probes used here were relatively large and may contain interaction sites for other, uncharacterized or unidentified proteins. Thus, zbtb7b binding to the motifs may be masked by additional interactions at non-overlapping sites on the same probe. However, in competition assays for larger probes using a 30 bp oligonucleotide that overlaps the suspected motif, no change in binding could be observed, which argues against recognition of the two motifs by any protein. Western blot analysis performed using Zbtb7b antibody resulted in a multi-banded pattern showing low specificity of the antibody. Lack of binding to these motifs in zebrafish protein extract combined with low specificity of antibody might mean that the site was genuinely recognized by a protein present in mouse brain extract yet the antibody binding was too unspecific to deduce that the binding partner was Zbtb7b. In addition, zbtb7b is the most abundantly expressed but not the only expressed member of the ZBTB family in the OE. Thus, other members of the family could have redundant function but may not be recognized by the antibody.

5.8. The Role of the Motif CNTCTGG in Inhibitory Function of i562

The best argument for a specific function of the CNTCTGG in i562 is the fact that it occurs once in each half of the fragments that have been deleted from i562 in functional assays. One consistent observation throughout all EMSA experiments performed in this study was that the binding pattern of 5bs was different from that of 3bs. The observed banding patterns were different for 5bs and 3bs assays and the del100 and del200 probes of 5bs showed less binding with RLM11 or mouse OE nuclear extracts, while no intensity difference were observed for 3bs del100 and del200 probes. The motifs found by Michaloski *et al*. (2011) contained either an A or C nucleotide in the second position of CNTCTGG.

The sequence for 5bs is CCTCTGG while the sequence for 3bs is CTTGCTGG, which is deviating from the observed preference in the mouse. More direct experiments to identify the true binding partner of the sequence, such as yeast-one-hybrid assays or ChIP experiments could shed more light onto the nature of the specific effects seen for i562.

5.9. Alternative Candidate Proteins with Repressive Function

The candidate binding sites of CNTCTGG of i562 are highly similar to Zbtb18 (RP58) motif of human (MA0698.1). However, when i562 were scanned for the Zbtb18 motif with a threshold of 80% none of the predicted sites (5bs and 3bs) was recognized. When the threshold was lowered to 60%, there was only one hit per predicted sites in a total of 66 hits (data not shown). Interestingly, some of the sites predicted by the motif were apart from each other at exact multiples of 10.5, which is approximately the number of nucleotides for a complete helical turn of DNA. While three of these sites resides apart from each other by 63 bp (two of which are upstream to 5bs the other one on 5bs), the most upstream one of these three sites and another one downstream to 5bs but upstream to 3bs reside 378 bp apart.

Zbtb18 is a POZ domain containing protein of BTB ZF family just as Zbtb7b is (Aoki *et al*., 1998). Zbtb18 plays a crucial role in proliferation, migration and differentiation during the development of cerebral cortex, and it is a repressor that performs this function by interacting with Dnmt3a. Repression is achieved through recruitment of HDAC by Dnmt3a, which in this case does not act as a demethylase but an intermediary between Zbtb18 and HDAC. Thus, by indirectly recruiting HDAC to the site, Zbtb18 acts as a repressor through chromatin remodeling (Ohtaka-Maruyama *et al*., 2007).

All the predicted sites mentioned above point to a mechanism of inhibition more likely driven by Zbtb18. This may explain why the deletion of 5bs resulted in loss of i562 inhibitory function. Yet, it would not explain the result of 3bs deletion. Multiple bands seen in EMSAs for wt100 and wt200 indicates the recognition of these probes by multiple proteins, which might be individual different proteins, proteins complexed with other proteins or simply

protein that are homodimerized. Overall, these points to a complex dynamics of protein interactions between the repressor proteins and i562. To support or disprove the possible role of Zbtb18, these recently predicted sites could be mutated in injection constructs, binding dynamics of these could be tested with EMSA or a conditional deletion of Zbtb18 in OE followed by in situ for *or101-1* can be performed.

5.10. Conclusion

Overall, the positive regulation of *or101-1* is possibly based on additional and so far unrecognized binding sites that were not studied here. Mutation analysis of O/E sites indicated that they have a function, though minor in overall regulation of the gene. Since individual deletions of predicted O/E motifs or the combined deletion of both motifs did not drastically the expression, it is possible that presence of two and more O/E sites creates redundancy. The presence of intron in close proximity to promoter is also crucial for the expression of *or101-1*. It should be noted that this gene structure is seen in about 50% of OR genes, however, the feature does not seem to be correlated with frequency of expression (Kaya, 2016). Similarly, the underlying mechanism of negative regulation by i562 are not understood. However, data in hand suggest that its functions may be more relevant for insulation of the OR115/OR101-1 cluster from neighboring genes than to decrease the expression (or rather frequency) of *or101-1*. The inherent difficulty in discerning the true function of the sequence lies with the compact nature of the zebrafish genome and the fact that *or101-1* is the gene at the border of the cluster. Thus, any sequence upstream of *or101-1* may equally well be a promoter, a border, or a foreign element.

REFERENCES

Abdus-Saboor, I., Al Nufal, M.J., Agha, M. V., Ruinart De Brimont, M., Fleischmann, A. & Shykind, B.M. 2016. An Expression Refinement Process Ensures Singular Odorant Receptor Gene Choice. *Current Biology*. 26(8):1083–1090. DOI: 10.1016/j.cub.2016.02.039.

Agarwal, N. & Ansari, A. 2016. Enhancement of Transcription by a Splicing-Competent Intron Is Dependent on Promoter Directionality. *PLoS Genetics*. 12(5):1–20. DOI: 10.1371/journal.pgen.1006047.

Ahuja, G., Ivandic, I., Saltürk, M., Oka, Y., Nadler, W. & Korsching, S.I. 2013. Zebrafish crypt neurons project to a single, identified mediodorsal glomerulus. *Scientific reports*. 3:2063. DOI: 10.1038/srep02063.

Ahuja, G., Nia, S.B., Zapilko, V., Shiriagin, V., Kowatschew, D., Oka, Y. & Korsching, S.I. 2014. Kappe neurons, a novel population of olfactory sensory neurons. *Scientific reports*. 4:4037. DOI: 10.1038/srep04037.

Alioto, T.S. & Ngai, J. 2005. The odorant receptor repertoire of teleost fish. *BMC genomics*. 6:173. DOI: 10.1186/1471-2164-6-173.

Almada, A.E., Wu, X., Kriz, A.J., Burge, C.B. & Sharp, P.A. 2013. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature*. 499(7458):360–363. DOI: 10.1038/nature12349.

Bargmann, C.I., Hartwieg, E. & Horvitz, H.R. 1993. Odorant-selective genes and neurons mediate olfaction in C. elegans. *Cell*. 74(3):515–527. DOI: 10.1016/0092-8674(93)80053-H.

Barth, A.L., Dugas, J.C. & Ngai, J. 1997. Noncoordinate expression of odorant receptor genes tightly linked in the zebrafish genome. *Neuron*. 19(2):359–369. DOI: 10.1016/S0896-6273(00)80945-9.

Bayramli, X., Kocagöz, Y., Sakizli, U. & Fuss, S.H. 2017. Patterned Arrangements of Olfactory Receptor Gene Expression in Zebrafish are Established by Radial Movement of Specified Olfactory Sensory Neurons. *Scientific Reports*. 7(1):1–16. DOI: 10.1038/s41598-017-06041-1.

Belluscio, L., Koentges, G., Axel, R. & Dulac, C. 1999. A map of pheromone receptor activation in the mammalian brain. *Cell*. 97(2):209–220. DOI: 10.1016/S0092-8674(00)80731-X.

Berg, M.G., Singh, L.N., Younis, I., Liu, Q., Pinto, A.M., Kaida, D., Zhang, Z., Cho, S., *et al*. 2012. U1 snRNP determines mRNA length and regulates isoform expression. *Cell*. 150(1):53–64. DOI: 10.1016/j.cell.2012.05.029.

Bozza, T., McGann, J.P., Mombaerts, P. & Wachowiak, M. 2004. In vivo imaging of neuronal activity by targeted expression of a genetically encoded probe in the mouse. *Neuron*. 42(1):9–21. DOI: 10.1016/S0896-6273(04)00144-8.

Bozza, T., Vassalli, A., Fuss, S., Zhang, J.J., Weiland, B., Pacifico, R., Feinstein, P. & 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(2):220–233. DOI: 10.1016/j.neuron.2008.11.010.

Bren, A. & Eisenbach, M. 2000. How Signals Are Heard during Bacterial Chemotaxis : Protein-Protein Interactions in Sensory Signal Propagation MINIREVIEW How Signals Are Heard during Bacterial Chemotaxis : Protein-Protein Interactions in Sensory Signal Propagation. *Journal of bacteriology*. 182(24):6865–6873. DOI: 10.1128/JB.182.24.6865-6873.2000.Updated.

Bressel, O.C., Khan, M. & Mombaerts, P. 2016. Linear correlation between the number of olfactory sensory neurons expressing a given mouse odorant receptor gene and the total volume of the corresponding glomeruli in the olfactory bulb. *Journal of Comparative Neurology*. 524(1):199–209. DOI: 10.1002/cne.23835.

Buck, L. & Axel, R. 1991. A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell*. 65(1):175–187. DOI: 10.1016/0092-8674(91)90418-X.

Bushdid, C., Magnasco, M.O., Vosshall, L.B. & Keller, A. 2014. Humans can discriminate more than 1 trillion olfactory stimuli. *Science*. 343(6177):1370–1372. DOI: 10.1126/science.1249168.

Celik, A., S. H. Fuss, and S. I. Korsching, 2002, "Selective Targeting of Zebrafish Olfactory Receptor Neurons by the Endogenous OMP Promoter", *Neuroscience*, Vol. 15, pp. 798–806.

Chess, A., Simon, I., Cedar, H. & Axel, R. 1994. Allelic inactivation regulates olfactory receptor gene expression. *Cell*. 78(5):823–834. DOI: 10.1016/S0092-8674(94)90562-2.

Clowney, E.J., Legros, M.A., Mosley, C.P., Clowney, F.G., Markenskoff-Papadimitriou, E.C., Myllys, M., Barnea, G., Larabell, C.A., *et al*. 2012. Nuclear aggregation of olfactory receptor genes governs their monogenic expression. *Cell*. 151(4):724–737. DOI: 10.1016/j.cell.2012.09.043.

Dalton, R.P. & Lomvardas, S. 2015. Chemosensory Receptor Specificity and Regulation. *Annual Review of Neuroscience*. 38(1):331–349. DOI: 10.1146/annurev-neuro-071714-034145.

Dalton, R.P., Lyons, D.B. & Lomvardas, S. 2013. Co-Opting the Unfolded Protein Response to Elicit Olfactory Receptor Feedback. *Cell*. 155(2):321–332. DOI: 10.1016/j.cell.2013.09.033.

Damgaard, C.K., Kahns, S., Lykke-Andersen, S., Nielsen, A.L., Jensen, T.H. & Kjems, J. 2008. A 5′ Splice Site Enhances the Recruitment of Basal Transcription Initiation Factors In Vivo. *Molecular Cell*. 29(2):271–278. DOI: 10.1016/j.molcel.2007.11.035.

Dugas, J.C. & Ngai, J. 2001. Analysis and characterization of an odorant receptor gene cluster in the zebrafish genome. *Genomics*. 71(1):53–65. DOI: 10.1006/geno.2000.6415.

Dulac, C. & Axel, R. 1995. A novel family of genes encoding putative pheromone receptors in mammals. *Cell*. 83(2):195–206. DOI: 0092-8674(95)90161-2.

Egawa, T. & Littman, D.R. 2008. ThPOK acts late in specification of the helper T cell lineage and suppresses Runx-mediated commitment to the cytotoxic T cell lineage. *Nature Immunology*. 9(10):1131–1139. DOI: 10.1038/ni.1652.

Eggan, K., Baldwin, K., Tackett, M., Osborne, J., Gogos, J., Chess, A., Axel, R. & Jaenisch, R. 2004. Mice cloned from olfactory sensory neurons. *Nature*. 428(6978):44–49. DOI: 10.1038/nature02375.

Ehlers, M., Laule-Kilian, K., Petter, M., Aldrian, C.J., Grueter, B., Würch, A., Yoshida, N., Watanabe, T., *et al*. 2003. Morpholino antisense oligonucleotide-mediated gene knockdown during thymocyte development reveals role for Runx3 transcription factor in CD4 silencing during development of CD4-/CD8+ thymocytes. *Journal of immunology (Baltimore, Md. : 1950)*. 171:3594–3604. DOI: 10.4049/jimmunol.171.7.3594.

Feinstein, P. & Mombaerts, P. 2004. A contextual model for axonal sorting into glomeruli in the mouse olfactory system. *Cell*. 117(6):817–831. DOI: 10.1016/j.cell.2004.05.011.

Firestein, S. 2001. How the olfactory system makes sense of scents. *Nature*. 413(6852):211–218. DOI: 10.1038/35093026.

Freitag, J., Ludwig, G., Andreini, I., Rössler, P. & Breer, H. 1998. Olfactory receptors in aquatic and terrestrial vertebrates. *Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology*. 183(5):635–650. DOI: 0.1007/s003590050287.

Friedrich, R.W. & Korsching, S.I. 1997. Combinatorial and chemotopic odorant coding in the zebrafish olfactory bulb visualized by optical imaging. *Neuron*. 18(5):737–752. DOI: 10.1016/S0896-6273(00)80314-1.

Furger, A., O'Sullivan, J.M., Binnie, A., Lee, B.A. & Proudfoot, N.J. 2002. Promoter proximal splice sites enhance transcription. *Genes and Development*. 16(21):2792–2799. DOI: 10.1101/gad.983602.

Fuss, S.H. & Ray, A. 2009. Mechanisms of odorant receptor gene choice in Drosophila and vertebrates. *Molecular and Cellular Neuroscience*. 41(2):101–112. DOI: 10.1016/j.mcn.2009.02.014.

Fuss, S.H., Omura, M. & Mombaerts, P. 2007. Local and cis Effects of the H Element on Expression of Odorant Receptor Genes in Mouse. *Cell*. 130(2):373–384. DOI: 10.1016/j.cell.2007.06.023.

Fuss, S.H., Zhu, Y. & Mombaerts, P. 2013. Odorant receptor gene choice and axonal wiring in mice with deletion mutations in the odorant receptor gene SR1. *Molecular and Cellular Neuroscience*. 56:212–224. DOI: 10.1016/j.mcn.2013.05.002.

Guiro, J. & O'Reilly, D. 2015. Insights into the U1 small nuclear ribonucleoprotein complex superfamily. *Wiley Interdisciplinary Reviews: RNA*. 6(1):79–92. DOI: 10.1002/wrna.1257.

Hamdani, E.H. & Døving, K.B. 2007. The functional organization of the fish olfactory system. *Progress in Neurobiology*. 82(2):80–86. DOI: 10.1016/j.pneurobio.2007.02.007.

Hamdani, E.H., Alexander, G. & Døving, K.B. 2001. Projection of sensory neurons with microvilli to the lateral olfactory tract indicates their participation in feeding behaviour in crucian carp. *Chemical senses*. 26(9):1139–1144. DOI: 10.1093/chemse/26.9.1139.

Hanchate, N.K., Kondoh, K., Lu, Z., Kuang, D., Ye, X., Qiu, X., Pachter, L., Trapnell, C., *et al*. 2015. Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. *Science*. 350(6265):1251–1255. DOI: 10.1126/science.aad2456.

Hansen, A. & Zeiske, E. 1998. The peripheral olfactory organ of the zebrafish, Danio rerio: An ultrastructural study. *Chemical Senses*. 23(1):39–48. DOI: 10.1093/chemse/23.1.39.

Hansen, A., Rolen, S.H., Anderson, K., Morita, Y., Caprio, J. & Finger, T.E. 2003. Correlation between olfactory receptor cell type and function in the channel catfish. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 23(28):9328–9339. DOI: 23/28/9328.

Hirota, J. & Mombaerts, P. 2004. The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory sensory neurons. *Proc Natl Acad Sci U S A*. 101(23):8751–8755. DOI: 10.1073/pnas.0400940101.

Hirota, J., Omura, M. & Mombaerts, P. 2007. Differential impact of Lhx2 deficiency on expression of class I and class II odorant receptor genes in mouse. *Molecular and Cellular Neuroscience*. 34(4):679–688. DOI: 10.1016/j.mcn.2007.01.014.

Hoppe, R., Frank, M.E., Breer, H. & Strotmann, J. 2003. The clustered olfactory receptor gene family 262: genomic organization, promoter elements, and interacting transcription factors. *Genome Res*. 13:2674–2685. DOI: 10.1101/gr.1372203.1.

Hoppe, R., Breer, H. & Strotmann, J. 2006. Promoter motifs of olfactory receptor genes expressed in distinct topographic patterns. *Genomics*. 87(6):711–723. DOI: 10.1016/j.ygeno.2006.02.005.

Hussain, A., Saraiva, L.R. & Korsching, S.I. 2009. Positive Darwinian selection and the birth of an olfactory receptor clade in teleosts. *Proceedings of the National Academy of Sciences*. 106(11):4313–4318. DOI: 10.1073/pnas.0803229106.

Ishii, T., Serizawa, S., Kohda, A., Nakatani, H., Shiroishi, T., Okumura, K., Iwakura, Y., Nagawa, F., *et al*. 2001. Monoallelic expresion of the odourant receptor gene and axonal projection of olfactory sensory neurones. *Genes to Cells*. 6:71–78. DOI: 10.1046/j.1365-2443.2001.00398.x.

Iwata, R., Kiyonari, H. & Imai, T. 2017. Mechanosensory-Based Phase Coding of Odor Identity in the Olfactory Bulb. *Neuron*. 96(5):1139–1152.e7. DOI: 10.1016/j.neuron.2017.11.008.

Iwema, C.L. & Schwob, J.E. 2003. Odorant receptor expression as a function of neuronal maturity in the adult rodent olfactory system. *Journal of Comparative Neurology*. 459(3):209–222. DOI: 10.1002/cne.10583.

Jacoby, E., Bouhelal, R., Gerspacher, M. & Seuwen, K. 2006. The 7TM G-protein-coupled receptor target family. *ChemMedChem*. 1(8):760–782. DOI: 10.1002/cmdc.200600134.

Jung, D. Alt, F. W. 2004. Unraveling V(D)J Recombination Insights into Gene Regulation. *Cell*. 116(2):299–311. DOI: 10.1016/S0092-8674(04)00039-X.

Kaida, D., Berg, M.G., Younis, I., Kasim, M., Singh, L.N., Wan, L. & Dreyfuss, G. 2010. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature*. 468(7324):664–668. DOI: 10.1038/nature09479.

Kappes, D.J., He, X. & He, X. 2006. Role of the transcription factor Th-POK in CD4:CD8 lineage commitment. *Immunological Reviews*. 209:237–252. DOI: 10.1111/j.0105-2896.2006.00344.x.

Kazci, Y.E. 2015. In Vivo Characterization of the OR101-1 Gene Proximal Promoter Region in Zebrafish. Bogazici University.

Khan, M., Vaes, E. & Mombaerts, P. 2011. Regulation of the probability of mouse odorant receptor gene choice. *Cell*. 147(4):907–921. DOI: 10.1016/j.cell.2011.09.049.

Killackey, H.P., Rhoades, R.W. & Bennett-Clarke, C.A. 1995. The formation of a cortical somatotopic map. *Trends in Neurosciences*. 18(9):402–407. DOI: 10.1016/0166-2236(95)93937-S.

Kobayakawa, K., Kobayakawa, R., Matsumoto, H., Oka, Y., Imai, T., Ikawa, M., Okabe, M., Ikeda, T., *et al*. 2007. Innate versus learned odour processing in the mouse olfactory bulb. *Nature*. 450(7169):503–8. DOI: 10.1038/nature06281.

Krautwurst, D., Yau, K.W. & Reed, R.R. 1998. Identification of ligands for olfactory receptors by functional expression of a receptor library. *Cell*. 95(7):917–926. DOI: 10.1016/S0092-8674(00)81716-X.

Kubick, S., Strotmann, J., Andreini, I. & Breer, H. 1997. Subfamily of olfactory receptors characterized by unique structural features and expression patterns. *Journal of neurochemistry*. 69(2):465–475. DOI: 10.1046/j.1471-4159.1997.69020465.x.

Kudrycki, K., Stein-Izsak, C., Behn, C., Grillo, M., Akeson, R. & Margolis, F.L. 1993. Olf-1-binding site: characterization of an olfactory neuron-specific promoter motif. *Molecular and cellular biology*. 13(5):3002–3014. DOI: 10.1128/MCB.13.5.3002.

Kudrycki, K.E., Buiakova, O., Tarozzo, G., Grillo, M., Walters, E. & Margolis, F.L. 1998. Effects of mutation of the Olf-1 motif on transgene expression in olfactory receptor neurons. *Journal of Neuroscience Research*. 52(2):159–172. DOI: 10.1002/(SICI)1097-4547(19980415)52:2<159::AID-JNR4>3.0.CO;2-9.

Kwek, K.Y., Murphy, S., Furger, A., Thomas, B., O'Gorman, W., Kimura, H., Proudfoot, N.J. & Akoulitchev, A. 2002. U1 snRNA associates with tfiih and regulates transcriptional initiation. *Nature Structural Biology*. 9(11):800–805. DOI: 10.1038/nsb862.

Landry, D.B., Engel, J.D. & Sen, R. 1993. Functional GATA-3 binding sites within murine CD8 alpha upstream regulatory sequences. *The Journal of experimental medicine*. 178(September):941–949. DOI: 10.1084/jem.178.3.941.

Laska, M., Seibt, a & Weber, a. 2000. "Microsmatic" primates revisited: olfactory sensitivity in the squirrel monkey. *Chemical senses*. 25(1):47–53. DOI: 10.1007/s00464-013-3109-y.

Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M.J. 2001. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO Journal*. 20(17):4987–4997. DOI: 10.1093/emboj/20.17.4987.

Lewcock, J.W. & 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(4):1069–1074. DOI: 10.1073/pnas.0307986100.

Li, J., Ishii, T., Feinstein, P. & Mombaerts, P. 2004. Odorant receptor gene choice is reset by nuclear transfer from mouse olfactory sensory neurons. *Nature*. 428(6981):393–399. DOI: 10.1038/nature02433.

Liberles, S.D. & Buck, L.B. 2006. A second class of chemosensory receptors in the olfactory epithelium. *Nature*. 442(7103):645–650. DOI: 10.1038/nature05066.

Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J. & Axel, R. 2006. Interchromosomal Interactions and Olfactory Receptor Choice. *Cell*. 126(2):403–413. DOI: 10.1016/j.cell.2006.06.035.

Lu, S. & Cullen, B.R. 2003. Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells. *Rna*. 618–630. DOI: 10.1261/rna.5260303.export.

Lyons, D.B., Allen, W.E., Goh, T., Tsai, L., Barnea, G. & Lomvardas, S. 2013. XAn epigenetic trap stabilizes singular olfactory receptor expression. *Cell*. 154(2):325–336. DOI: 10.1016/j.cell.2013.06.039.

Magklara, A., Yen, A., Colquitt, B.M., Clowney, E.J., Allen, W., Markenscoff-Papadimitriou, E., Evans, Z.A., Kheradpour, P., *et al*. 2011. An epigenetic signature for monoallelic olfactory receptor expression. *Cell*. 145(4):555–570. DOI: 10.1016/j.cell.2011.03.040.

Malnic, B., Hirono, J., Sato, T. & Buck, L.B. 1999. Combinatorial receptor codes for odors. *Cell*. 96(5):713–23. DOI: 10.1016/S0092-8674(00)80581-4.

Malnic, B., D., Gonzalez-Kristeller, and L., Gutiyama, 2010, "7 Odorant Receptors", *Neurobiology of Olfaction*, Vol. 12, p. 181.

Manuscript, A. & Magnitude, S. 2013. NIH Public Access. 31(9):1713–1723. DOI: 10.1109/TMI.2012.2196707.Separate.

Maresh, A., Rodriguez Gil, D., Whitman, M.C. & Greer, C.A. 2008. Principles of glomerular organization in the human olfactory bulb - Implications for odor processing. *PLoS ONE*. 3(7). DOI: 10.1371/journal.pone.0002640.

Markenscoff-Papadimitriou, E., Allen, W.E., Colquitt, B.M., Goh, T., Murphy, K.K., Monahan, K., Mosley, C.P., Ahituv, N., *et al*. 2014. Enhancer Interaction Networks as a Means for Singular Olfactory Receptor Expression. *Cell*. 159(3):543–557. DOI: 10.1016/j.cell.2014.09.033.

McIntyre, J.C., Bose, S.C., Stromberg, A.J. & McClintock, T.S. 2008. Emx2 stimulates odorant receptor gene expression. *Chemical Senses*. 33(9):825–837. DOI: 10.1093/chemse/bjn061.

Michaloski, J.S., Galante, P.A.F. & Malnic, B. 2006. Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences. *Genome Research*. 16:1091–1098. DOI: 10.1101/gr.5185406.sults.

Michaloski, J.S., Galante, P.A.F., Nagai, M.H., Armelin-Correa, L., Chien, M.S., Matsunami, H. & Malnic, B. 2011. Common promoter elements in odorant and vomeronasal receptor genes. *PLoS ONE*. 6(12). DOI: 10.1371/journal.pone.0029065.

Miyamichi, K. 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. *Journal of Neuroscience*. 25(14):3586–3592. DOI: 10.1523/JNEUROSCI.0324-05.2005.

Mombaerts, P. 2004a. Genes and ligands for odorant, vomeronasal and taste receptors. *Nature Reviews Neuroscience*. 5(4):263–278. DOI: 10.1038/nrn1365.

Mombaerts, P. 2004b. Odorant receptor gene choice in olfactory sensory neurons: The one receptor-one neuron hypothesis revisited. *Current Opinion in Neurobiology*. 14(1):31–36. DOI: 10.1016/j.conb.2004.01.014.

Mombaerts, P., Wang, F., Dulac, C., Chao, S.K., Nemes, A., Mendelsohn, M., Edmondson, J. & Axel, R. 1996. Visualizing an Olfactory Sensory Map. 87:675–686.

Lin, H., and R. Grosschedl, 2002, "Failure of B-Cell Differentiation in Mice Lacking the Transcription Factor EBF", *Nature*, Vol. 376, No. 6537, pp. 263–267.

Niimura, Y. & Nei, M. 2007. Extensive gains and losses of olfactory receptor genes in mammalian evolution. *PLoS ONE*. 2(8). DOI: 10.1371/journal.pone.0000708.

Nishizumi, H., Kumasaka, K., Inoue, N., Nakashima, A. & 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(50):20067–20072. DOI: 10.1073/pnas.0706544105.

Niimura, M. & Nei, M. 2007. Evolutionary dynamics of olfactory receptor genes in Drosophila species. *Proc Natl Acad Sci U S A*. 104(17):7122–7127. DOI: 0702133104 [pii]\r10.1073/pnas.0702133104.

O'Gorman, W., Thomas, B., Kwek, K.Y., Furger, A. & Akoulitchev, A. 2005. Analysis of U1 small nuclear RNA interaction with cyclin H. *Journal of Biological Chemistry*. 280(44):36920–36925. DOI: 10.1074/jbc.M505791200.

Oka, Y., Saraiva, L.R. & Korsching, S.I. 2012. Crypt neurons express a single v1r-related ora gene. *Chemical Senses*. 37(3):219–227. DOI: 10.1093/chemse/bjr095.

Olender, T., Fuchs, T., Linhart, C., Shamir, R., Adams, M., Kalush, F., Khen, M. & Lancet, D. 2004. The canine olfactory subgenome. *Genomics*. 83(3):361–372. DOI: 10.1016/j.ygeno.2003.08.009.

Piasecka, A., Brzuzan, P., Woźny, M., Ciesielski, S. & Kaczmarczyk, D. 2014. Splice-acceptor site mutation in p53 gene of hu888 zebrafish line. *Journal of Applied Genetics*. 56(1):115–121. DOI: 10.1007/s13353-014-0239-4.

Plessy, C., Pascarella, G., Bertin, N., Akalin, A., Carrieri, C., Vassalli, A., Lazarevic, D., Severin, J., *et al*. 2012. Promoter architecture of mouse olfactory receptor genes. *Genome Res*. 22(3):486–497. DOI: 10.1101/gr.126201.111.

Prasad, B.C. & Reed, R.R. 1999. Chemosensation: Molecular mechanisms in worms and mammals. *Trends in Genetics*. 15(4):150–153. DOI: 10.1016/S0168-9525(99)01695-9.

Qasba, P. & Reed, R.R. 1998. Tissue and zonal-specific expression of an olfactory receptor transgene. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 18(1):227–236.

Ressler, K.J., Sullivan, S.L. & Buck, L. 1993. A Zonal O rganization of Odorant Receptor Gene Expression in the O lfactory Epithelium. 73:597–609.

Ressler, K.J., Sullivan, S.L. & Buck, L.B. 1994. Information coding in the olfactory system: evidence for a stereotyped and higly organized epitopc map in the olfactory bulb. *Cell*. 79:1245–1255.

Rodriguez, I., Feinstein, P. & Mombaerts, P. 1999. Variable patterns of axonal projections of sensory neurons in the mouse vomeronasal system. *Cell*. 97(2):199–208. DOI: 10.1016/S0092-8674(00)80730-8.

Rothman, A., Feinstein, P., Hirota, J. & Mombaerts, P. 2005. The promoter of the mouse odorant receptor gene M71. *Molecular and cellular neurosciences*. 28(3):535–46. DOI: 10.1016/j.mcn.2004.11.006.

Rouquier, S., Taviaux, S., Trask, B.J., Brand-Arpon, V., van den Engh, G., Demaille, J. & Giorgi, D. 1998. Distribution of olfactory receptor genes in the human genome. *Nature genetics*. 18(3):243–50. DOI: 10.1038/ng0398-243.

Sancer, G. 2015. Regulation of Olfactory Receptor Gene Choice By Negative Regulatory Elements. Bogazici University.

Sato, Y. 2005. Mutually Exclusive Glomerular Innervation by Two Distinct Types of Olfactory Sensory Neurons Revealed in Transgenic Zebrafish. *Journal of Neuroscience*. 25(20):4889–4897. DOI: 10.1523/JNEUROSCI.0679-05.2005.

Sato, Y., Miyasaka, N. & Yoshihara, Y. 2007. Hierarchical regulation of odorant receptor gene choice and subsequent axonal projection of olfactory sensory neurons in zebrafish. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 27(7):1606–15. DOI: 10.1523/JNEUROSCI.4218-06.2007.

Seki, T., Namba, T., Mochizuki, H. & Onodera, M. 2007. Clustering, migration, and neurite formation of neural precursor cells in the adult rat hippocampus. *J Comp Neurol*. 502(2):275–290. DOI: 10.1002/cne.

Serizawa, S., Ishii, T., Nakatani, H., Tsuboi, A., Nagawa, F., Matsuda, Y., Suzuki, M., Yamamori, T., *et al*. 2000. Mutually exclusive expression of odorant receptor transgenes. 687–693.

Serizawa, S., Miyamichi, K. & Sakano, H. 2016. One neuron – one receptor rule in the mouse olfactory system One neuron – one receptor rule in the mouse olfactory system. 302(March):2088–2095. DOI: 10.1016/j.tig.2004.09.006.

Shaul, O. 2017. How introns enhance gene expression. *International Journal of Biochemistry and Cell Biology*. 91:145–155. DOI: 10.1016/j.biocel.2017.06.016.

Shepherd, G.M. 2004. The human sense of smell: Are we better than we think? *PLoS Biology*. 2(5):572–575. DOI: 10.1371/journal.pbio.0020146.

Shi, P. & Zhang, J. 2007. Comparative genomic analysis identifies an evolutionary shift of vomeronasal receptor gene repertoires in the vertebrate transition from water to land. *Genome Research*. 17(2):166–174. DOI: 10.1101/gr.6040007.

Shykind, B.M., Rohani, S.C., O'Donnell, S., Nemes, A., Mendelsohn, M., Sun, Y., Axel, R. & Barnea, G. 2004. Gene switching and the stability of odorant receptor gene choice. *Cell*. 117(6):801–815. DOI: 10.1016/j.cell.2004.05.015.

Song, H.G., Jae, Y.K., Hyung, S.H., Bae, Y.C. & Moon, C. 2008. First contact to odors: Our current knowledge about odorant receptors. *Sensors*. 8(10):6303–6320. DOI: 10.3390/s8106303.

Söğünmez, N. 2012. Transgenic Analysis of the Zebrafish or101-1 Gene Promoter Promoter. Bogazici University.

Stormo, G.D. 2013. Modeling the specificity of protein-DNA interactions. *Quantitative Biology*. 1(2):115–130. DOI: 10.1007/s40484-013-0012-4.

Sullivan, S.L., Adamson, M.C., Ressler, K.J., Kozak, C. a & 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(2):884–888. DOI: 10.1073/pnas.93.2.884.

Talavage, T.M., Ledden, P.J., Benson, R.R., Rosen, B.R. & Melcher, J.R. 2000. Frequency-dependent responses exhibited by multiple regions in human auditory cortex. *Hearing Research*. 150(1–2):225–244. DOI: 10.1016/S0378-5955(00)00203-3.

Tapscotr, S.J., Davis, R.L., Thayer, M.J., Cheng, P., Weintraub, H. & Lassar, A.B. 1988. Homology Region to Convert Fibroblasts to Myoblasts. 295(1987):405–411.

Thommesen, G. 1983. Morphology, distribution, and specificity of olfactory receptor cells in salmonid fishes. *Acta physiologica Scandinavica*. 117(2):241–9. DOI: 10.1111/j.1748-1716.1983.tb07203.x.

Tian, H. & Ma, M. 2008. Differential development of odorant receptor expression patterns in the olfactory epithelium: A quantitative analysis in the mouse septal organ. *Developmental Neurobiology*. 68(4):476–486. DOI: 10.1002/dneu.20612.

Tinaztepe, E. 2009. Class restriction in odorant receptor gene regulation. Bogazici University.

Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature*. 302(5909):575–581. DOI: 10.1038/302575a0.

Tsuboi, A., Yoshihara, S., Yamazaki, N., Kasai, H., Asai-Tsuboi, H., Komatsu, M., Serizawa, S., Ishii, T., *et al*. 1999. Olfactory neurons expressing closely linked and homologous odorant receptor genes tend to project their axons to neighboring glomeruli on the olfactory bulb. *Journal of Neuroscience*. 19(19):8409–8418.

Tsuboi, A., Miyazaki, T., Imai, T. & Sakano, H. 2006. Olfactory sensory neurons expressing class I odorant receptors converge their axons on an antero-dorsal domain of the olfactory bulb in the mouse. *European Journal of Neuroscience*. 23(6):1436–1444. DOI: 10.1111/j.1460-9568.2006.04675.x.

Valencia, P., Dias, A.P. & Reed, R. 2008. Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*. 105(9):3386–91. DOI: 10.1073/pnas.0800250105.

Vassalli, A., Rothman, A., Feinstein, P., Zapotocky, M. & Mombaerts, P. 2002. Minigenes Impart Odorant Receptor-Specific Axon Guidance in the Olfactory Bulb. *Neuron*. 35(4):681–696. DOI: 10.1016/S0896-6273(02)00793-6.

Vassalli, A., Feinstein, P. & Mombaerts, P. 2011. Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice. *Molecular and cellular neurosciences*. 46(2):381–96. DOI: 10.1016/j.mcn.2010.11.001.

Vassar, R. & Ngai, J. 1993. Spatial Segregation of Odorant Receptor Expression in the Mammalian Olfactory Epithelium. 74.

Vassar, R., Chao, S.K., Sitcheran, R., Nuiiez, M., Vosshall, L.B. & Axel, R. 1994. Topographic O rganization of Sensory Projection to the O lfactory Bulb. *Cell*. 79:981–991. DOI: 10.1016/0092-8674(94)90029-9.

Verbeurgt, C., Wilkin, F., Tarabichi, M., Gregoire, F., Dumont, J.E. & Chatelain, P. 2014. Profiling of olfactory receptor gene expression in whole human olfactory mucosa. *PLoS ONE*. 9(5):21–26. DOI: 10.1371/journal.pone.0096333.

Wagner, S., Gresser, A.L., Torello, A.T. & Dulac, C. 2006. A Multireceptor Genetic Approach Uncovers an Ordered Integration of VNO Sensory Inputs in the Accessory Olfactory Bulb. *Neuron*. 50(5):697–709. DOI: 10.1016/j.neuron.2006.04.033.

Wang, S.S. 2004. Genetic disruptions of O/E2 and O/E3 genes reveal involvement in olfactory receptor neuron projection. *Development*. 131(6):1377–1388. DOI: 10.1242/dev.01009.

Wang, F., Nemes, A., Mendelsohn, M. & Axel, R. 1998. Odorant receptors govern the formation of a precise topographic map. *Cell*. 93(1):47–60. DOI: 10.1016/S0092-8674(00)81145-9.

Weth, F., Nadler, W. & Korsching, S. 1996. Nested expression domains for odorant receptors in zebrafish olfactory epithelium. *Proceedings of the National Academy of Sciences of the United States of America*. 93(23):13321–13326. DOI: 10.1073/pnas.93.23.13321.

Yamamoto, M. and K. Ueda, 1978. Comparative morphology of fish olfactory epithelium-III. Cypriniformes. Bull. Japan Soc. Sci. Fish., 44: 1201-1206.

Yoshihara, Y., Kawasaki, M., Tamada, a, Fujita, H., Hayashi, H., Kagamiyama, H. & Mori, K. 1997. OCAM: A new member of the neural cell adhesion molecule family related to zone-to-zone projection of olfactory and vomeronasal axons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 17(15):5830–5842.

Young, J.M. 2002. Different evolutionary processes shaped the mouse and human olfactory receptor gene families. *Human Molecular Genetics*. 11(5):535–546. DOI: 10.1093/hmg/11.5.535.

Young, J.M. & Trask, B.J. 2002. The sense of smell: genomics of vertebrate odorant receptors. *Human molecular genetics*. 11(10):1153–1160. DOI: 10.1093/hmg/11.10.1153.

Young, J.M., Kambere, M., Trask, B.J. & Lane, R.P. 2005. Divergent V1R repertoires in five species: Amplification in rodents, decimation in primates, and a surprisingly small repertoire in dogs. *Genome Research*. 15(2):231–240. DOI: 10.1101/gr.3339905.

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., *et al*. 2017. Ensembl 2018. *Nucleic Acids Research*. 46(February):754–761. DOI: 10.1093/nar/gkx1098.

Zhang, X. & Firestein, S. 2002. The olfactory receptor gene superfamily of the mouse. *Nature neuroscience*. 5(2):124–133. DOI: 10.1038/nn800.

Zhang, X., Rodriguez, I., Mombaerts, P. & Firestein, S. 2004. Odorant and vomeronasal receptor genes in two mouse genome assemblies. *Genomics*. 83(5):802–811. DOI: 10.1016/j.ygeno.2003.10.009.

Zhang, X., Rogers, M., Tian, H., Zhang, X., Zou, D.-J., Liu, J., Ma, M., Shepherd, G.M., *et al*. 2004. High-throughput microarray detection of olfactory receptor gene expression in the mouse. *Proc Natl Acad Sci USA*. 101:14168–14173. DOI: 10.1073/pnas.0405350101.

Zhang, X., Zhang, X. & Firestein, S. 2007. Comparative genomics of odorant and pheromone receptor genes in rodents. *Genomics*. 89(4):441–450. DOI: 10.1016/j.ygeno.2007.01.002.

Zhang, X., and S. Firestein, 2009, "Genomics of Olfactory Receptors", *Results and Problems in Cell Differentiation*, Vol. 47, pp. 25–36.

Zozulya, S., Echeverri, F. & Nguyen, T. (in press). The human olfactory receptor repertoire. *Genome biology*. 2(6):research0018.1–12. DOI: 10.1186/gb-2001-2-6-research0018.

APPENDIX A: EQUIPMENTS

Table A. 1. The table of equipments.

Product Name	**Manufacturer Company**
FemtoJet	Eppendorf, Germany
Tetramin	Tetra, Germany
Aquatic Habitats	Pentair Aquatic Eco-systems, Inc., USA
Stemi 2000-C stereomicroscope	Zeiss, Germany
GELoader tips, 0.5-20 µl	Eppendorf, Germany
Crystal Sea Marinemix	Marine Enterprises International, USA
P-97 Micropipette Puller	Sutter Instrument, Co., USA
Centrifuge 5424	Eppendorf, Germany
Centrifuge 5417R	Eppendorf, Germany
C1000 Thermal Cycler	Bio-Rad, USA
SZ61 stereomicroscope	Olympus, USA
Tissue-Plus O.C.T compound	Fisher HealthCare, USA
IKA Color Squid magnetic stirrer	IKA works, Inc., USA
05-090-128 Mini Centrifuges	Fisher Scientific, Korea
Vortex-Genie 2	Scientific Industries, Inc., USA
Mini Trans-Blot Cell	Bio-Rad, USA
Filter tips	
(10 µl, 20 µl, 100 µl, 200 µl, 1000 µl)	Griener Bio-One, Germany
Electrical Balance	Sartorius, Germany
Microwave oven	Vestel, Turkey
-20˚C Freezer	Ugur, Turkey
-20˚C Freezer	Arcelik, Turkey
Refrigerator	Arcelik, Turkey
Magnetic stirrer, RH B 2	IKA workd, Inc., USA
pH-meter, pH315i	WTW, Germany
Unstirrer water bath, JB Aqua 12, Grant	Wolflabs, UK
Stage Vacuum Pump, Ve115	Value, China
Universal Incubator	Binder, Germany
Sodium Bicarbonate	Aquatic Eco-Systems, Inc., USA
Agarose Gel Tank, Mini-Sub Cell GT Cell	BioRad, China
Capillary glass (1.00 mm x 0.5 mm x 4’’)	A-M Systems, Inc., USA
Single channel micropipettes	Eppendorf, Germany
(10 µl, 20 µl, 100 µl, 200 µl, 1000 µl)	
Super Pap Pen	Liquid Blocker, Japan
Knete	Idena, Germany
High Pure PCR purification Kit	Roche, Germany
Serological pipettes	
(5ml, 10ml, 25ml, 50ml)	Greiner Bio-One, Germany
Glass bottle	Isolab, Germany
Cryostat CM3050S	Leica Biosystems, Germany
Adhesion slides, Superfrost Plus	VWR, Germany

Table A.1. The table of equipments (cont'd).

Bottle Warmer, 12312	Weewell, China
Dishwasher, Melabor G 7783	Miele, Germany
Swiftlock Front loading Autoclave	Astell, UK
Laboratory Drying Oven, KD 200	Nüve, Turkey
Tabletop Autoclaves	
(CV-EL 12L/18L, ES1.0038)	CertoClav, Austria
Discard Container Autoclave, Midas 55	Priorclave, UK
Orbital Shaker, Rotamax120	Heidolph, Germany
Nylon Membrane, Positively Charged	Roche, Germany
Biotin 3' End Labelling Kit, 89818	Thermo Scientific, USA
Syringe filter, 0.22 µm, 99722	TPP, Switzerland
Dual Intensity Transilluminator	UVP, USA
Analytical Balance, ME54	Mettler Toledo, India
Ice Flaker	Brema, Italy
Sonicator, Sonoplus HD2070	Bandeline, Germany
UV Crosslinker, Spectrolinker XL-1000	Spectroline, USA
Colibri Microvolume spectrometer	Titertek Berthold, Germany
UV-visible spectrophotometer, 8453	Agilent, USA
Shake 'n' stack hybridization oven, 6240	Thermo Fisher Scientific, USA
Rnase-ExitusPlus, A7153,0500	PanReac Applichem, Spain
SuperScript First Strand Synthesis	Invitrogen, USA
System for RT-PCR	
Transcriptor High Fidelity	
cDNA Synthesis Kit	Roche, Germany
-86C ULT Freezer	ThermoForma, USA
Allegra X-30R Benchtop Centrifuge	Beckman Coulter Inc., Germany
Floor Model Centrifuge, L2-21	Beckman Coulter, USA
Incubated Benchtop Orbital Shaker	Thermo Scientific, USA
Liquid Nitrogen Dewar, 10LD	Taylor-Wharton, UK
Brine Shrimp Artemia Cysts	
(A. franciscana)	Salt Lake Aquafeed, USA
Waving platform shaker, Polymax 2040	Heidolph, Germany
Microscope Camera for	
Basic Microscopy, EC4	Leica Microsystems, Switzerland
Polypropylene Tubes with two position	Greiner Bio-One, Germany
Vent Stopper, 187 261	
MicroCentrifuge tubes 1.5 ml	Capp ApS, Denmark
VWR micro cover glass, 22x50 mm	VWR, USA
Parafilm™	Parafilm, USA
Acrylamide/Bis-Acrylamide solution, 30%	Sigma-Aldrich, Germany
N',N',N',N' – Tetramthyl ethylenediamine	
(TEMED),10 ml	Merck, Germany
GeneJET Plasmid Miniprep Kit	Thermo Scientific, Lithuania
WesternBright™ Sirius, K-12043-D10	Advansta, USA
Pellet Pestle Motor	Kimble Kontess, USA

Table A.1. The table of equipments (cont'd).

Whetstone	Dan's Whetstone Campany,Inc., USA
Forceps, FST	Dumont, Switzerland
Instrument oil, No.29055-00	FST, Switzerland
Graduate cylinder	Isolab, Germany
Beaker	Isolab, Germany
Polypropylene centrifuge tubes	
(15 ml, 50 ml)	Sarstedt, Germany
PCR® strip Tubes	Axygen, USA

APPENDIX B: CONSUMABLES

Table B. 1. The table of consumables.

Product Name	**Manufacturer Company**
1 kb DNA ladder	New England Biolabs, U.S.A. (N3232)
100 bp DNA ladder	New England Biolabs, U.S.A. (N3231
Acetic acid	Sigma-Aldrich, Germany (71251-5ML-F)
Acrylamide/Bisacrylamide solution	BioRad, U.S.A (1610156)
Agarose SeaKem®	Cambrex, U.S.A. (50004)
Ampicillin	Sigma-Aldrich,Germany (10835242001)
Anti-Zbtb7B	Genetex, U.S.A (GTX48861)
Bovine Serum Albumine (BSA)	New England Biolabs, U.S.A. (B9001)
Chloroform	Sigma-Aldricg, Germany (288306-1L)
Dithiothreitol (DTT)	Sigma-Aldrich, Germany (10197777001)
EDTA Disodium salt	Sigma-Aldrich., U.S.A. (E5134 - 1 kg)
Ethanol Absolute	Sigma-Aldrich, U.S.A. (34870)
Ethidium bromide	Sigma Life Sciences, U.S.A. (E1510-1 ml)
Glycerol	Sigma-Aldrich, U.S.A. (G5516-500 ml)
HEPES	Sigma-Aldrich, Germany (H3375-100G)
LB agar	Sigma Life Sciences, U.S.A. (SL08394)
LB Broth	Sigma-Aldrich, U.S.A. (L7658- 1 kg)
Magnesium Chloride, Molecular Grade	Promega, U.S.A. (A3511)
Magnesium Sulfate	Sigma-Aldrich, U.S.A. (M7506)
Onetaq® DNA polymerase	New England Biolabs, U.S.A (M0480L)
pGEM®-T Easy Vector System	Promega, U.S.A. (A1360)
Pierce™ 3’ End Labelling DNA Biotinylation Kit	Thermo Fisher Scientific, USA (89818)
Potassium Chloride	Sigma-Aldrich, U.S.A. (P9541)
Q5® High Fidelity DNA Polymerase	NEB, England (M0491S)
SacI	Invitrogen (15240-201)
SacII	New England Biolabs, U.S.A. (R0157S)
Sodium Acetate	Sigma-Aldrich, U.S.A. (S8625)
Sodium Chloride	Sigma-Aldrich, U.S.A. (S7653 - 1 kg)
Sodium Hydroxide	Sigma-Aldrich, U.S.A. (S8045 - 1 kg)
SpeI	New England Biolabs, U.S.A (R0133L)
T4 DNA ligase	New England Biolabs, U.S.A (M0202L)
Tris-HCl	Sigma-Aldrich, U.S.A. (T6066)

APPENDIX C: RAW INJECTION DATA

Table C.1. Raw injection data.

p1.2 eYFP - pOmp mCherry					
	5.08.2016	6.08.2016	13.08.2016	19.08.2016	20.08.2016
injected	68	110	274	69	146
alive	24	43	89	24	77
pOmp positive	7	21	49	20	50
p1.2 positive	4	13	28	15	39

5.08.2016							
p1.2 pOmp double positive				pomp (single) positive			
OE1		OE2		OE1		OE2	
pOmp	p1.2	pOmp	p1.2	pOmp	p1.2	pOmp	p1.2
16	11	21	16	5	3	1	1
3	1	6	4	8	6	3	3
15	11	5	2	10	5	2	1
1	1						

6.08.2016							
pOmp	p1.2	pOmp	p1.2	pOmp	p1.2	pOmp	p1.2
15	10	28	11	0	0	6	1
2	2	7	4	4	1	0	0
0	0	3	5	0	0	8	5
4	3	5	2	9	6	0	0
17	9	8	6	1	0	0	0
16	11	13	10	0	0	6	5
5	2	0	0	2	2	1	0
4	2	0	0	1	0	2	0
1	0	0	0				
5	3						
11	6	7	7				
0	0	10	6				
2	1	0	0				

13.08.2016							
pOmp	p1.2	pOmp	p1.2	pOmp	p1.2	pOmp	p1.2
2	2	18	14	13	9	10	5
3	1	0	0	0	0	1	0
25	22	23	21	7	2	7	4
5	3	3	3	0	0	0	0
8	5	9	5	0	0	2	0
3	7	9	8	5	2	3	3
5	2	0	0	4	1	3	3
14	9	7	1	4	0	5	2
2	1	2	3	0	0	3	2
9	7	13	8	22	11	24	3
4	5	1	1	7	6	8	7
12	9	13	10	11	6	8	7
7	5	18	13	6	7	0	0
6	3	7	2	0	0	14	9
7	4	11	7	2	0	3	0
6	2	6	2	10	7	8	5
8	5	3	2	0	0	9	3
5	2	0	0	0	0	3	0
4	2	12	8	5	1	3	3
11	10	5	1	3	0	5	2
14	13	9	8	0	0	6	2
9	6	9	5				
10	7	6	4				
7	4	4	3				
8	4	14	8				
16	12	10	11				

19.08.2016							
pOmp	p1.2	pOmp	p1.2	pOmp	p1.2	pOmp	p1.2
14	6	10	7	7	2	8	4
13	7	11	4	11	5	5	0
25	10	32	16	5	1	4	1
20	8	28	17	7	2	13	3
7	3	4	2	16	13	12	4
23	15	15	12				
11	6	21	9				
29	20	32	28				
18	11	11	7				

3	1	20	16
11	5	54	36
8	4	37	21
17	8	16	11
24	18	24	24
x	x	35	27

20.08.2016							
pOmp	p1.2	pOmp	p1.2	pOmp	p1.2	pOmp	p1.2
1	0	9	2	12	7	3	3
11	6	3	1	5	3	0	0
20	11	6	3	12	6	14	6
19	13	17	12	2	0	0	0
12	8	15	12	12	5	0	0
12	4	14	12	1	1	0	0
7	5	5	2	26	18	19	11
6	3	0	0	4	1	5	2
14	10	17	14	1	0	0	0
9	7	16	8	0	0	2	0
18	6	15	10	2	1	0	0
6	3	0	0				
11	7	10	8				
11	7	1	0				
10	6	4	2				
11	9	6	2				
10	8	6	2				
6	1	22	12				
15	5	14	5				
6	5	0	0				
13	2	11	6				
20	9	8	0				
10	13	1	1				
11	10	19	18				
0	0	3	2				
6	3	11	8				
15	13	12	10				
14	10	14	15				
12	5	10	8				
0	0	1	1				
0	0	7	6				
3	1	16	9				
0	0	8	6				
7	5	0	0				
1	0	6	6				
11	9	7	3				
7	4	13	7				
5	6	27	15				
24	13	11	11				

mut p1.2 eYFP pomp mCherry							
	12.11.2016	18.11.2016	3.02.2017	4.02.2017	11.03.2017	17.03.2017	24.03.2017
injected	130	51	26	117	127	314	87
alive	51	24	6	40	52	124	32
pOmp positive	22	6	3	18	16	64	21
p1.2 positive	5	4	2	13	6	32	16

9.04.2017	14.04.2017	21.04.2017
224	131	77
26	9	23
17	6	21
8	5	13

double positive					pomp positive			
OE1		OE2			OE1		OE2	
12.11.2016								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
18	10				7	2	0	0
11	3				1	0	6	1
5	1	14	2		4	2	7	1
7	3	3	2		7	4	8	5
8	7	4	2		3	0	2	1
					2	1	4	2
					2	0		
					4	4	1	0
					11	5	5	0
					9	1	2	0
					4	4	0	0
					17	10	0	0
					1	0	7	5
					5	1	11	2
					11	4	9	0

					2	0	18	13
18.11.2016								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
15	6	10	1		0	0	5	2
0	0	10	7		5	2	5	3
3	1	3	0					
1	1	8	3					
3.02.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
15	7	2	0		3	0	0	0
15	10	13	12					
4.02.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
6	2	10	0		6	4	2	0
5	1	7	2		4	3	0	0
0	0	5	2		3	1	0	0
4	2	1	0		0	0	2	0
9	6	8	5		9	6	7	1
4	4	3	3					
2	1	0	0					
7	5	2	0					
1	0	3	3					
0	0	2	2					
2	2	5	3					
2	0	5	2					
2	0	3	1					
11.03.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
1	1	0	0		5	4	1	0
7	5	4	0		8	1	0	0
26	19	13	9		0	0	2	0
9	7	10	10		1	1	0	0
9	6	0	0		1	1	0	0
18	9	0	0		5	0	9	2
					0	0	1	0
					8	2	7	4
					3	0	4	1
					0	0	1	0
17.03.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
11	2	7	5		15	8	0	0
10	6	0	0		4	3	0	0
8	4	15	4		13	7	0	0
9	6	6	4		3	1	3	3
4	3	9	5		0	0	1	0
7	4	7	6		0	0	2	1
3	1	6	2		3	2	4	1
0	0	11	10		1	0	5	2
6	2	7	3		0	0	5	2
2	1	0	0		6	1	5	4
0	0	11	10		4	1	0	0
3	2	0	0		5	1	7	2
14	10	11	6		0	0	1	0
7	7	4	1		1	0	3	0
6	4	4	4		10	4	4	0
9	8	9	2		3	2	9	3
9	7	0	0		1	0	3	0
0	0	1	1		10	4	4	0
3	2	8	2		3	2	9	3
3	3	1	1		3	1	5	1
3	2	8	3		0	0	8	4
4	4	15	7		2	0	1	0
7	2	10	2		5	1	2	0
8	6	8	6		2	0	9	0
14	8	7	4		2	1	0	0
6	0	5	2		2	0	2	1
7	4	6	3		0	0	3	0
36	21	27	16					
0	0	4	3					
7	5	0	0					
24.03.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
38	23	9	6		9	7	3	2
9	7	9	5		26	11	13	7
2	2	4	10		15	8	5	0
6	2	0	0					
6	1	7	1					
0	0	3	0					

12	4	2	2
7	2	9	2
8	3	13	3
8	4	21	11
21	10	11	3
22	13	8	4
13	8	21	9
18	14	12	8

9.04.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
0	0	1	1		0	0	9	5
0	0	1	1		1	0	1	1
3	0	15	4		1	0	1	0
9	4	0	0		0	0	2	0
14	10	10	8		1	1	2	0
9	5	23	15		0	0	1	0
2	1	0	0		14	0	12	1
					0	0	3	0
					4	3	0	0

14.04.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
0	0	8	8		7	4	3	0
13	9	7	6					
18	18	24	24					
7	2	2	2					
2	2	2	1					

21.04.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
30	21	43	29		5	2	12	8
8	4	5	2		0	0	1	0
21	15	6	2		4	2	2	1
9	4	11	7		1	0	1	0
43	33	47	40		4	1	6	1
17	13	7	7		1	1	4	2
8	3	5	2		1	1	2	0
8	7	5	3		4	4	6	3
7	3	11	5					
6	4	8	5					
15	12	21	18					
6	2	4	1					
15	6	5	4					

ddp1.2 eYFP pomp mCherry

	19.05.2017	20.05.2017	21.05.2017	26.05.2017	27.05.2017	3.06.2017
injected	165	230	200	186	150	94
alive	98	51	54	96	60	15
pOmp positive	78	32	46	43	49	10
p1.2 positive	43	17	33	22	24	7

double positive					pomp positive			
OE1		OE2			OE1		OE2	

19.05.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
24	12	15	4		39	14	23	11
18	16	5	5		18	11	18	16
24	16	14	11		15	7	6	4
9	9	8	8		6	1	0	0
24	21	18	11		13	4	8	5
13	9	8	7		19	5	15	5
37	39	16	18		0	0	4	0
19	10	10	2		14	3	6	1
12	15	14	14		18	10	18	3
23	8	17	4		21	13	14	5
36	29	44	31		7	4		
10	3	17	8		6	4	12	4
16	11	23	16		26	17	13	5
21	17	24	22		12	7	11	8
28	17	51	24		7	2	13	8
20	17	21	7		6	3	15	9
21	21	32	32		4	3	7	2
9	6	12	8		6	4	11	3
17	6	26	7		3	1	0	0
		17	8		5	3	19	8
9	4	1	0		12	3	15	9
38	30	11	11		19	8	4	2
37	25	36	23		5	3	1	0
19	11	27	9		2	2	30	25
32	23	16	17		2	0	1	0

9	6	24	20
16	10	8	4
4	3	14	0
3	0	0	0
28	16	25	12
10	7	8	5
8	1	26	13
3	0	9	1
7	3	10	9
4	3		

14	10	12	9
12	0	5	0
14	6	8	5
0	0	8	2
4	1	0	0
5	2	5	0
15	0	13	3
13	7	11	6
7	3	4	0
8	1	3	0

16	4	14	11
9	2	9	1
28	22	29	20
5	1	8	3
39	21	37	14
5	1	8	3
2	1	0	0
17	6	39	15

20.05.2017

pOmp	p1.2	pOmp	p1.2
5	0	2	0
6	4	7	3
13	6	1	0
1	0	8	4
6	3	5	1
1	0	9	2
1	1	0	0
7	5	2	1
0	0	2	1
7	9	18	16
6	2	4	3
0	0	12	10
24	24	21	15
15	2	16	6
4	0	0	0
16	10	9	7
18	7	17	6

pOmp	p1.2	pOmp	p1.2
3	0	0	0
5	0	4	2
0	0	2	0
1	0	1	0
7	3	10	4
17	1	2	0
3	2	4	3
5	2	3	0
9	2	5	5
0	0	7	3
6	1	12	0
3	0	0	0
6	0	4	1
7	4	22	12
7	1	3	1

21.05.2017

pOmp	p1.2	pOmp	p1.2
13	8	6	6
35	16	15	9
37	26	48	30
4	2	6	2
23	12	16	5
20	11	20	17
12	9	28	20
38	22	13	10
12	7	10	7
24	18	29	17
9	3	14	9
17	11	34	25
9	4	13	10
3	3	22	8
14	14	26	25
6	2	8	4
21	17	21	15
16	13	17	7
15	14	25	17
11	2	7	3
10	6	19	10
17	11	8	6
9	8	22	13
46	33	5	3
3	3	0	0
40	26	19	12
32	24	19	13
25	11	26	12
17	14	43	30
14	7	11	5
20	4	28	4
21	9	32	13
18	12	9	3

pOmp	p1.2	pOmp	p1.2
13	2	9	4
1	0	6	1
4	1	2	0
6	2	7	6
5	0	8	1
16	14	8	0
10	1	7	2
16	11	10	5
22	4	30	19
18	7	19	7
12	8	10	11
4	3	3	0
17	3	6	0

26.05.2017

pOmp	p1.2	pOmp	p1.2
1	0	6	1
4	1	2	0
7	3	0	0
4	3	7	2

pOmp	p1.2	pOmp	p1.2
4	1	9	3
0	0	0	0
15	6	7	3
2	1	9	7

0	0	4	1		11	4	11	7
5	2	1	1		4	1	6	3
3	2	6	0		0	0	3	0
1	1	9	6		3	1	0	0
13	8	0	0		6	2	1	1
3	2	3	1		8	3	3	1
12	7	11	3		9	9	6	3
19	3	16	9		9	2	11	7
25	6	26	20		0	0	2	1
10	4	9	3		2	0	3	1
12	10	9	6		3	0	2	0
23	10	2	0		3	1	0	0
18	5	15	7		5	5		
5	4	1	1					
27.05.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
15	2	18	3		5	0		
7	3	5	1		3	0	0	0
11	2	12	8		4	1	2	0
14	7	6	1		9	3	13	5
17	10	12	7		0	0	1	0
33	24	32	13		3	0	2	0
3	0	12	3		11	2	15	3
14	5	14	3		0	0	2	1
13	3	16	4		6	1	2	0
12	3	3	0		8	4	10	4
11	4	13	6		2	2	0	0
14	5	7	2		3	0	2	1
8	3	12	3		13	8	20	6
13	11	23	19		0	0	4	1
10	3	4	1		22	9	40	12
1	1	2	1		4	4	7	1
14	4	16	3		0	0	3	0
7	5	2	2		6	2	3	1
11	4	11	3		0	0	2	0
4	6	10	9		17	8	23	6
16	4	5	3		3	1	1	0
2	2				0	0	5	1
14	8	8	3		4	2	0	0
11	9	2	0		3	1	7	3
					2	0	7	3
3.06.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
2	1	3	2		1	0	1	0
7	6	3	1		6	1	3	0
2	0	4	1		1	0	0	0
5	2	0	0					
3	0	1	1					
28	21	12	10					
42	27	6	8					

p1.2 myod1 eYFP pomp mCherry

	1.04.2017	2.04.2017	8.04.2017	13.04.2017	14.04.2017	15.04.2017	22.04.2017
injected	114	165	138	47	156	115	123
alive	19	30	29	23	73	10	15
pOmp positive	13	23	26	22	57	8	8
p1.2 positive	10	16	18	14	35	5	7

double positive					pomp positive			
OE1		OE2			OE1		OE2	
1.04.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
3	1	5	3		0	0	8	5
12	11	1	0		2	0	0	0
6	3	5	4					
12	7	0	0					
0	0	21	8					
0	0	15	13					
8	3	9	6					
8	7	6	3					
4	1	6	4					
2.04.2017								
pOmp	p1.2	pOmp	p1.2		pOmp	p1.2	pOmp	p1.2
18	15	29	18		6	0	1	1
8	4	8	5		10	3	0	0
24	17	19	13		0	0	2	0
14	12	15	14		0	0	6	0
11	7	37	32		8	0	0	0

12	5	16	11
45	20	18	11
0	1	19	12
19	16	29	27
9	6	7	5
17	13	15	9
12	12	19	15
21	16	35	36
30	23	16	12
19	9	12	10
21	16	7	2

0	0	1	4
9	6	7	5

8.04.2017

pOmp	p1.2	pOmp	p1.2
12	7	0	0
0	0	2	3
2	1	9	5
23	18	20	11
8	5	12	9
3	2	14	8
10	8	11	5
12	8	23	18
9	4	12	7
0	0	2	2
17	14	14	10
0	0	15	11
10	6	14	11
20	17	18	14
13	9	22	12
3	1	12	6
3	2	6	6
15	11	7	5

pOmp	p1.2	pOmp	p1.2
19	11	7	5
14	7	5	3
14	6	5	2
5	2	10	5
0	0	4	0
18	11	15	11
7	3	1	1
3	0	5	5

13.04.2017

pOmp	p1.2	pOmp	p1.2
9	4	6	5
0	0	12	8
9	6	8	5
16	14	0	0
15	10	0	0
20	5	13	10
29	25	31	19
12	2	4	1
1	0	15	2
22	12	0	0
0	0	16	11
10	5	14	3
11	6	28	18
11	10	12	6

pOmp	p1.2	pOmp	p1.2
1	0	0	0
14	7	8	2
2	0	12	6
14	5	25	9
3	0	5	0
12	5	14	3
5	2	18	9
19	3	23	6

14.04.2017

pOmp	p1.2	pOmp	p1.2
3	1	8	3
61	41	20	41
18	13	20	14
16	14	8	6
16	7	27	19
9	3	5	3
7	2	12	4
18	11	20	13
6	4	12	7
10	8	38	27
30	20	14	5
27	20	31	31
16	13	15	3
11	0	16	6
10	2	12	10
10	5	17	12
4	0	4	2
5	4	6	5
1	0	6	4
31	25	28	27
25	17	8	7
14	11	11	6
16	14	14	12
3	3	25	11
6	2	6	3
7	7	35	34
9	7	5	3
10	9	8	7
14	6	15	14

19	14	18	12
11	6	13	11
24	23	6	6
0	0	16	15
1	0	13	9

15.04.2017							
	p1.2	pOmp	p1.2	pOmp	p1.2	pOmp	p1.2
4	1	10	8	4	3	17	13
5	3	2	2	11	10	8	8
9	8	12	5	0	1	5	5
3	2	2	0				
15	13	15	12				

22.04.2017							
pOmp	p1.2	pOmp	p1.2	pOmp	p1.2	pOmp	p1.2
7	5	1	0	3	0	0	0
1	0	9	8				
8	6	1	0				
26	21	9	5				
21	16	0	0				
1	0	8	5				
5	6	5	4				

p2.5 eYFP pomp mCherry

	3.06.2017	8.06.2017	9.06.2017	15.06.2017	16.06.2017	17.06.2017
injected	122	60	63	56	76	95
alive	53	9	32	32	40	42
pOmp positive	45	5	14	30	26	35
p1.2 positive	22	3	13	23	17	20

double positive				pomp positive			
OE1		OE2		OE1		OE2	
pOmp	p2.5	pOmp	p2.5	pOmp	p2.5	pOmp	p2.5
15	6	11	5	3	0	0	0
2	0	7	3	0	0	1	0
4	4	5	4	16	9	8	4
4	1	15	4	9	5	16	9
11	3	20	3	13	6	14	5
6	5	9	4	10	3	7	2
32	17	31	15	4	1	9	3
8	5	11	4	6	3	2	2
5	2	0	0	23	8	10	5
14	7	6	0	0	0	3	0
6	4	13	6	6	2	0	0
11	6	14	7	2	0	4	1
14	7	8	3	8	7	4	1
36	19	20	5	2	2	5	3
26	23	12	6	8	4	11	6
10	6	11	3	4	2	4	0
6	5	3	3	4	1	9	3
5	2	3	2	5	0	6	0
5	0	7	1	8	4	4	1
		10	3	13	6	8	3
8	3	5	2	0	0	7	2
20	12	14	11	0	0	2	0
				0	0	2	0

8.06.2017							
pOmp	p1.2	pOmp	p1.2	pOmp	p1.2	pOmp	p1.2
15	8	16	8	10	4	5	5
7	4	3	2	7	5	9	4
				4	1	7	4

9.06.2017							
pOmp	p2.5	pOmp	p2.5	pOmp	p2.5	pOmp	p2.5
27	8	23	12	9	3	1	0
13	10	17	2	8	6	12	9
9	2	2	1	2	2	8	2
7	3	22	8	4	2	13	9
12	5	20	4	4	2	5	0
11	5	15	15	11	9	25	16
24	14	2	1	8	4	16	4
33	31			6	2	7	2
22	10	13	5	11	3	5	5
30	15	19	12	2	0	3	0
4	2	12	8	1	0	2	0
10	8	0	0				
9	3	12	13				

15.06.2017							
pOmp	p2.5	pOmp	p2.5	pOmp	p2.5	pOmp	p2.5
10	4	9	3	18	10	16	8
2	1	10	7	15	8	21	8

39	26	42	19
2	1	19	11
19	9	16	10
15	7	9	2
5	1	7	3
26	9	16	9
8	5	34	23
22	15	2	2
17	6	15	3
14	8	20	5
5	1	6	1
25	20	22	7
21	10	15	6
6	3	9	3
21	8	5	3
31	18	12	1
13	6		
22	9	16	10
5	3		
13	10	22	16
5	2	13	9

25	4	15	5
36	14	16	2
0	0	0	0
23	12	4	2
7		4	

16.06.2017

pOmp	p2.5	pOmp	p2.5
9	3	9	7
11	3	5	1
6	3	3	2
5	4	15	12
6	2	5	2
6	4	5	5
9	5	32	24
7	2		
18	12	6	4
12	11	9	7
18	11	4	2
11	4	6	1
11	11	7	3
10	5	3	1
6	4	10	10
7	6	5	2
4	2	0	0

pOmp	p2.5	pOmp	p2.5
5	1	2	0
15	7	15	9
5	2	6	1
0	0	0	0
2	0	0	0
1	0	0	0
1	0	0	0
14	2	14	1

17.06.2017

pOmp	p2.5	pOmp	p2.5
24	14	20	11
23	2	10	2
21	16	8	2
17	5	11	7
12	12	16	14
67	40	31	15
21	10	21	11
2	1	11	6
31	7	24	9
7	2	12	6
15	10	20	10
32	20	28	20
25	13	39	20
28	12	15	8
19	9	20	10
4	2	5	3
15	7	12	4
44	26	29	28
21	6	29	10
5	0	8	3

pOmp	p2.5	pOmp	p2.5
5	0	6	2
4	1	5	2
16	2	16	3
9	4	20	7
0	0	3	3
5	1	4	0
4	2	5	1
0	0	6	0
4	2	9	4
5	1	9	5
18	6	20	11
31	23	28	21
0	0	2	0
4	2	5	1
8	6	8	5

APPENDIX D: DETAILED FIGURE LEGENDS

Figure 4.1. Transgene expression of the p1.2-eYFP construct. A. Representative confocal micrographs of an olfactory epithelium expressing the p1.2-eYFP (green, left) or the pomp-mCherry reference construct (red, center). Strong expression can be observed in individual olfactory sensory neurons. B. Schematic representation of the p1.2-eYFP construct; candidate O/E binding sites are shown as black boxes, the intron is marked by a dotted line. C. Efficiency of p1.2-eYFP transgene expression per embryo (top) or number of cells (bottom) relative to pomp-mCherry.

Figure 4.11. Transgene expression of the mut-p1.2-eYFP construct. A. Representative confocal micrographs of an olfactory epithelium expressing the mut-p1.2-eYFP (green, left) or the pomp-mCherry reference construct (red, center). B. Schematic representation of the mut-p1.2-eYFP construct; the upstream and deleted intronic O/E binding site are shown as black and red boxes, respectively, the intron is marked by a dotted line. C. Efficiency of mut-p1.2-eYFP transgene expression per embryo (top) or number of cells (bottom) relative to pomp-mCherry.

Figure 4.12. Transgene expression of the ddp1.2-eYFP construct. A. Schematic representation of the ddp1.2-eYFP construct; the deleted upstream and intronic O/E binding site are shown as red boxes, the intron is marked by a dotted line. B. Efficiency of ddp1.2-eYFP transgene expression per embryo (top) or number of cells (bottom) relative to pomp-mCherry.

Figure 4.13. Transgene expression of the p1.2-myod1-eYFP construct. A. Representative confocal micrographs of an olfactory epithelium expressing the p1.2-myod1-eYFP (green, left) or the pomp-mCherry reference construct (red, center). B. Schematic representation of the p1.2-myd1-eYFP construct; the replaced intron is marked by a dotted line. C. Efficiency of p1.2-myod1-eYFP transgene expression per embryo (top) or number of cells (bottom) relative to pomp-mCherry.

Figure 4.14. Summary and statistical analysis of p1.2-derived transgenic constructs. Expression frequency of various transgenic constructs at the level of embryos (left) and per cell analysis (right). The bars indicate means ± SEM, the brackets indicate statistical significance levels.

Figure 4.15. Schematic genomic organization of the OR115/OR101 gene cluster. The or101-1 gene (blue) is located at the far end of the combined OR115/OT101 gene cluster on chromosome 21. The arrows indicate direction of transcription, bars detail relative expression frequencies as determined by RNA-seq. The i562 sequence may comprise a boundary element of the cluster.

Figure 4.16. Biotinylation efficiency of EMSA probes. Raw chemiluminescence (top) of serial dilutions (1x, 0.1x, 0.01x) for 5' (left) and 3' (right) probe sets and quantification of biotinylation efficiency (bottom).

Figure 4.17. Western blot analysis for Zbtb7b protein. Nuclear (nc.) protein extracts from RLM11 (left), mouse OE, brain (br) and thymus (th) were probed with a human-specific anti-Zbtb7b antibody. The indicated numbers shows the 75 kDa and 50 kDa bands from the protein marker (ladder).

Figure 4.18. EMSA for 5' binding site with RLM11 protein extracts. Wild type (wt) and mutated (del) DNA probes covering the 5' CNTCTGG sequence were incubated with RLM11 nuclear protein extract (RLM11 nuc. ext.) and analyzed by non-denaturing polyacrylamide gel electrophoresis. The symbols above lanes indicate the presence (+) or omission (-) of protein extract from the reaction.

Figure 4.19. EMSA for 3' binding site with RLM11 protein extracts. Wild type (wt) and mutated (del) DNA probes covering the 5' CNTCTGG sequence were incubated with RLM11 nuclear protein extract (RLM11 nuc. ext.) and analyzed by non-denaturing polyacrylamide gel electrophoresis. The symbols above lanes indicate the presence (+) or omission (-) of protein extract from the reaction.

Figure 4.11. EMSA for 5' and 3' binding site with protein extracts from mouse tissues. Probes covering the 5' (left) or 3' (right) CNTCTGG sequence were incubated with nuclear protein extract (mouse nuc. ext.) from various mouse tissues (top) and analyzed by non-denaturing polyacrylamide gel electrophoresis (- = no protein control).

Figure 4.12. Supershift assay for 5' and 3' binding site with OE and brain protein extracts. Mouse OE (left) or brain (center, right) protein extracts were incubated before (pre) or after (post) addition of the wt30 probes for the 5' or 3' (indicated at top) binding sites.

Figure 4.13. Supershift assay for 5' and 3' binding sites with brain protein extracts. EMSA with 5' and 3' probes and mouse brain protein extracts in the presence (+) and absence (-) of anti-Zbtb7b antibody.

Figure 4.14. EMSA analysis of 5' probes with mouse OE protein extracts. Probes for the 5' binding site of various length that did (wt30, wt100, wt200) or did not (del100, del200) include the CNTCTTG motif were incubated with mouse olfactory epithelium protein extract.

Figure 4.15. EMSA analysis of 5’ and 3’ probes with zebrafish nose whole protein extract. 5‘ probes (wt100, wt200), 3’ probes (wt100, wt200) and a random probe were incubated with zebrafish nose whole protein extract. *an experimental is present in these lanes.

Figure 4.16. EMSA analysis of 5’ and 3’ probes in the presence of competitor with zebrafish nose whole protein extract. 5’ probes (wt100, wt200) and 3’ probes (wt100, wt200) were incubated with zebrafish nose whole protein extract in the presence (+) or absence (-) of oligo competitor.

Figure 4.17. EMSA analysis of 5’ and 3’ probes with different competitor concentrations and with zebrafish nose whole protein extract. 5’ and 3’ probes (wt200) were incubated with zebrafish nose whole protein extract in the presence of increasing oligo competitor concentrations (250X, 500X 750X).

Figure 4.18. EMSA analysis 3’ probes with different competitor concentrations and with zebrafish nose whole protein extract. 3’ probes (wt200) were incubated with zebrafish nose whole protein extract in the presence of increasing oligo competitor concentrations (250X, 500X 750X).