IN VIVO CHARACTERIZATION OF THE OR101-1 GENE PROXIMAL PROMOTER REGION IN ZEBRAFISH

by

Yusuf E. Kazcı

B.S., Molecular Biology and Genetics, Bilkent University, 2012

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

2015

IN VIVO CHARACTERIZATION OF THE OR101-1 GENE PROXIMAL PROMOTER REGION IN ZEBRAFISH

APPROVED BY:

Assoc. Prof. Stefan H. Fuss ....

(Thesis Supervisor)

Assoc. Prof. Uygar H. Tazebay ....

Assoc. Prof. N.C. Tolga Emre ....

Date of Approval: 22.05.2015

ACKNOWLEDGEMENT

I would like to thank to Stefan Fuss for presenting me the opportunity to work in his lab and for the phlogenetic footprinting performed in this study.

I would like to thank to Ahmet Burak Kaya for the bioinformatics analyzes presented in this study.

I was supported by the Scientific and Technological Research Council of Turkey (TÜBITAK) with the scholarship in conjunction with “2210-E Doğrudan Yurt İçi Yüksek Lisans Burs Programı (2012-2)”.

Work carried out in this project was supported by the Scientific and Technological Research Council of Turkey (TÜBITAK) Grant number: 112T168 - "Zebra balığında koku reseptör gen düzenleyici bölgelerin işlevsel karakterizasyonu".

ABSTRACT

IN VIVO CHARACTERIZATION OF THE OR101-1 GENE PROXIMAL PROMOTER REGION IN ZEBRAFISH

In vertebrates, each olfactory sensory neuron (OSN) of the nose expresses only a single allele of a single molecular chemoreceptor gene from a large genomic repertoire. This monogenic mode of olfactory receptor (OR) expression in OSNs is critical for the sense of smell because it determines the sensitivity of OSNs to odorants and instructs OSN axons to establish appropriate synaptic connections in the brain. However, the molecular mechanisms regulating OR expression are not fully understood. Over the past five years a principle mechanism to explain monogenic OR expression has emerged that is based on LSD1-mediated changes in the epigenetic signatures of OR gene loci. In this model, LSD1 would modify repressive histone marks with low probability to ensure expression of only a single OR per OSN. Yet, not all ORs are expressed with the same probability and / or within the same spatial domain of the olfactory tissue, suggesting that LSD1 is somehow 'guided' to OR gene loci in a biased fashion. Two types of OR locus-biased control mechanisms were proposed so far; long-range control by distant *cis*-acting locus control regions and short-range control by proximal promoter elements. Interestingly, both mechanisms share identical families of homeodomain (Lhx2/Emx2) and Olf-1/Ebf-1 transcription factors. To better understand the contribution of proximal promoter sequences in OR gene expression, the regulation of the zebrafish OR101-1 gene was studied by transient transgenic promoter assays. The proximal 1.2 kb of sequence upstream of the OR101-1 coding region drives expression of fluorescence reporter proteins in zebrafish OSNs with high efficiency. To pinpoint positive regulatory sites within this sequence, a series of related transgenic constructs was generated in which specific sequences were mutated or deleted. The OR101-1 gene promoter appears to be rather compact. Only 212 bp ups. of OR101-1 TSS are sufficient to maintain high efficiency of transgene expression. Sites included in this sequence and resembling an O/E-like binding motif appear not to contribute to expression. Interestingly, however, removal of a 345 bp intron resulted in a 50% loss of efficiency of the promoter construct, though it remains unanswered whether an included O/E-like site or splicing *per se* contribute to expression.

ÖZET

OR101-1 GENİNİN YAKIN PROMOTÖR BÖLGESİNİN ZEBRABALIĞINDA *İN VİVO* KARAKTERİZASYONU

Omurgalı hayvanların burunlarında bulunan her bir koku algılayıcı sinir hücresi geniş gen repertuarından yalnızca bir tane koku reseptör (KR) geninin tek bir alelini ifade eder. Bu KR genlerinin sinir hücrelerinde tek olarak ifade edilmesi koku duyu sisteminde çok önemlidir çünkü bu reseptör proteinler hem sinir hücrelerinin koku moleküllerine olan duyarlılığını belirler hem de sinir hücresi aksonlarına beyinde uygun sinaptik bağlantılar kurması için yol gösterir. KR genlerinin ifade edilişini düzenleyen moleküler mekanizmalar henüz anlaşılmamıştır. Son beş yıl içinde KR genlerinin tekli ifade edilişini açıklamak için LSD1 enzimi aracılığıyla KR gen bölgelerinin epigenetik işaretlerinin değiştirilmesi üzerine temellenmiş bir mekanizma öne sürüldü. Bu modelde LSD1 baskılayıcı histon işaretlerini düşük olasılıkla değiştirerek sadece tek bir KR geninin hücrelerde ifade edilişini sağlamaktadır. Fakat tüm KR genlerinin aynı olasılıkla ve burun dokusunun aynı alanında ifade edilmemesi LSD1'nin KR gen bölgelerine "yönlendirildiğini" öne sürmektedir. Şu ana kadar iki tip KR gen-bölgesi merkezli kontrol mekanizması teklif edildi; uzaktaki *cis*-aktiviteli bölge kontrol kısımları tarafından uzun menzilli kontrol ve yakın promotör bölgesi unsurlarıyla yakın menzilli kontrol. Yakın promotör sekansının KR geni ifade edilişine katkısını daha iyi anlamak için zebrabalığının OR101-1 geninin ifade edilişinin düzenlenmesi geçici transgenik promotör tahlili ile çalışıldı. OR101-1'in kodlayan bölgesinin yukarısındaki 1.2 kb'lik kısmı zebrabalığı koku sinir hücrelerinde florasan proteinlerin ifade edilişini yüksek verimlilikle tetiklemektedir. Bu sekanstaki pozitif düzenleyici bölgelerin yerini tespit etmek için özel kısımların silindiği ya da mutasyona uğratıldığı çeşitli transgenik yapılar oluşturuldu. OR101-1 promotörünün oldukça sıkı olduğu ve transkripsiyon başlangıç noktasının yukarısındaki ilk 212 bazlık kısmın florasan proteinin yüksek verimlilikte ifade edilişi için yeterli olduğu bulunmuştur. Bu sekanstaki Olf1/Ebf1 proteinini bağlayıcı motife benzeyen kısmın ifade edilişe katkısı bulunmamıştır. Diğer yandan 345 bazlık intronun silinmesi promotör aktivitesinde yüzde 50'lik bir düşüşe sebep olmuştur. Bunun esas sebebi araştırılmaktadır.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv
ABSTRACT iv
ÖZET v
LIST OF FIGURES viii
LIST OF TABLES x
LIST OF ACRONYMS / ABBREVIATIONS xi
1. INTRODUCTION 1
1.1. The Vertebrate Olfactory System 1
1.2. Odorant Receptors 4
1.3. Odorant Receptor Gene Expression 6
1.3.1. Monogenic and Monoallelic Expression 6
1.3.2. Zonal Expression of OR genes 8
1.3.3. Regulation of Singular Expression of OR genes 9
1.3.3.1. Regulation by an Epigenetic Silencing Mechanism 9
1.3.3.2. Regulation by Locus Control Regions 11
1.3.3.3. Regulation by Proximal Elements 13
1.3.3.4. Negative Feedback Regulation. 16
2. PURPOSE 20
3. MATERIALS AND METHODS 21
3.1. Materials 21
3.1.1. Fish 21
3.1.2. Equipment and Supplies 21
3.1.3. Buffers and Solutions 21
3.2. Methods 22
3.2.1. Fish Maintenance 22
3.2.2. Microinjection into Zebrafish Oocytes 22
3.2.3. Polymerase Chain Reaction (PCR) 22
3.2.4. Restriction Endonuclease Digests of DNA 23
3.2.5. Agarose Gel Electrophoresis 23
3.2.6. Gel Extraction 24

3.2.7. PCR Purification 24
3.2.8. Ligation of DNA Fragments into Vectors 24
3.2.9. Preparation of Competent Cells for Transformation 25
3.2.10. Transformation of Plasmid DNA into Competent Cells 25
3.2.11. Plasmid Isolation 26
3.2.12. Imaging of Zebrafish Embryos 26
3.2.13. Bioinformatics Analysis of Genomic Sequences 26
4. RESULTS 30
4.1. Zebrafish OR 101-1 Gene 30
4.2. Injection of Transgenic Constructs 31
4.2.1. Analysis of 1.2 kb Upstream Genomic Sequence 33
4.2.2. Analysis of Sequences Upstream of Transcription Start Site 34
4.2.2.1. Analysis of Transgene Expressing Cells in Embryos 37
4.2.3. Bioinformatics Analysis of Sequences Upstream of TSS 39
4.2.3.1. Analysis of Mutations within the proximal O/E Site 42
4.2.4. Analysis of Sequences Downstream of Transcription Start Site 44
4.2.4.1. Analysis of Transgene Expressing Cells in Embryos. 45
4.2.4.2. Bioinformatics Analysis of 5'-Intronic Sequences. 46
4.2.5. Analysis of Expression with Olf1/Ebf1 Gene Knock Down 50
5. DISCUSSION 57
5.1. Short-range vs long-range control of a zebrafish OR gene 57
5.2. The basic OR101-1 promoter is contained within 212 bp upstream of the TSS 58
5.3. The OR101-1 core promoter contains O/E and homeodomain binding motifs 59
5.4. The proximal O/E site upstream of the TSS is dispensable for promoter function. 60
5.5. The first intron of OR101-1 is critical for gene expression 61
5.6. Regulatory sites in introns or regulation by intron splicing? 63
5.7. No effect of O/E sites or no effect of Olf1/Ebf1? 64
APPENDIX A: EQUIPMENT 66
APPENDIX B: SUPPLIES 68
APPENDIX C: RESULTS FOR INDIVIDUAL EXPERIMENTS 71
REFERENCES 86

LIST OF FIGURES

Figure 1.1. An illustration of neuronal wiring in vertebrate olfactory system. 2

Figure 3.1. General scheme of injection experiments. 29

Figure 4.1. Phlogenetic relationship among zebrafish and human odorant receptors. 31

Figure 4.2. Transgene expression analysis of 1.2 kb promoter of OR101-1 gene. 32

Figure 4.3. Expression analysis of the $p_{1.2kb}$OR101-1::EYFP-pA construct. 34

Figure 4.4. Expression analysis of the $p_{1.0kb}$OR101-1::EYFP-pA construct. 35

Figure 4.5. Expression analysis of the $p_{0.8kb}$OR101-1::EYFP-pA construct. 36

Figure 4.6. Motif-based analysis of genomic upstream sequences of OR101-1 TSS. 41

Figure 4.7. Expression analysis of the $p_{1.2kb\text{-}\Delta O/E}$OR101-1::EYFP-pA construct. 43

Figure 4.8. Expression analysis of the $p_{0.6kb\text{-}\Delta UTR\text{-}\Delta intron}$OR101-1::EYFP-pA construct. ...47

Figure 4.9. Expression analysis of the $p_{0.8kb\text{-}\Delta intron}$OR101-1::EGFP-pA construct. 48

Figure 4.10. Motif-based analysis of downstream sequences of OR101-1 TSS. 49

Figure 4.11. Scan of the 1.2 kb promoter region for the motifs representing O/E-binding sites identified in human (green), murine (blue) and fish (red). 50

Figure 4.12. Expression analysis of the $p_{1.2kb}$OR101-1::EYFP-pA construct with co-injection of Ebf1b translation blocking morpholino. 53

Figure 4.13. Summary of expression analysis of transgenic constructs throughout Section 4.2. 54

Figure 4.14. Summary of cell number analysis of transgenic constructs throughout Section 4.2. 55

Figure 4.15. Summary of alternative cell number analysis of transgenic constructs throughout Section 4.2. 56

LIST OF TABLES

Table 3.1. List of the primers used in this study. 28

Table 3.2. List of the morpholinos used in this study. 28

Table 4.1. Expression levels of Ebf1a and Ebf1b genes in olfactory tissue at log_2 scale, which were normalized to brain tissue. 51

Table 6.1. Equipment. 66

Table 6.2. List of Supplies. 68

Table 6.3. Number of Transgenic Cells. 71

LIST OF ACRONYMS / ABBREVIATIONS

BAC	Bacterial Artificial Chromosome
bp	Base Pair
cAMP	Cyclic Adenosinemonophosphate
DNA	Deoxyribonucleic Acid
GFP	Green Fluorescent Protein
GPCR	G-protein-coupled Receptor
kb	Kilobase Pair
mRNA	Messenger Ribonucleic Acid
OB	Olfactory Bulb
OE	Olfactory Epithelium
OMP	Olfactory Marker Protein
OR	Odorant Receptor
OSN	Olfactory Sensory Neuron
PCR	Polymerase Chain Reaction
RNA	Ribonucleic Acid
UTR	Untranslated Region

1. INTRODUCTION

1.1 The Vertebrate Olfactory System

The sense of smell is mediated by the olfactory system, which, in vertebrates, employs specialized chemosensory neurons in the nasal cavity to detect odorants and pheromones. Information about the nature of the external stimulus is then transmitted to the central nervous system where it can elicit important endocrine or behavioral responses, such as food finding, social interaction, predator avoidance, or mating.

In the vertebrate olfactory system, the detection of odorants occurs in the main olfactory epithelium (OE), while detection of pheromones is mediated by the vomeronasal organ (VNO). Olfactory sensory neurons (OSNs) of the main OE interact with the inhaled odorant molecules through odorant receptors (ORs), which are concentrated in OSN cilia that protrude into the nasal cavity. As a general rule, each OSN expresses only one OR gene, despite the presence of a very large genomic repertoire of OR genes (Chess *et al.*, 1994; Malnic *et al.*, 1999). This fundamental principle, the 'one neuron - one receptor' rule, restricts the receptive range of each OSN to that of the expressed OR. Information about a particular odor is then encoded by the specific combination of OSN populations that are activated by the odor stimulus. This virtually limitless combinatorial system also constitutes the biological basis for the astonishing odor discrimination abilities of certain species.

OSNs extend long and unbranched axons that project to the main olfactory bulb (OB) of the forebrain. The OB is the first relay station of olfactory information. Interestingly, OSNs expressing the same OR gene project their axons to the same glomeruli, which often have invariant positions. Glomeruli are discrete bundles of neuropil where OSN axons form synapse with projection neurons, such as mitral and tufted cells, that project to the olfactory cortex and higher brain centers and local interneurons (Figure 1.1; Buck, 2000).

As a result of the connectivity between OSNs in the nose and glomeruli in the OB, the identity of an odorant is now represented by the combination of activated glomeruli at the level of the OB (Mombaerts *et al.*, 1996; Wang *et al.* 1998; Buck, 2004).

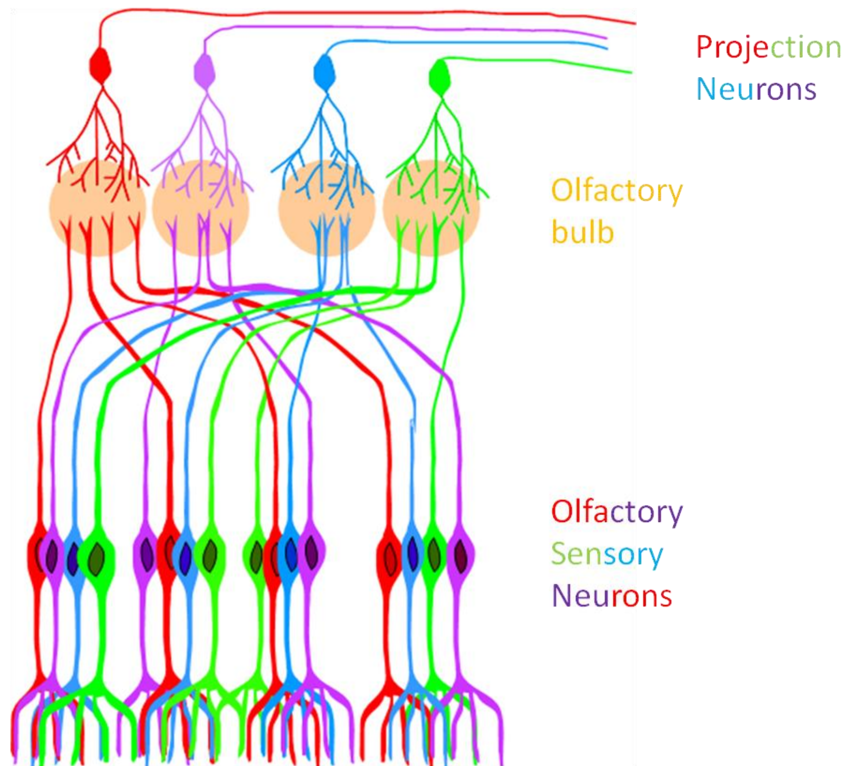


Figure 1.1. An illustration of neuronal wiring in vertebrate olfactory system. Four different populations of OSNs, each expressing a different OR, are represented by different colors. OSNs expressing same ORs project their axons to the same glomeruli in the Olfactory Bulb.

In certain vertebrates, a second chemosensory tissue, the VNO, is specialized in the detection of pheromones and kairomones, essentially signaling molecules that mediate intra- or interspecies communication, respectively (Firestein, 2001). Different from OSNs, which project to the main OB, sensory neurons of the VNO project their axons to a defined subregion called the accessory OB (Belluscio *et al.*, 1999; Rodriguez *et al.*, 1999). The vomeronasal receptor neurons express the V1R and V2R chemoreceptor genes depending on their position within the VNO tissue (Dulac and Axel, 1995; Herrada and

Dulac, 1997). Axonal projections of VNO neurons are not as strict as OSN projections and VNO cells expressing the same VR gene converge onto multiple glomeruli in the accessory OB (Mombaerts *et al.*, 1996; Wang *et al.*, 1998).

The two basic principles of the olfactory system, the one neuron – one receptor rule and the innervation of glomeruli by same axons, most likely also apply to zebrafish (Barth *et al.*, 1997; Sato *et al.*, 2007). However, certain differences between the olfactory systems of fish and mouse exist. For instance, there is no anatomical segregation of odor and pheromone recognizing systems, i.e. between the main OE and the VNO, and both types of signals are detected by sensory neurons located in the same olfactory tissue. The single rosette-shaped zebrafish OE thus contains different types of OSNs that resemble sensory neurons of the main OE and VNO. Four distinct types of chemosensory neurons have been identified in the zebrafish OE so far. The subtypes differ in their overall morphology, their apical to basal positions in the OE, as well as in their expression of chemosensory receptors and molecular marker proteins (see Yoshihara, 2009 for a review).

The ciliated OSNs, which are located at the basal layer of OE, possess long dendrites and numerous cilia are similar to OSNs of the main OE in mammals. They express either classical OR or trace amine-associated receptor (TAAR) genes in addition to the universal marker olfactory marker protein (OMP; Celik *et al.*, 2002; Sato *et al.*, 2005). The second major type of zebrafish OSNs, the microvillous neurons, are similar to VNO neurons. They are located more apically in the OE, bear short dendrites that terminate in a microvillous knob (Hansen *et al.*, 2003), and express the orthologues of mammalian V2R genes along with the signal transduction component TrpC2 (Hansen *et al.*, 2004; Sato *et al.*, 2005). Interestingly, zebrafish microvillus OSNs were reported to respond to food odors, thus blurring the functional distinction between the main OE and the VNO as odor and pheromone detection systems, respectively (Sato *et al.*, 2005). In addition to these two types of OSNs, less abundant cell populations have been described with the crypt and kappe cells. The globose Crypt cells are found in the most apical layer of the OE, they possess both cilia and microvillae (Hansen and Zeiske, 1998) and it was shown that one member of V1R-like genes; ORA4 is expressed in almost all Crypt cells (Oka *et al.*, 2011).

Very recently, a fourth sensory neuron population was identified in zebrafish OE (Ahuja *et al.*, 2014) with the Kappe neurons by means of their G_o-like immunoreactivity. They resemble Crypt neurons but exhibit a different spatial distribution in the OE and target a different glomerulus in the OB.

1.2. Odorant Receptors

The multigene family that encodes ORs was discovered by Linda Buck and Richard Axel in 1991 (Buck and Axel, 1991) via a cloning strategy based on the presumption that ORs were G protein-coupled receptors (GPCR). The assumption was founded on physiological studies with odorant-activated OSNs, the identification of the olfactory-specific G protein subunit $G_\alpha$olf (Jones and Reed, 1989) and a cAMP-gated ion channel (Dhallan *et al.*, 1990).

Vertebrate ORs belong to the group of class A GPCRs. A typical feature of vertebrate ORs is that they have intronless coding sequences and show extreme amino acid sequence diversity, especially in transmembrane regions that are supposed to form the odorant binding pocket.

The mouse OR gene repertoire consists of 1.375 genes, which constitutes by far the largest coherent gene family in the genome (Zhang *et al.*, 2007) and comprises about 3% of all genes. OR genes are distributed throughout 43 different loci on all chromosomes except chromosomes 12 and Y. A total of 1.194 OR genes, or 87 %, have intact open reading frames and encode functional proteins, while the remaining 13% of genes are pseudogenes. Comparable numbers were observed for the rat OR genes (Zhang *et al.*, 2007), yet, the size of the repertoire and the proportion of functional genes vary across species and most likely reflect their evolutionary history and ecological needs. For instance, the human olfactory genome contains 802 OR genes, more than half of which (415) are pseudogenes (Nimura and Nei, 2007). Similar to the mouse, human OR genes are distributed across almost all chromosomes, except chromosome 20 and Y (Glusman *et al.*,

2001). Interestingly, dogs, reputed with their very good sense of smell, have less number of functional OR genes (811) than rodents (Nimura and Nei, 2007). Thus, the difference in the number of intact OR genes does not strictly reflect the olfactory abilities of the given species, suggesting the existence of additional factors (see Malnic *et al.*, 2010 for a review).

Vertebrate ORs can be classified into two large groups, namely the class I and class II ORs; a distinction based on amino acid sequences (Glusman *et al.*, 2001; Zhang and Firestein, 2002). Class I ORs are supposed to be evolutionary more ancient and show stronger sequence similarity with the ORs originally identified in fish (Ngai *et al.*, 1993) and are responsive to water soluble odorants (Malnic *et al.*, 1999). These “fish-like” ORs, are not just an evolutionary relic in mammalian genomes, but constitute about 10 % of the functional human or mouse OR repertoire. Interestingly, class I ORs are localized to one large cluster on a single chromosome in both species (Glusman *et al.*, 2001; Zhang *et al.*, 2004). Furthermore, most of these class I ORs have intact open reading frames and so are functionally involved in mammalian olfaction. In contrast to “fish-like” class I ORs, however, class II ORs are specific to terrestrial vertebrates and constitute the majority of mammalian OR repertoire, with high responsivity to volatile odorants.

In contrast to the large OR repertoire of mammals, the zebrafish genome contains only 179 OR genes, 143 of which are potentially intact. The majority of zebrafish OR genes are restricted to five large clusters on three chromosomes. Despite its limited size, the zebrafish OR repertoire can be subdivided into eight different gene families based on amino acid sequence similarity. When zebrafish OR genes are compared to mouse or human OR genes, they almost exclusively cluster with class I ORs, tightly or distantly. Interestingly, a single zebrafish OR gene, OR 101-1 shows high sequence similarity with the “mammalian-like” class II ORs of mouse (Alioto and Ngai, 2005).

1.3. Odorant Receptor Gene Expression

1.3.1. Monogenic and Monoallelic Expression

As a fundamental principle in the vertebrate olfactory system, each mature OSN expresses only one functional OR gene from the entire genomic repertoire. This principle is often referred to as the "one neuron - one receptor" rule, although conclusive proof for the universal validity of the rule is hard to provide because of the vast dimension of the OR repertoire and the even greater number of possible combinations of OR coexpression. One of the studies providing more direct evidence for the rule was conducted by Malnic *et al* (1999). In this study, murine OSNs were isolated from the MOE and single-cell RT-PCR using degenerate primers was performed to amplify OR fragments. When the relevant OR PCR products were sequenced, by and large only one OR cDNA sequence could be read per individual OSN, suggesting that only one OR gene is expressed by these OSNs.

The mutually exclusive expression of mammalian OR genes is also supported by double in situ hybridization (ISH) assays performed for various murine OR genes to examine their expression patterns in the OE. In one of these studies, the expression patterns in the OE were analyzed for three different members of the OR37 gene family. Kubick *et al*. (1997) determined the number of neurons expressing each of the individual receptors by hybridizing sections of the main OE with receptor-specific probes or a degenerate probe corresponding to the conserved coding region of all subtypes. Consequently, the number of neurons which were visualized by the hybridization with the degenerate probe correlated very well with the sum of the neurons being visualized by the signals obtained from specific subtype probes. Similar results were also obtained in another study analyzing the expression patterns of three adjacent murine OR genes on the same cluster, in which no co-expressing OSN could be observed for these ORs in ISH assays (Tsuboi *et al.*, 1999). Although the comprehensive double ISH analysis for all combinations of OR genes is not practical due to the large size of mammalian OR repertoire, the results which were obtained so far in the expression pattern analyzes of the selected OR genes are highly consistent with the one neuron – one receptor rule.

Another feature of mammalian OR gene expression is that each selected OR gene is transcribed only from the maternal or the paternal allele. This was first shown in a study in which OR transcripts obtained by single-cell RT-PCR were assayed for allelic polymorphisms (Chess *et al*., 1994). It was found that OR transcipts from a single OSN did not show any sequence polymorphisms and were derived from only one of the parental alleles. In contrast, different cells would express either one of the two alleles. Monoallelic expression of OR genes was also confirmed in gene-targeted mice (Strotmann *et al.*, 2000) in which both alleles of the OR gene were tagged with different reporter genes. Among thousands of OSNs, which could be visualized, none were detected to express both markers, even though both markers were expressed by distinct OSN subpopulations. Lastly, fluorescent in situ hybridization (FISH) experiments directly demonstrated monoallelic activation of a single OR allele within OSN nuclei (Ishii *et al.*, 2001). Only one of the two genomic loci of the OR detected by DNA-FISH showed the cognate transcript by RNA-FISH.

Mutually exclusive expression of OR genes has also been shown to exist between endogenous and transgenic alleles (Serizawa *et al.*, 2000). Two versions of a murine OR gene (MOR28), endogenous and transgenic, were tagged with different markers. Almost no neurons were found to simultaneously express both reporters. More surprisingly, allelic exclusion was also observed between two transgenic alleles of the same gene (Serizawa *et al.*, 2000). These results demonstrated that mutually exclusive activation of OR genes is not restricted to endogenous OR gene loci but is even applicable for transgenic loci if they sufficiently resemble an OR gene locus.

In zebrafish, a comprehensive ISH analysis showed that typically an OR gene is expressed in a small population of OSNs scattered throughout the OE. OR-specific riboprobes detect between 0.5 - 2.5 % of OSNs. When compared to the total of ~140 intact OR genes in the zebrafish genome, these results supported the notion that the one neuron – one receptor rule may also be conserved in fish (Weth *et al*., 1996; Barth *et al*., 1996, 1997; Sato *et al*., 2007). More directly, double ISH experiments testing different

combinations of OR genes mostly resulted in the labeling of non-overlapping OSN populations (Barth *et al.*, 1997), with two exceptional cases (Sato *et al.*, 2007).

In addition to monogenic expression, the second basic principle of mammalian olfactory system, - the convergence of like axons to the target glomeruli (Mombaerts *et al.*, 1996), is also applicable to zebrafish. Although it is not yet elucidated whether there are specific glomeruli for specific ORs in zebrafish like in mammals, OSNs which express the OR genes belonging to particular OR subfamilies were observed to project their axons to a topographically fixed glomerular cluster in the OB (Sato *et al.*, 2007).

1.3.2. Zonal Expression of OR genes

Expression of each OR gene is restricted to a defined subregion of the OE, often referred to as a zone. Initially it was believed that there are four distinct expression zones with limited overlap in mouse and rat (Ressler *et al.*, 1993; Vassar *et al.*, 1993). The expression domains are organized along the dorsal-ventral axis of the OE but OSNs expressing the same OR appear to be randomly distributed within a domain. Expression patterns that resemble the zonal organization of the rodent OE were observed in the form of concentric domains in zebrafish (Weth *et al.*, 1996).

Several studies showed that most mouse class I OR genes are expressed in zone 1 (the most dorso-medial zone; Zhang *et al.*, 2004; Tsuboi *et al.*, 2006). As different from class I ORs, two studies with mouse class II OR genes revealed that their expression is mainly restricted to zone 2-4 (Zhang *et al.*, 2004; Miyamichi *et al.*, 2005). However, the latter study also showed that the segregation of OR gene expression into four conventional zone is not applicable to every OR gene (Miyamichi *et al.*, 2005). The expression pattern analysis of 80 class II OR genes in OE demonstrated that the expression areas are specific to each OR genes and exhibit considerable overlap in between.

Interestingly, zonal segregation of OR gene expression in the OE is also seen in the olfactory bulb by the axonal projection of OSNs. Thus, there is a zone to zone relationship between the OE and the OB, where axons of OSNs from one expression zone converge onto glomeruli in related dorsal/ventral positions in the OB (Miyamichi *et al.*, 2005; Tsuboi *et al.*, 2006). However, no apparent correlation was found between anterior-posterior positions of glomeruli in the OB and the spatial distributions of OSNs along the anterior-posterior axis of the OE (Miyamichi *et al.*, 2005).

1.3.3. Regulation of OR gene expression

The unique mode of OR gene expression (i.e. monogenic, monoallelic, and zonal expression) raises the question as to how OR expression in OSNs is controlled. Only during the last 5 years a plausible mechanism to explain certain aspects of OR gene expression has emerged. In this model, a single OR gene locus is released from epigenetic silencing through low level activity of a lysine-specific demethylase (LSD1), which modulates repressive histone marks that are typical for olfactory chromatin in immature OSNs (Lyons *et al.*, 2013). While this mechanism can, in principle, explain the expression of a single OR allele per OSN, it, by itself, cannot explain differences in OR expression. As explained above, not all ORs are expressed with the same frequency and different ORs are expressed in different zones of the OE. Thus, additional mechanisms are required that preferentially guide the activity of LSD1 to some OR loci but not to others.

1.3.3.1. Regulation by an Epigenetic Silencing Mechanism. It was investigated whether chromatin mediated silencing, which is an effective form of transcriptional repression, could play a role in OR expression (Magklara *et al.*, 2011). Interestingly, OR gene loci were found to be highly deposited with molecular markers of constitutive heterochromatin, such as H3K9 and H4K20 trimethylation. These epigentic marks appear to be established in immature OSNs, i.e. before onset of OR transcription and suggest that epigenetic silencing of OR loci is the default state. Prior to these studies it was assumed that OR gene loci are transcriptionally silenced late and after an OSN has made a choice for expression of a particular OR. The presence of an OR protein-elicited negative feedback signal was

assumed that changes transcriptional competence of non-expressed OR loci (Serizawa *et al*., 2003; Shykind *et al*., 2004; Nguyen *et al*., 2007).

In contrast to widespread epigenetic silencing, actively expressed OR alleles wer found to be associated with reduced level of silent and increased levels of active marks, such as H3K4me3 which is typically found on active promoters. Based on these results, a revised model was proposed to account for stable expression of a single OR by an OSN. By a process involving a switch from H3K9 methylation to H3K4 methylation, the selected OR allele is released from the pool of silenced OR genes and is transcribed (Magklara *et al*., 2011). In a recent study, a lysine-specific histone demethylase, LSD1 was shown to be highly expressed in immature OSNs (i.e. before initiation of OR gene expression), which can catalyze demethylation of both H3K9 and H3K4 (Lyons *et al*., 2013). This would release the Locus from repression and allow for transcription of the OR gene. To prevent expression of more than one OR per OSN, expression of LSD1 needs to be downregulated as soon as a successful OR choice has occurred.

Consistent with this model, lack of LSD1 activity prior to OR choice caused general loss of OR expression and OSNs failed to mature and to project axons to the brain. In contrast, lack of LSD1 activity in mature OSNs and following OR gene choice did not result in any detectable effect in OR expression and axonal projection to the targeted glomeruli, suggesting that LSD1 activity is strictly required during the initial derepression process and around OR gene choice. How is LSD1 then downregulated to prevent derepression of more than one OR allele per OSN? Once an OR gene is expressed by an OSN, it was shown to induce the subsequent expression of adenyl cyclase 3 (Adcy3), which in turn promotes LSD1 downregulation. Interestingly, mice lacking Adcy3 expression showed increased OR gene switching, suggesting that timely downregulation of LSD1 is required to stabilize the expression of the selected OR gene. In addition, when transgenic LSD1 was ectopically expressed in mature OSNs, it also perturbed the stability of OR choice, further supporting the role of Adcy3 to stabilize OR transcription by downregulating LSD1 (Lyons *et al*., 2013).

Following the observation that LSD1 downregulation is regulated by Adcy3 whose expression is induced upon OR gene activation, it was proposed that LSD1 expression is the target of OR protein-elicited negative feedback signal (Dalton *et al.*, 2013). In this study, it was proposed that OR proteins, once expressed in immature OSNs, generate a feedback signal by activating the unfolded protein response (UPR) pathway in the endoplasmic reticulum (ER). When the protein folding machinery of the cell in the ER is overloaded with the newly translated proteins, the UPR is activated both to transiently inhibit global translation of proteins and to increase the folding capacity of ER by upregulating genes encoding chaperon proteins. By analogy, accumulation of OR proteins stimulates the UPR in immature OSNs, which results in the induction of Adcy3 expression through a undefined signaling pathway. Eventually, Adcy3 expression would relieve the UPR and stabilize expression of the selected OR gene by LSD downregulation.

As outlined above, this general mechanism cannot explain all aspects of OR gene expression. Different OR genes are expressed with vastly different frequencies (Khan *et al.*, 2011), ranging from 0.1 to 50% within a zone (Tian and Ma, 2004; Fuss *et al.*, 2013). The LSD1 model also does not account for zonal expression of OR genes. Thus, it is conceivable that additional signaling mechanisms 'guide' the LSD1 to specific subsets of OR genes in different parts of the OE and with different efficiencies. Two types of control mechanisms, long-range control by locus control regions and short-range control by proximal promoter elements have been shown to affect OR gene expression.

<u>1.3.3.2. Regulation by Locus Control Regions.</u> When expression of the mouse MOR28 gene cluster was studied by generating YAC-transgenic mice it was found that a DNA region 40 to 150 kb upstream of MOR28 gene contains a *cis*-acting regulatory element that is necessary for transgene expression (Serizawa *et al.*, 2000; 2003). Comparison of genomic sequences surrounding the MOR28 (mice) and HOR28 (human) gene clusters revealed a 2.1 kb region of sequence homology (denoted as H region), which is located 75 kb upstream of MOR28 gene. Deletion of the H region from the YAC constructs resulted in the abolishment of expression of all the seven genes from the adjacent MOR28 cluster. Reintroduction of H to the upstream end of the MOR28 cluster restored the expression of

all transgenes. Interestingly, repositioning of the H region closer to the MOR28 cluster resulted in a significant increase in the number of transgene expressing OSNs (Serizawa *et al*., 2003).

Based on analogy to regulation of the red-green opsin locus in photoreceptor cells (Nathans *et al*., 1989; Wang *et al*., 1992) a functional model was proposed in which a *cis*-acting locus control region (LCR) stochastically interacts with only one promoter site across its cluster and tehreby activates only a single OR gene from the cluster. It could also explain differences in the frequencies of OR expression across clustered OR genes, as the probability for interaction between the LCR and the OR promoter would depend on distance between the two loci (Dillon *et al*., 1997); a mechanism that is also compatible with recruiting LSD1 activity to specific OR promoters.

In a later study, the H region residing in chromosome 14 was also shown to associate with OR gene promoters located on other chromosomes using the chromosome conformation capture (CCC) technique (Lomvardas *et al*., 2006). A regulatory model was proposed in which H region interacts in *trans* with only one allele of a single OR gene in any chromosome and thereby activates the expression of only one OR gene in OSNs. However, the model was dismissed when it was shown that deletion of the H region only impacts expression of the MOR28 cluster (Fuss *et al*., 2007; Nishizumi *et al*., 2007). Mice with deleted H region display in their sensory neurons normal expression of all OR genes from the whole repertoire but only three genes directly downstream of H region out of all the seven genes within MOR28 cluster. Thus, the observed local effect of H region deletion on across its cluster demonstrates that it does not possess a *trans* activity but is only a *cis*-acting regulatory element. These studies also showed that the different OR genes from the MOR28 cluster exhibit different zonal expression patterns in the MOE. The proximal three OR genes are expressed in the ventral OE whereas the distal 4 genes are expressed in the dorsal OE, suggesting that the H element or any other LCR does control zonal expression.

In addition to H region, additional cis-acting regulatory elements that may function as LCR have been identified in the mouse (Khan *et al.*, 2011; Markenscoff-Papadimitriou et al., 2014). One such LCR, denoted as P element, was shown to regulate the probability of OR gene choice within P2 cluster Khan et al., 2011). Similar to mice, genomic regions with LCR properties were also found in a zebrafish OR cluster on chromosome 15 (E-15-1 and E15-2 regions on chr. 15; Nishizumi *et al.*, 2007). Interestingly, the activity of LCRs is highly conserved across species and mouse LCRs boost expression in zebrafish OSNs (Nishizumi *et al.*, 2007; Tastekin, 2012; Uzel, 2014; Markenscoff-Papadimitriou *et al.*, 2014).

1.3.3.3. Regulation by Proximal Elements. In addition to long range control by LCRs or related elements, some OR genes have been reported to be controlled by rather short sequences surrounding their transcription start site. In many cases short promoter constructs faithfully mimic expression of the endogenous gene and exhibit similar zonal expression patterns, suggesting that proximal elements might also be involved in the regulation of singular and zonal expression of OR genes (Qasba and Reed, 1998; Vassalli *et al.*, 2002; Rothman *et al.*, 2005). For instance, ~9 kb genomic segments of two mouse OR genes- MOR23 and M71, could reproduce the zonal expression pattern of their endogenous counterparts and axonal projections to the cognate glomeruli. By successive deletions, this ~9kb transgene of MOR23 could be shortened up to 2.2 kb sequence containing only 395 bp upstream of transcription start site, which could maintain in most transgenic lines the zonal expression pattern of its endogenous counterpart and mediate the axonal projection to the cognate glomeruli (Vassalli *et al.*, 2002). Similar results were also obtained for M71 gene; the transgenes containing 491 bp and even as short as 161 bp upstream of the transcription start displayed similar epithelial pattern and glomerular positions to the endogenous gene (Rothman *et al.*, 2005). The axonal innervation of same glomeruli by short transgene expressing OSNs suggested that these OSNs do not co-express any other endogenous OR genes which would result in the innervation of spatially different multiple glomeruli in the olfactory bulb. This was also supported by the combined expression pattern analysis of M71 transgene and the endogenous gene; their expressions were mutually exclusive in the main olfactory epithelia (Vassalli *et al.*, 2002).

The high efficiency of short DNA segments suggested a strong *cis*-regulatory influence by proximal promoter elements. The comparative analyzes of genomic upstream regions of murine and human OR genes demonstrated that they harbor Olf1/Ebf1 (O/E)- and homeodomain (HD)-like binding sites (Vassalli *et al*., 2002; Hoppe *et al*., 2003). A more comprehensive bioinformatic analysis revealed that O/E-($(Y)_3CA(R)_4$) and HD-like binding sites (TATTX) are virtually present in proximal promoter regions of almost all OR (Michaloski *et al*., 2006) and VR genes (Michaloski *et al*., 2012). Yet, O/E-like sites were also identified in the promoters of other genes ubiquitously expressed in OSN, such as OMP (Kudrycki *et al*., 1993), G_{olf}, type III adenyl cyclase(ACIII), and the olfactory-specific cyclic nucleotide activated ion channel I (OCNCI; Wang and Reed, 1993). Therefore, whether the function of these sites is to only restrict the expression to the olfactory epithelia or they might direct zonal expression of OR genes still remain to be elucidated. The specific arrangements of O/E-like sites in proximal promoters of class I OR genes suggested a functional correlation with zonal expression (Hoppe *et al*., 2006), since their expression is mostly restricted to zone I (most dorso-medial zone) in the olfactory epithelia (Zhang *et al.*, 2004; Tsuboi *et al.*, 2006). However, a comprehensive analysis of proximal promoters of 198 OR genes to reveal any correlation between the distributions of the O/E- and HD-like sites with the cognate zonal expression patterns failed, remaining the issue hypothetical (Michaloski *et al*., 2006).

In a more recent study, another comprehensive analysis with different computational approaches also did not succeed in identifying strong correlations between promoter motifs and zonal expressions of OR genes (Clowney *et al*., 2011). Interestingly, O/E- and HD-like sites were identified also in LCRs (H element; Hirota and Mombaerts, 2004; P element; Vassalli *et al*., 2011). Mutations in these sites resulted in loss of H element derived enhancer activity on transgene expression (Nishizumi *et al*., 2007). Conversely, multimerization of an HD-like site in P element resulted in a dramatic increase in expression of transgenes (Vassalli *et al*., 2011).

The most proximal O/E- and HD-like sites within the promoter region of mouse OR gene M71 were studied with a mutagenic approach, separately for the genomic and the

transgenic allele (Rothman *et al*., 2005). The mutations of either site did not lead to any dramatic effect on the expression of M71 transgene - only a decrease in number of expressing cells for HD site mutated transgene, but caused ventralization of the epithelial pattern of the transgene expressing OSNs. However, the combined mutations of both sites resulted in loss of transgene expression in the OSNs. When the same mutations were introduced to the promoter of the endogenous gene, they did not preclude expression completely but also resulted in a decrease of M71-expressing OSNs and a ventral shift in the epithelial pattern. These results suggested that these motifs might mainly be involved in the specification of zonal expression patterns of OR genes in the olfactory epithelia.

After identification of O/E- and HD-like binding sites in proximal promoters of OR genes, the proteins binding to these sites have been also analyzed in several studies for their roles to regulate OR gene expression. In an early study in which the O/E1 was knocked out from the genome, the mutant mice displayed normal morphology and gene expression in the olfactory epithelia (Lin and Grosschedl, 1995), which was interpreted as a result of possible redundancy and functional rescue by other O/E family members showing high sequence homology with O/E1 (Wang *et al*., 1997). Later, it was reported that both O/E2 and O/E3 knockout mice displayed defects in OSN-axonal projections to the dorsal olfactory bulb, which was also observed in O/E2/ O/E3 double heterozygous animals. However, no effect could be observed in the expression of tested OR genes (Wang *et al*., 2004). In addition to O/E proteins, some of the transcription factors binding to HD-like sites were also studied in the context of OR gene expression. Lhx2, LIM homeodomain protein was identified to bind the most proximal HD-like site in M71 OR gene promoter and mice with Lhx2 null mutation displayed defects in OSN development (Hirota and Mombaerts, 2004). Additionally, these mice abolished expression of class II OR genes in their OSNs, although expression of class I OR genes remained intact (Hirota *et al*., 2007). However, the high embryonic lethality of Lhx2-deficient mice precludes insight as to whether Lhx2 plays a general role in OSN maturation or is more specifically involved in OR gene expression, even though its uniform expression across the OE does not suggest a direct implication in zonal patterning of OR expression (Hirota and Mombaerts, 2004). Another transcription factor, Emx2 was also implicated in OR gene regulation (McIntyre *et al*. 2008). This homeobox protein was shown to bind HD-like sites

in OR promoters (Hirota and Mombaerts, 2004) and to be expressed in the OE (Nedelec *et al.*, 2004). In this study (McIntyre *et al.* 2008), Emx2 mutant mice showed a reduction in the number of mature OSNs and an altered OR gene expression profile. Interestingly, Emx2 deficiency had differential impact on expression of OR genes. The majority of OR genes were observed to be expressed in fewer OSNs and a few OR genes exhibited expression in more OSNs. Together with the observed reduction only in mature OSNs, these results suggest that Emx2 more directly functions in OR gene expression, than Lhx2 whose deficiency causes severe reduction both mature and immature OSNs, implying mainly a developmental function for it.

1.3.3.4. Negative Feedback Regulation. Once an OSN has made a successful OR gene choice it has to prevent activation of additional OR genes in the same cell. It was proposed that OSNs employ a negative feedback mechanism to achieve the mutually exclusive expression of OR genes (Serizawa *et al.*, 2003). When the coding sequence of MOR28 transgene was deleted, expression of other ORs was observed in the transgene expressing OSNs. Similarly, the introduction of frameshift mutations into an OR coding sequence also resulted in activation of secondary OR gene expression, suggesting that OR protein, once expressed in an OSN, have an inhibitory role in expression of additional OR genes (Serizawa *et al.*, 2003). This was further supported by the following studies (Lewcock and Reed, 2004; Shykind *et al.*, 2004). The activation of an OR promoter in the absence of an OR protein does not prevent the expression from additional OR promoters in the OSNs. However, when an intact OR coding region is introduced to the downstream, it restores the mutually exclusive expression. This suggests that a negative feedback signal is elicited by a functional OR protein, which abrogates expression of other OR genes (Lewcock and Reed, 2004). This OR protein-elicited feedback signal does not only function in stabilization of the selected OR gene-expression and restriction of additional OR gene activation but also performs a quality check (Shykind *et al.*, 2004; Rodriguez, 2013).

As previously mentioned, pseudogenes constitute a significant proportion of the OR genes in mammalian genomes (Zhang *et al.*, 2007; refer to Section 1.2, for detail). The stable expression of pseudogenes would result in lack of proper olfactory functioning in

the relevant OSNs. To avoid this possibility, it was proposed that by this feedback mechanism, an OSN which initially selected a pseudogene for OR expression would undergo a second choice, an event called “OR switching” (Shykind *et al*., 2004). By a lineage tracing experiment, it was shown that immature OSNs initially expressing a mutant MOR28 gene subsequently terminated its expression and switched at high frequencies to express another OR gene. Interestingly, this second OR choice could also be observed, though in low frequencies, within the wildtype MOR28 expressing immature OSNs which have not yet generated stable synaptic connections to the olfactory bulb. These observations suggested that OR switching is an inherent property of OR gene choice (Shykind *et al*., 2004).

Several studies analyzing the OSNs which initially express different mutant OR genes and so are presumed to undergo second OR choice showed that most OSNs displayed a restricted pattern of glomerular projection, suggesting that this second choice is not random among murine OR genes (Serizawa *et al*., 2003; Shykind *et al*., 2004; Feinstein and Mombaerts, 2004). Later, it was shown also to be class restricted (Bozza *et al*., 2009). OSNs expressing a nonfunctional class I OR gene projected their axons to the glomerular regions which are exclusively innervated by OSNs expressing class I OR genes. Similarly, OSNs expressing a nonfunctional class II OR gene projected their axons to class II specific glomerular domains in the olfactory bulb.

The restriction in second choice was also observed in zebrafish (Sato *et al*., 2007). In this study, a BAC transgenic line was generated, carrying an OR gene cluster in which the coding sequences of two OR genes were replaced with different fluorescent proteins. When the OSNs expressing these reporter genes were analyzed, it was found that their axons were projected to a topographically fixed cluster of glomeruli in the medial olfactory bulb. Further examination of these OSNs revealed that they mostly co-expressed the OR genes within the same subfamily of the cluster. Basing on these findings, a model of hierarchical regulation was suggested to be present in zebrafish OR gene choice. According to this model, an OR subfamily is initially selected and then an OR gene is expressed among the limited members of this subfamily. If the selected gene does not

encode a functional OR protein, then the second choice is also restricted to the limited members of this subfamily.

Although the exact nature of negative feedback regulation needs to be revealed, there is strong evidence for OR proteins to be strongly involved in this process (Nguyen *et al.*, 2007). It was observed that the forced expression of an OR transgene from a synthetic promoter (TetO) in OSNs is sufficient to repress the expression of endogenous OR genes, if this promoter is activated before the onset of endogenous OR expression, suggesting that OR protein has a direct role in preventing secondary OR activation. In contrast, the precedential expression of endogenous ORs were observed to suppress OR expression from the TetO promoter. This surprising result strongly suggested that OR coding sequence itself could be a target for the negative feedback signal to ensure monogenic expression (Nguyen *et al.*, 2007). However, although OR proteins appear to be strongly involved in negative feedback regulation, it is unlikely that they function in this process through G-protein mediated signaling (İmai *et al.*, 2006; Nguyen *et al.*, 2007). The mutation of G-protein activating site in OR transgenes did not result in co-expression of multiple OR genes. Recently, OR proteins have been proposed to generate this feedback, once expressed in OSNs, by activating Unfolded Protein Response which temporally alters the translational landscape of the relevant OSNs (Daltons *et al.*, 2013; for detail, refer to section 1.3.3.1).

With these recent and previous reports, a picture emerged about different layers of control that appear to co-regulate OR gene expression. The mechanisms employed include proximal elements within OR gene promoters, *cis*-acting LCRs to activate specific OR gene clusters, an OR protein-elicited negative feedback signal to prevent additional OR gene expression and epigenetic silencing and desilencing of selected OR gene loci. However, in zebrafish, the mechanisms regulating singular OR expression are not as clearly defined as in mice. Sato *et al*. (2007) proposed a model in which OR gene choice is hierarchically regulated in an OSN. Upon selection of an OR subfamily, expression of an OR gene is restricted to the members of this subfamily and this restriction also extends to the second choice if needed. However, the roles of additional components that might be

involved in this regulation are not well established, yet. One of them is the role of proximal elements in this regulation. To address this issue, in this study we focused on the proximal promoter of zebrafish OR101-1 gene.

2. PURPOSE

The aim of this study was to experimentally pinpoint positive regulatory sites within the highly efficient OR101-1 gene promoter. A1.2 kb sequence upstream of the OR101-1 gene coding frame has previously been shown to promote strong and specific expression in subsets of OSNs. The sequence comprises 600 bp of sequence upstream of the TSS, 196 bp of a 5'-noncoding exon, 345 bp of 5'-intronic sequence, and 47 bp of untranslated sequence at the beginning of exon 2. To test the contribution of these components to promoter activity, a variety of transgenic constructs was generated in which specific sequences have been mutated or excluded from the reporter construct. By scoring the transgenic efficiency of different variants it was aimed to identify and test sequences that may constitute transcription factor binding sites, which are critical for the regulation of OR101-1 expression.

3. MATERIALS AND METHODS

3.1. Materials

3.1.1. Fish

AB/AB and AB/TÜ strains of zebrafish were mostly used in experiments, which were obtained from the Zebrafish International Resource Center (ZIRC), at the University of Oregon, Eugene, USA. Wild type fish obtained from a local pet shop were also used during the project. Fish were grown and maintained at the Boğaziçi University Animal Facility.

3.1.2. Equipment and Supplies

The list of equipment and chemicals can be found in Appendix A and Appendix B.

3.1.3. Buffers and Solutions

Buffers and solutions for general molecular biology techniques were derived from manufacturers of molecular biology kits. Few unsupplied buffers were prepared according to Sambrook and Russell (1989). Solutions needed for zebrafish maintenance were prepared according to Westerfield (2007).

3.2. Methods

3.2.1. Fish Maintenance

Zebrafish are kept at 28ºC under 13/11 hours light/dark cycle. Adults are fed three times a day, with live brine shrimps at mornings/nights and with flake food at noon. To obtain eggs for microinjection experiments, an adult zebrafish male and a female are put in 1.5 L mating boxes with a separator between them in the evening. These boxes also include a container with holes through which fertilized eggs can infiltrate and accumulate at the bottom. Next day, separators are removed at the beginning of light cycle. For the case of injection, fertilized eggs are collected with a Pasteur pipette and lined up in an agarose injection mold. For other cases, eggs are transferred into an E3 medium including petri dish, which prepared according to Westerfield (2007).

3.2.2. Microinjection

Fertilized eggs are collected with a Pasteur pipette and lined up in an agarose injection mold. Glass capillary needle is filled with injection solution which contains 100 mM KCI, 0.01% Phenol Red, 50 ng/ µl of each plasmid. Needle is connected to the appendage of FemtoJet® Express pressure injector (Eppendorf) and injection is done with proper pressure and time conditions. Approximately, 4 nanoliter of the solution is injected into an embryo. Then, injected embryos are transferred into an E3 medium including petri dish (Westerfield, 2007).

3.2.3. Polymerase Chain Reaction

For Polymerase chain reactions (PCR), One Taq DNA Polymerase of NEB (New England Biolabs) was generally used according to manufacturer's instructions. PCR

reactions are carried out at 50 µl final volumes and include following reagents: 5X Standard Taq Buffer, dNTP (10 mM), 0.5 µM forward and 0.5 µM reverse primer, 50-100 ng template DNA and 1-3 units of Taq Polymerase. 1.5 mM $MgCl_2$ is also included if the used Taq Buffer does not contain per se (as in the case of Colony PCR). PCR reaction starts with 3 min. of initial denaturation step at 95 ºC. It continues with denaturation for 30 sec. at 95 ºC, annealing for 30.sec. at calculated temperature (usually the lowest Tm of the primers minus 1 ºC) and elongation for the calculated time interval according to the formula- 1 min./ 1 kb of target amplicon- at 72 ºC for 24 cycles. After the cycling, reaction ends up with 10 min. of elongation at 72 ºC and finally incubated at 4 ºC. At Colony PCR reactions, we used 10X Fermentas Taq Buffer and Home-made Taq polymerase and initial denaturation step was taken for 8 min. These reactions were also conducted for 34 cycles.

3.2.4. Restriction Endonuclease Digests of DNA

Restriction enzymes from Promega, New England Biolabs (NEB), Invitrogen and Fermentas were used at reactions. Generally, 1X buffers recommended and supplied by manufacturers are used at reactions but if not supplied, suggested NEB buffer was used instead. The used amount of enzyme was determined according to its units/ml concentration. 1-5 units of enzyme were usually applied per µg of DNA. Reactions were incubated at 37 ºC for 3 to 5 hours.

3.2.5. Agarose Gel Electrophoresis

DNA samples were run in 1 % Agarose gel which contains EtBr (0.5 µg/ml). 1 kb ladder of NEB was used as molecular weight marker. After running enough for intended band separation, gels were visualized under UV light and gel images were taken in TIF/JPEG format.

3.2.6. Gel Extraction

High Pure PCR Purification Kit (Roche, USA) was used to extract DNA fragments from 1% agarose gels after intended band separation obtained through electrophoresis. Under Ultra-Violet light, the band of interest is cut via a scalpel, trying to take the whole DNA fragment within a thin slice of agarose as much as possible. Then, the protocol supplied by the manufacturer was applied in the rest of extraction. Finally, DNA was eluted with either TE Elution Buffer or ddH_2O (only for injection plasmids).

3.2.7. PCR Purification

PCR products were purified via High Pure PCR Purification Kit (Roche, USA). Purifications were done according to the protocol supplied by manufacturer, except that following washing steps were taken for 10 min. and 5 min., respectively.

3.2.8. Ligation of DNA Fragments into Vectors

We used 1:3 molar ratio of vector to insert in all ligation reactions. If insert DNA was obtained from another vector by restriction enzyme digestion, it was directly ligated to the final vector plasmid. Concentrations of vector and insert were calculated via Nanodrop Spectrophotometer and their amounts in reaction did not exceed 100 ng, respectively. Ligation reactions are carried out at 20 µl final volumes and include insert and vector DNAs (1:3 in molarities), 1 µl of T4 DNA ligase (NEB), 10x Ligase Buffer (NEB) and dH_2O to achieve the final volume. Reaction mix is incubated at 22 ºC for 1 hour. Then 10 µl of mix is transferred into a fresh tube for transformation and rest is incubated o/n at16 ºC.

In cases where insert DNA was obtained from a PCR reaction, insert was firstly ligated into a pGEM-T Easy vector (Promega). This reaction is carried out at 10 µl final volume and includes 3.5 µl of PCR mix of insert (or purified PCR product), 2x Rapid Ligase Buffer (Promega) and dH_2O to complete the final volume. Reaction mix is incubated at 22 ºC for 1 hour and transformed into competent cells.

3.2.9. Preparation of Competent Cells for Transformation

Top10 MRF' bacteria strain was used to amplify plasmid vectors. Rubidium Chloride Method was applied to render bacteria competent for transformation. Basically, a single colony of the Top10 MRF' was picked and inoculated o/n at 37 ºC in 5 ml LB. Then, we used 500 µl of this culture to inoculate 500 ml of fresh LB medium and incubated at 37ºC on a shaker until the desired optical density (0.6) at 550 nm was reached. Chilling the culture on ice for 15 min. and centrifuging at 3000 rpm for 10 min. at 4 ºC, supernatant was removed and pellet was gently resuspended. We added 500 µl CT1 solution (30 mM KCH_3COO, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM RbCl and 15% glycerol) and mixture was incubated on ice for 30 min. After another centrifugation at 3000 rpm for 10 min. at 4 ºC, supernatant was removed and pellet was resuspended at 20 µl of CT2 solution (10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl and 15% glycerol). Finally, 50 µl aliquots of bacteria mixture were prepared and immediately shock frozen in liquid nitrogen. These aliquots were stored in -80 ºC until use for transformation.

3.2.10. Transformation of Plasmid DNA into Competent Cells

50 µl aliquots of competent cells being incubated at -80 ºC were taken and immediately put on ice and thawed for 5 min. Vector plasmid (50-100 ng) or 10 µl of ligation reaction was mixed with competent cells and resuspended. After incubation on ice for 30 minutes, mixture was heat shocked on a pre-adjusted heat block at 42 ºC for 90 seconds and then moved on ice for 5 min. Following heat shock, 300 µl LB was added and

the cells were left on a shaker at 37 ºC for 1 hour to recovery (not needed for plasmids). All the mixture was spreaded on a proper antibiotic resistant selection plate.

3.2.11. Plasmid Isolation

Plasmid Mini Gene Isolation Kit (Thermo Scientific) was used for minipreps. QIAgen MaxiPrep Kit was used for maxipreps.

3.2.12. Imaging of Zebrafish Embryos

Zebrafish embryos were anesthetized with 0.04 % MS222 (Sigma, USA) at 3 dpf. To visualize the embryos, two different methods were used in different experiments. At first method, anesthetized zebrafish embryos were lined within MS222 containing water droplets on a transparent plate lid and observed individually under SP5-AOBS laser scanning confocal microscope (Leica, Germany). In order to take images, anesthetized embryos were mounted in low-melting agarose (2%) and covered with a coverslip. Mounted embryos were then imaged under confocal microscope. At second method, we used an agarose (2 %) mold-developed in our lab, which includes 0.7 mm width of holes (equal to approximate width of a 3dpf zebrafish embryo). Zebrafish embryos were put into these holes in the proper way to directly expose the noses to light of the objective through the coverslip. This versatile mold was used for both visualizing and imaging.

3.2.13. Bioinformatics Analysis of Genomic Sequences

In the motif-scan of 2.0 kb genomic sequences upstream of OR101-1 TSS, RSAT servers were used (http://rsat.sb-roscoff.fr/RSAT_portal.html). In RSAT, oligo-analysis and matrix-scan tools were used according to the instructions provided in the website tutorials. In the analysis, Danio Rerio was chosen as the background model. Position

Weight Matrix (PWM) for Ebf1 motif was obtained from the JASPAR database (http://jaspar.genereg.net). PWM for Lhx2 motif was obtained from UniProbe database (http://uniprobe.org). In these analyses, statistical threshold value was always set “1” as default.

Same tools were also used in the scan of 1.2 kb genomic sequences upstream of OR101-1 coding sequence for Ebf1 motif. However, in addition to human and murine PWMs obtained from the JASPAR database, a fish-specific PWM for Ebf1 motif was also used, which was formed by phlogenetic footprinting performed on the putative OMP-promoter regions from 14 different fish species.

Table 3.1. List of the primers used in this study.

Primer	**Sequence 5' to 3'**
101-1_1.2kb_BamHI-Fw	GGATCCCATTGTGAAGAGAACATTC
101-1_1000_F	GCTACTGGGGAAAGGAAAGAGC
101-1_800_F	TTACCTGAACCTTACAACAGCC
101-1_nTSS_NcoI_R	CCATGGTATAGTATGTGTGTGTGTGCGC
101-1_ExonI_HindIII_R	AAGCTTATAACACAAGTCAATTCTGTTGTTTC
101-1_Exon2_HindIII_F	AAGCTTAGTGCATATAAGTTGTCCAGGTC
GFP_R_5'	AACTTGTGGCCGTTTACGTCGC
M13_F	CGCCAGGGTTTTCCCAGTCACGAC
M13_R	CAGGAAACAGCTATGACC
101-1_mO-E_F	GAG ATG TTAAAAGGGAAATCGCCTGCACAAC
101-1_mO-E_R	GAT TTC CCT TTT AAC ATC TCT CAA AGT GCA TCA TTA AGA GG

Table 3.2. List of the morpholinos used in this study.

Morpholino	**Sequence 5' to 3' according**
O/E-Transblocking-MO	GGATGCTCTCCTGAATCCCAAACAT
O/E-Spliceblocking-MO	TAGAAAAAGTCTCACCTGCACATGA

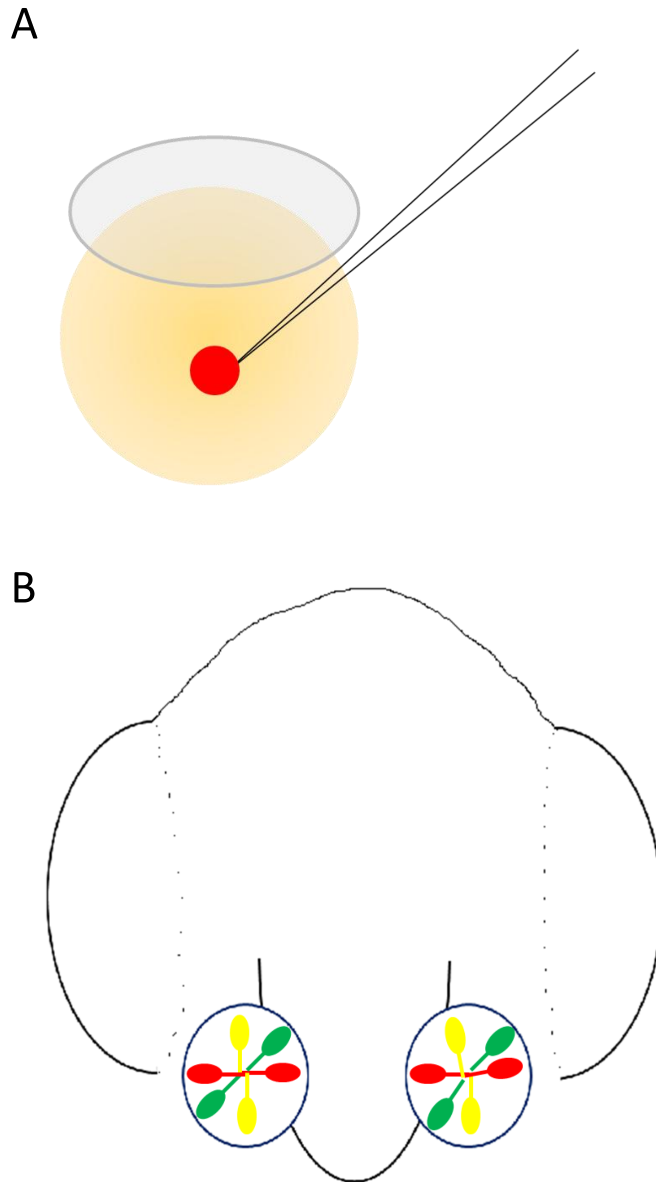


Figure 3.1. General scheme of injection experiments. (A) Injection of plasmid DNA into one-cell stage zebrafish embryo (Adapted from Kardash E., 2012). (B) Olfactory sensory neurons expressing the injected-plasmids are represented in different colours at right and left noses of fish at 3 dpf (Adapted from Yoshihara, 2009).

4. RESULTS

4.1. Zebrafish OR 101-1 Gene

The zebrafish genome contains approximately 179 Odorant Receptor (OR) genes (Alioto and Ngai, 2005). In contrast to higher vertebrates, such as mammals, whose OR genes can be classified into two broad classes - namely the class I and class II ORs - zebrafish OR genes can be subdivided into eight different gene families based on their amino acid sequence similarity (Alioto and Ngai, 2005). An alignment of zebrafish and mouse OR genes showed that most zebrafish ORs cluster tightly or distantly with the class I ORs of the mouse (Alioto and Ngai, 2005). Similar results were also obtained for the alignment of zebrafish and human OR genes (Tınaztepe, 2009). Interestingly, out of 179 zebrafish ORs, only a single gene, OR101-1, shows close sequence homology with human class II ORs. Typically, class II ORs expanded massively in the mammalian lineage and constitute up to 90 % of the OR repertoire in higher vertebrates (Alioto and Ngai, 2005, Tınaztepe, 2009, Figure 4.1). The OR101-1 gene is interesting, because its close alignment with mammalian class II ORs suggests that it may represent one of the first steps of evolutionary transition from class I to class II ORs in the teleost olfactory system.

In this study, we focused on the zebrafish OR 101-1 gene promoter to understand the role of proximal cis-regulatory regions that affect OR gene expression. Previously in our lab, a promoter bashing approach was employed to pinpoint regulatory sequences within the first 3.5 kb upstream of the OR101-1 coding sequence (CDS) using a transient transgenic assay. Among constructs of different length including various parts of the genomic upstream sequence, the first1.2 kb of genomic sequence upstream of the OR101-1 coding region was identified to drive robust and high frequency expression of fluorescent reporter proteins in zebrafish OSNs (61.3 %; Söğünmez, 2012). The 1.2kb construct comprises 600bp of genomic upstream sequence, 196 bp of a 5'-noncoding exon, 345 bp of 5'-intronic sequence, and 47 bp of untranslated sequence directly in front of the coding frame (Figure 4.2a-b). In the following experiments, we set out to further dissect the 1.2 kb

sequence by promoter bashing to identify positive regulatory sites and to understand the contribution of intronic and untranslated sequences to OR101-1 expression.

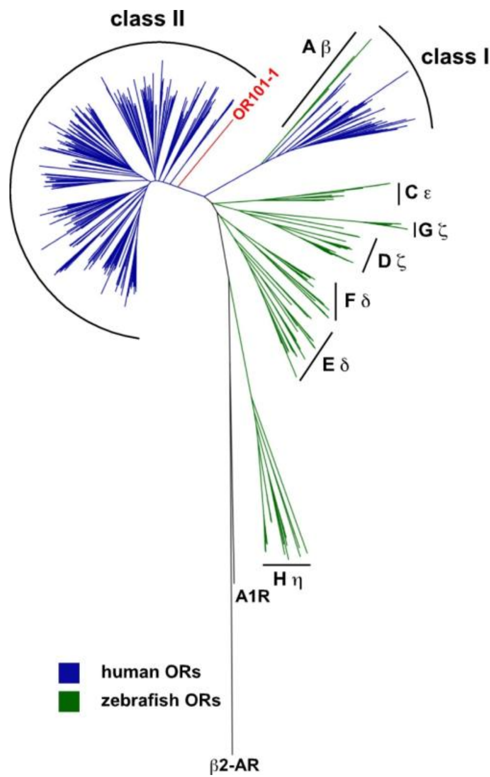


Figure 4.1. Phlogenetic relationship among zebrafish and human odorant receptors (Tınaztepe, 2009).

4.2. Injection of Transgenic Constructs

As a general design of the study, all transgenic constructs to be tested were co-injected with $p_{1.3kb}$OMP::mCherry-pA as a positive control for injection efficiency and normalization of results. The constructs were simultaneously co-injected at equal concentration into zebrafish oocytes at the one-cell stage. The $p_{1.3kb}$OMP::mCherry-pA vector includes 1.3 kb of genomic sequence upstream of the Olfactory Marker Protein(OMP) gene, and drives expression of fluorescent mCherry proteins in ciliated OSNs of zebrafish with high efficiency (Celik *et al.*, 2002; Sato *et al.*, 2005). The proximal

OMP promoter drives expression of fluorescent reporter proteins in up to % 90 percent of injected zebrafish oocytes at 3dpf, which makes it a robust and reliable control to assess injection efficiency (Söğünmez, 2012). Typically, the efficiency of expression of test constructs was quantified against the OMP promoter, either as the fraction of $p_{1.3kb}$OMP::mCherry-pA-expressing fish that express the test construct (penetrance) or as the number of EYFP-positive cells relative to the number of mCherry-positive cells per OE at 3 dpf.

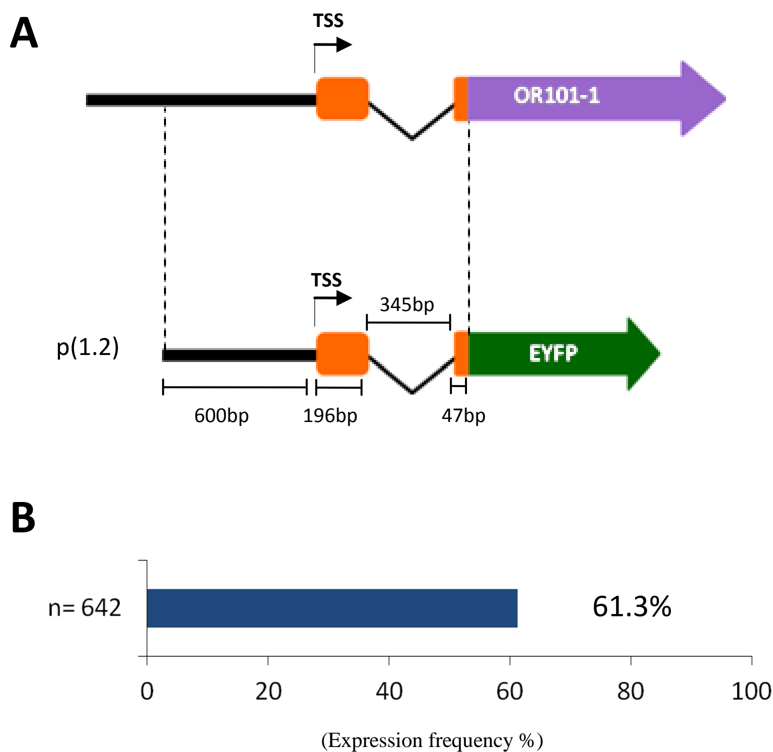


Figure 4.2. Transgene expression analysis of 1.2 kb promoter of OR101-1 gene. (A) Overview of the $p_{1.2kb}$OR101-1::EYFP-pA construct (B) Expression frequency of p(1.2)-positive zebrafish embryos at 3 dpf relative to $p_{1.3kb}$OMP::mCherry-pA-positive ones. n: mCherry (+) embryos (Söğünmez, 2012).

4.2.1. Analysis of 1.2 kb Upstream Genomic Sequence

As a first step to further dissect the 1.2kb OR101-1 promoter sequence, it was necessary to recapitulate the results obtained by Söğünmez (2012) to set the baseline for the analysis. The $p_{1.2kb}$OR101-1::EYFP-pA construct includes 1.2 kb upstream of the OR101-1 coding sequence followed by a sequence coding for the eYFP reporter gene and a poly-adenylation signal (polyA; Figure 4.3a).

When the $p_{1.2kb}$OR101-1::EYFP-pA construct was co-injected with the $p_{1.3kb}$OMP::mCherry-pA control plasmid into 355 fertilized zebrafish oocytes at the one-cell stage (four independent experiments: 84, 39, 135 and 97 embryos, respectively), 209 embryos (67, 23, 58 and 61, respectively) survived until the time of analysis at 3 dpf. 193 of the 209 survived embryos were further analyzed. Of those, 101 embryos were positive for mCherry and 81.1 % (82 of 101) expressed eYFP in OSNs (Figure 4.3c). Therefore, the relatively high expression frequency of this construct could be reproduced successfully, although a slightly higher efficiency was obtained. The high transgenic efficiency of the construct suggests that the 1.2 kb sequence upstream of OR101-1 includes strong positive regulators of OR101-1 expression.

This 1.2 kb promoter region includes 600 bp of genomic sequence upstream of the TSS of OR101-1 gene as well as downstream untranslated and intronic sequences. Similar results of expression by rather short sequences flanking the TSS of OR genes were also reported for OR111-7 in zebrafish (Mori *et al.,* 2000) and for MOR23 and MOR71 genes in the mouse (Vassalli *et al.,* 2002 and Rothman *et al.,* 2005), suggesting a strong role for these short sequences in regulation of vertebrate OR gene expression.

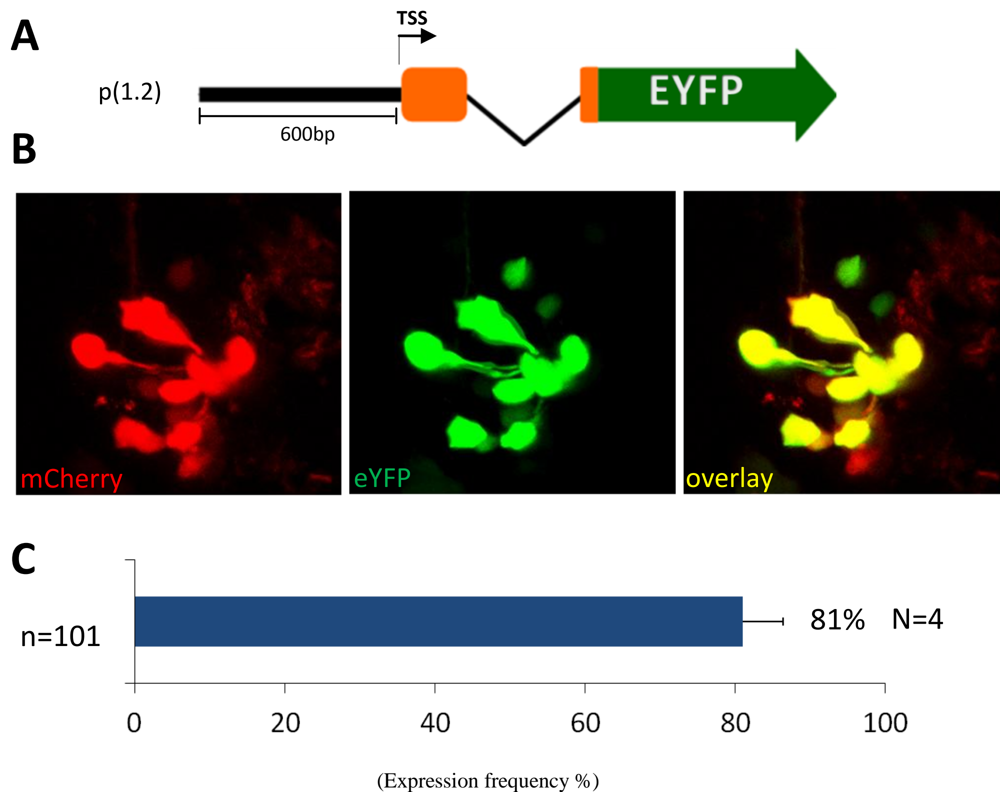


Figure 4.3. Expression analysis of the $p_{1.2kb}$OR101-1::EYFP-pA construct. (A) Overview of the construct. (B) Projection of a confocal z-stacks of a 3 dpf OE. (C) Expression frequency of the construct at 3 dpf relative to $p_{1.3kb}$OMP::mCherry-pA. n: mCherry (+) embryos. N: experiments.

4.2.2. Analysis of Sequences Upstream of Transcription Start Site

To assess the contribution of sequences upstream of the TSS to the high expression from the 1.2 kb promoter construct, the sequence was successively shortened from its distal end. In a first set of experiments, the first 206 bp from the 5'-end of the construct were removed. The newly generated construct, $p_{1.0kb}$OR101-1::EYFP-pA thus included 394 bp of sequence upstream of the TSS along with untranslated and intronic sequences (Figure 4.4a). The construct was then co-injected with the $p_{1.3kb}$OMP::mCherry-pA reference plasmid into one-cell stage zebrafish oocytes. 122 zebrafish embryos that

survived the injection were analyzed at 3dpf (five independent experiments: 16, 10, 19, 64 and 13, respectively). Of those, 71 embryos were found to be mCherry-positive and 86 % of them (61 out of 71) expressed eYFP in OSNs (Figure 4.4c). This result indicates that the shorter construct, which contains only 394 bp of sequence upstream of TSS is sufficient to drive the expression of reporter protein at a similar high frequency as the longer $p_{1.2kb}$OR101-1::EYFP-pA construct and it suggests that the deleted 206 bp of sequence does not play an essential role in regulation of OR101-1 expression.

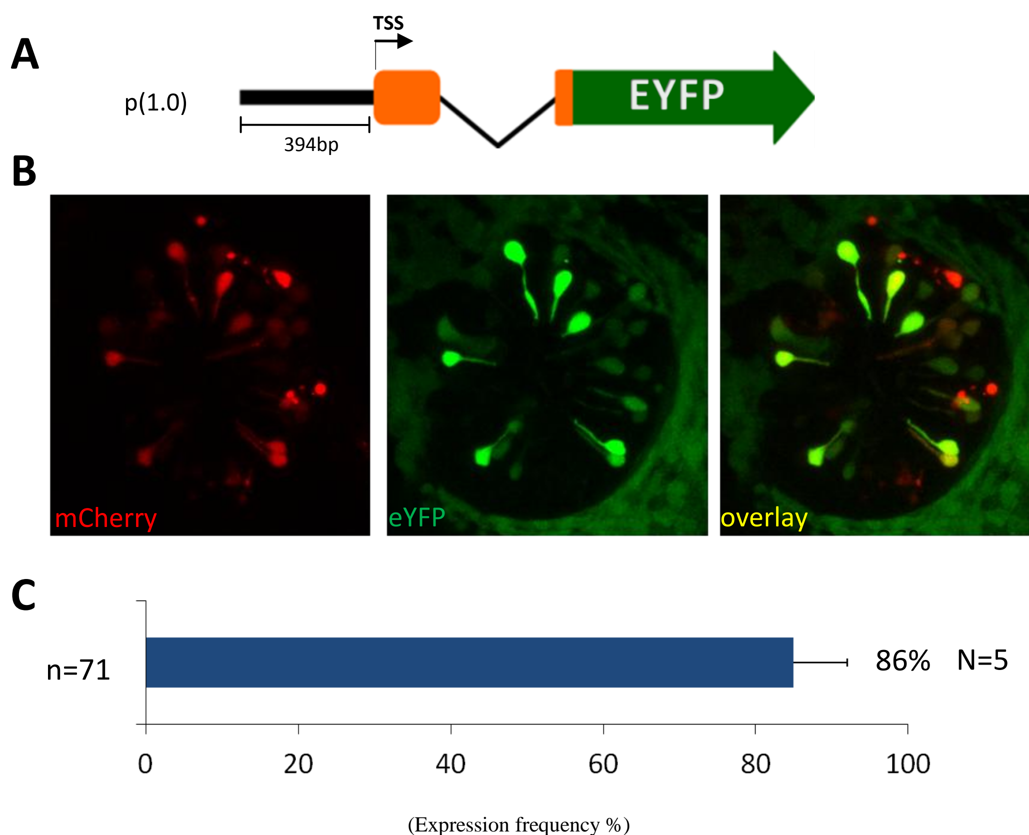


Figure 4.4. Expression analysis of the $p_{1.0kb}$OR101-1::EYFP-pA construct. (A) Overview of the construct. (B) Projection of a confocal z-stacks of a 3 dpf OE. (C) Expression frequency of the construct at 3 dpf relative to $p_{1.3kb}$OMP::mCherry-pA. n: mCherry (+) embryos. N: experiments.

Next, an additional 182 bp of sequence was deleted from the 5'-end to generate $p_{0.8kb}$OR101-1::EYFP-pA. The newly generated construct includes 212 bp genomic sequence upstream of the TSS along with untranslated and intronic sequences (Figure

4.5a). The construct was co-injected with the OMP promoter reference plasmid into 395 one-cell stage zebrafish oocytes (four independent experiments: 60, 90, 162 and 83, respectively) of which 179 embryos (18, 58, 46 and 57, respectively) survived until the time of analysis at 3 dpf. In total, 139 of 179 survived embryos were analyzed. Of those, 80 embryos were found to be mCherry-positive and 75 % of those (60 out of 80) expressed eYFP in OSNs (Figure 4.5c). Thus, the proximal 212 bp upstream of TSS of the OR101-1 gene are entirely sufficient to control OR101-1 expression or additional regulatory sites are located within the intronic or untranslated sequence.

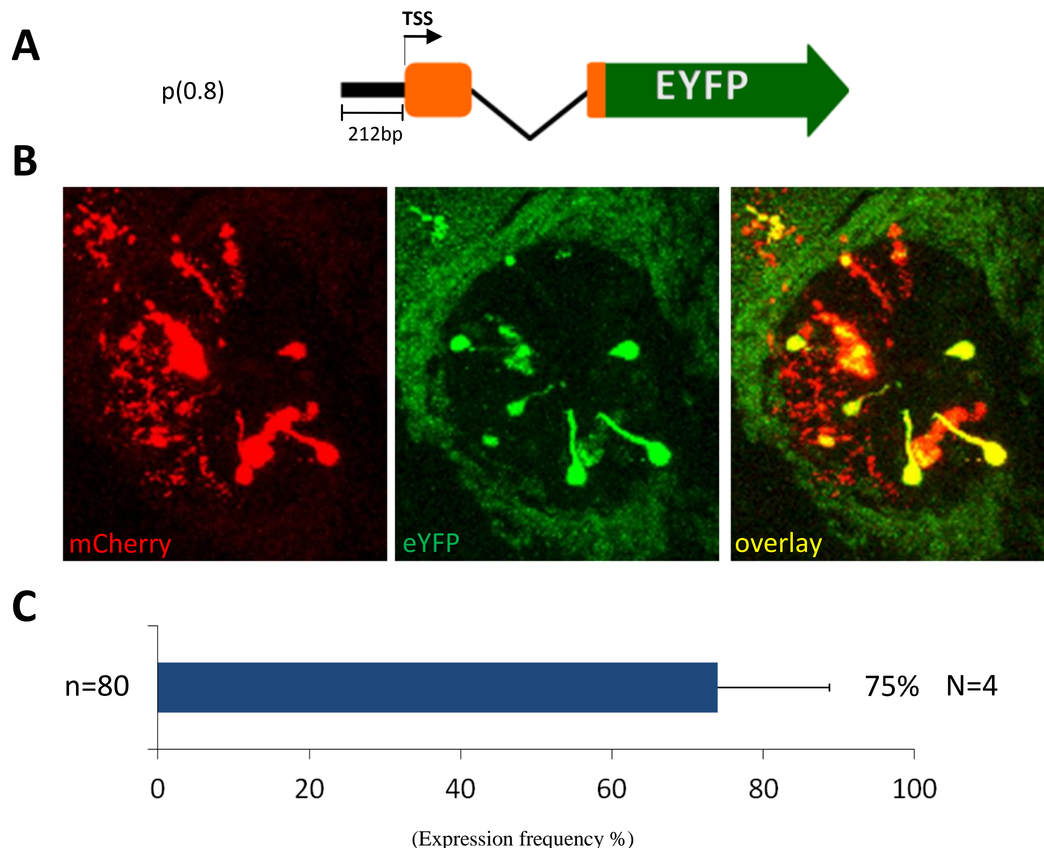


Figure 4.5. Expression analysis of the $p_{0.8kb}$OR101-1::EYFP-pA construct. (A) Overview of the construct. (B) Projection of a confocal z-stack of a 3 dpf OE. (C) Expression frequency of the construct at 3 dpf relative to $p_{1.3kb}$OMP::mCherry-pA. n: mCherry (+) embryos. N: experiments.

In summary, the genomic sequences upstream of TSS of OR101-1 gene were functionally dissected by promoter bashing to test their contribution to OR101-1 promoter activity. By using a transient expression approach in zebrafish embryos, positive regulatory sequences could be narrowed down to a sufficient core of 212 bp upstream of the TSS without drastically lowering promoter activity. When the longest and the shortest construct are compared, only a modest overall decrease in promoter efficiency of 6% (from 81% in $p_{1.2kb}$OR101-1::EYFP-pA to 75% in $p_{0.8kb}$OR101-1::EYFP-pA) was observed. Similar results were also observed for the mouse M71 odorant receptor gene (Rothman *et al.,* 2005), suggesting that, at least in some OR genes, cis-regulatory motifs are contained with short proximal sequences. However, the analysis so far does not address whether critical cis-activating motifs could be located downstream of the TSS, either in the intron or in untranslated regions.

4.2.2.1. Analysis of Transgene Expressing Cells in Embryos. Expression analysis of the three constructs that contained 1.2, 1.0, and 0.8 kb upstream of the OR101-1 coding sequence based on the penetrance of transgene efficiency suggests that there is only modest decrease in efficiency for the shortest construct. To alternatively assess the effect of successive shortening of the sequence upstream of the TSS on expression the average number of transgene-expressing OSNs was analyzed. To do so, the average ratio of eYFP-positive OSNs to mCherry-positive OSNs per embryo were determined for each construct and a Student's t-test was applied to determine the level of significance of the results. On average, 4±0.33 mCherry and 3.16±0.33 eYFP expressing OSNs could be observed in the OE of $p_{1.2kb}$OR101-1::EYFP-pA-injected embryos ($n = 101$), which results in a ratio of 0.76±0.049 eYFP-/mCherry-positive OSNs. Similar results were observed for $p_{1.0kb}$OR101-1::EYFP-pA-injected embryos with on average 3.62±0.39 mCherry- and 2.48±0.25 eYFP-expressing OSNs ($n = 71$), yielding a ratio of 0.78±0.075 eYFP-/mCherry-positive OSNs. Statistical analysis showed that the difference between the average proportion of eYFP cells(eYFP/mCherry) for these two constructs is not significant ($p = 0.818$), which suggests that the first 206 bp of sequence in the longest construct does not play an essential role in driving expression. This result is also reflected by the difference in the normalized number of expressing embryos for these constructs (81 % and 86 %, respectively). However, when $p_{0.8kb}$OR101-1::EYFP-pA-injected embryos

were analyzed a significant difference was found. On average, 4.34±0.36 mCherry- and 2.4±0.31 eYFP-expressing OSNs could be observed (n = 80), giving a ratio of 0.56±0.050 eYFP-/mCherry-positive cells. The average proportion of eYFP cells in $p_{0.8kb}$OR101-1::EYFP-pA-injected embryos is significantly lower than those of the 1.2 and 1.0 kb constructs (0.76±0.049 and 0.78±0.075 for 1.2 and 1.0 kb constructs, respectively), which means that 0.8 kb upstream sequence could drive the expression of reporter gene in significantly less number of OSNs ($p < 0.01$ for both). The observed difference in the ratio of transgene expressing cells is correlated with the difference observed in the penetrance of transgene-expressing embryos for 1.0 kb and 0.8 kb constructs (86 % and 75 %), however, the difference in penetrance between $p_{0.8kb}$OR101-1::EYFP-pA and $p_{1.0kb}$OR101-1::EYFP-pA did not reach statistical significance.

As previously mentioned, each construct was co-injected with $p_{1.3kb}$OMP::mCherry-pA into zebrafish embryos and so each embryo is expected to express both transgenes (eYFP and mCherry) in most cases. However, there were some fish which only expressed mCherry. To correct for this observation, we independently assessed the cell number ratios for fish that expressed both transgenes and excluded those that did not express eYFP. In this analysis, the average ratio of eYFP cells (eYFP-/mCherry-positive cells) increased for $p_{1.2kb}$OR101-1::EYFP-pA (0.94±0.040, n = 82) and $p_{1.0kb}$OR101-1::EYFP-pA (0.91±0.076, n = 61). Again the difference between those constructs was not statistically significant (p = 0.723). The average proportion of eYFP cells (eYFP-/mCherry-positive cells) increased also for $p_{0.8kb}$OR101-1::EYFP-pA (0.75±0.045, n = 60) when the correction was applied. The results of $p_{0.8kb}$OR101-1::EYFP-pA remain statistically different from $p_{1.2kb}$OR101-1::EYFP-pA but not from $p_{1.0kb}$OR101-1::EYFP-pA (p = 0.002 and p = 0.072, respectively).

In conclusion, the combined analysis of transgene expression in cells and embryos showed that the first 212 bp of genomic sequence upstream of the OR101-1 TSS is most relevant for transgene expression. Additional sequences further upstream affect transgene efficiency only moderately, for instance an increase from 75 to 81% when the two constructs at each extreme were compared (Figure 4.3 and Figure 4.5). However, there

might be a slight decrease in the overall number of transgene-expressing OSNs for the shortest construct $p_{0.8kb}$OR101-1::EYFP-pA and this difference reached significance under some conditions. As will be detailed further below, the breakpoint between $p_{0.8kb}$OR101-1::EYFP-pA and $p_{1.0kb}$OR101-1::EYFP-pA harbors a candidate homeodomain site and exclusion of this site may be the reason for the observed phenotype. Regardless, the OR101-1 promoter is rather compact and does not require sequences more distant than 212 bp from the TSS to promote high efficiency of expression.

4.2.3. Bioinformatics Analysis of Sequences Upstream of TSS

Following the observation that the first 212 bps of genomic sequence upstream of the OR101-1 TSS appear to be most relevant for transgene expression ($p_{0.8kb}$OR101-1::EYFP-pA) while sequences further upstream might additionally increase the transgene efficiency at least on the level of transgene-positive cells, the first 2 kb of sequence upstream of the OR101-1 TSS were subjected to a bioinformatic analysis to find out if it contains any cis-regulatory sites that promote high efficiency of expression. Because no homologous sequences of the OR101-1 gene locus are available from other fish species, the analysis was limited to known transcription factor binding sites in vertebrates. Screening of this sequence revealed primarily two DNA sequence motifs (Figure 4.6a) which were also shown to be overrepresented in promoter regions of mouse OR and VR genes belonging to the Olf1/Ebf1 (O/E) and homeodomain transcription factor (HD) families (Michaloski *et al.*, 2006; 2012). An OE-like motif "TTACCAGGGA" can be identified 88 bp upstream of TSS and a homeodomain-binding site "TAATTA" 215 bp upstream of the OR101-1 TSS (Figure 4.6b). While a total of 3 additional candidate homeodomain-binding sites could be identified within the genomic upstream sequences contained within the 1.2 kb promoter construct, no additional sequences resembling OE sites could be found upstream of the OR101-1 TSS. The next O/E-like site in this analysis locates about 1.6 kb upstream of the TSS, thus about 1 kb upstream of the highly efficient 1.2 kb promoter. O/E-like binding sites were first identified in the OMP gene promoter and shown to be critical for olfactory tissue specific expression (Kudrycki *et al.*, 1993). Later, the same motif was also identified in the proximal promoter regions of other OSN-specific genes such as G_{olf}, type III adenyl

cyclase, and the olfactory-specific cyclic nucleotide activated ion channel (Wang and Reed, 1993). O/E-like sites were also recognized to be overrepresented within the upstream regions of mammalian (M71, M72 and MOR23 *etc.*) and zebrafish OR genes (OR2.1; Vassalli *et al.*, 2002). A comprehensive bioinformatic analysis revealed that O/E-like sites can be identified in 87% of mouse OR and all mouse VR genes that were analyzed (Michaloski *et al.*, 2006; 2012). Both, O/E- and homeodomain-binding sites were functionally studied in the context of the mouse M71 OR gene promoter. Individual loss-of-function mutations of either O/E- or homeodomain-binding sites within the minimal promoter region of M71 revealed no dramatic effect in the expression of a transgene (except a decrease in cell numbers for homeodomain-binding site-mutated transgene) but consistently resulted in a ventralization of the epithelial pattern of transgene-expressing OSNs. Yet, combined mutation of both sites completely abolished expression of the transgene (Rothman *et al.*, 2005).

Review of the $p_{1.2kb}$OR101-1::EYFP-pA, $p_{1.0kb}$OR101-1::EYFP-pA, and $p_{0.8kb}$OR101-1::EYFP-pA constructs tested so far reveals that the most proximal homeodomain-binding site resides directly at the breakpoint between $p_{0.8kb}$OR101-1::EYFP-pA and $p_{1.0kb}$OR101-1::EYFP-pA constructs; hence, the $p_{0.8kb}$OR101-1::EYFP-pA construct excludes the critical "AA" nucleotides of the TAATTA motif and which have been shown to be essential for DNA binding (Figure 4.6c; Hirota and Mombaerts, 2004). Therefore, it is conceivable that the slight decrease in the overall number of transgene-expressing OSNs between the $p_{1.0kb}$OR101-1::EYFP-pA and $p_{0.8kb}$OR101-1::EYFP-pA constructs might be caused by the disruption of protein binding to the homeodomain-binding site (Figure 4.6c). This observation is consistent with the previously suggested role of homeodomain-binding sites, which are assumed to modulate the probability of OR gene choice (Vassalli *et al.*, 2011). However, even in the absence of a functional homeodomain-binding site in the $p_{0.8kb}$OR101-1::EYFP-pA construct, overall high levels of transgene expression are observed, suggesting that the definition of homeodomain-binding sites are not well conserved in zebrafish or that additional, as of yet undefined, sequence motifs are more critical to control OR101-1 expression. Because of the largely unperturbed high transgene efficiency of $p_{0.8kb}$OR101-1::EYFP-pA which contains only a single candidate

O/E-like site upstream of the TSS (Figure 4.6c), the O/E motif was subjected to further analysis using a mutagenesis approach.

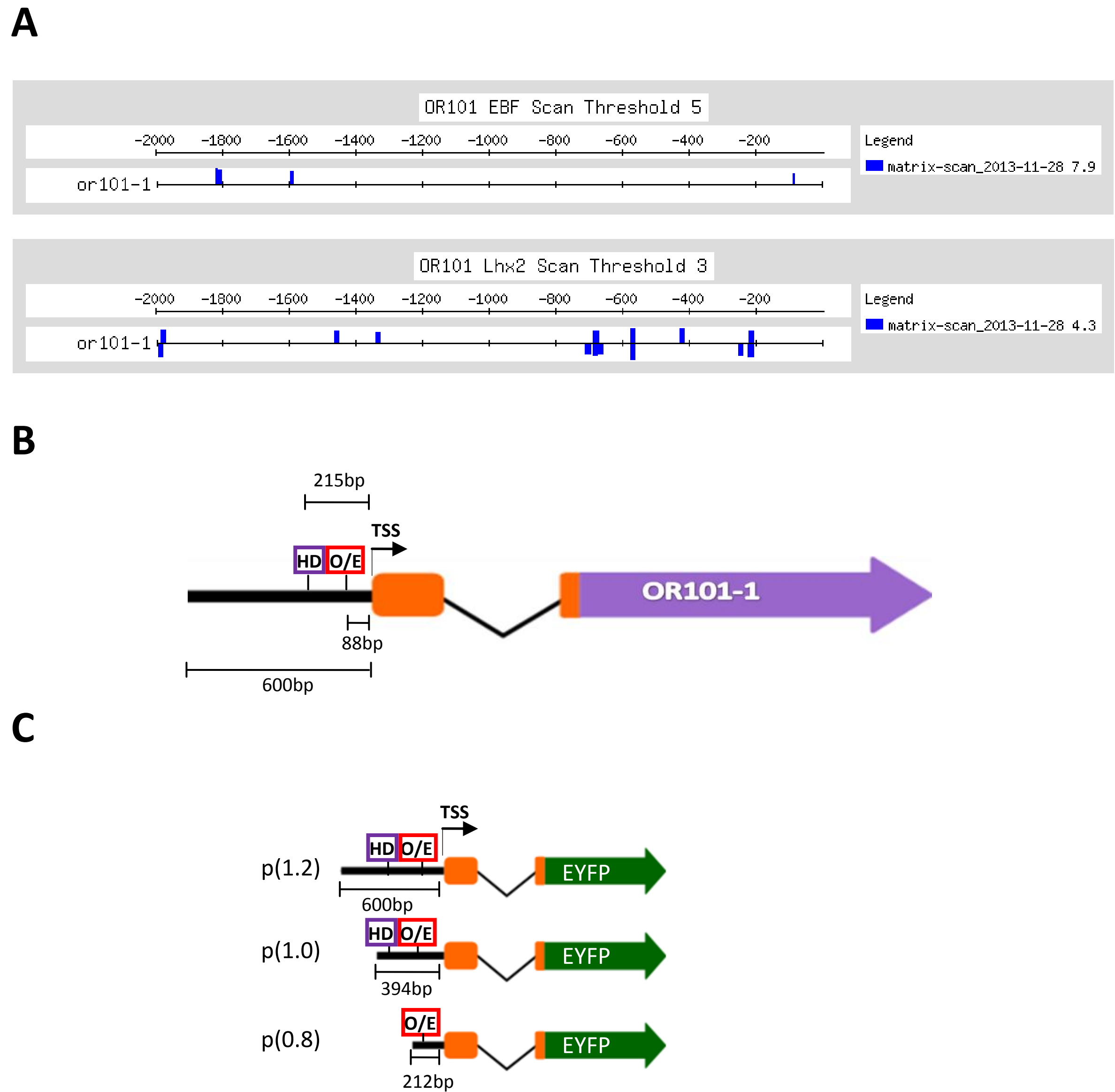


Figure 4.6. Motif-based analysis of genomic upstream sequences of OR101-1 TSS. (A) Scan of 2.0 kb of sequences upstream of TSS for O/E-like site (EBF) and Homeodomain-binding site (Lhx2). (B-C) Representation of the candidate O/E and most proximal HD-binding site in gene and constructs.

4.2.3.1. Analysis of Mutations within the proximal O/E Site. To further test the role of the candidate O/E-like site on transgene expression functionally, the site was subjected to site-directed mutagenesis in the context of the $p_{1.2kb}$OR101-1::EYFP-pA construct. Using overlap extension PCR, the original TTACCAGGGA sequence was converted to TTAAAAGGGA. These substituted CC nucleotides were previously shown to be critical for O/E protein binding (Rothman *et al.*, 2005) and O/E binding should be abolished in the mutated sequence. Then, the mutated construct, $p_{1.2kb\text{-}\Delta O/E}$OR101-1::EYFP-pA (Figure 4.7a), was co-injected with the OMP promoter reference plasmid into a total of 225 one cell-stage zebrafish oocytes (four independent experiments: 54, 58, 45 and 68, respectively) of which 190 embryos (42, 51, 36 and 61, respectively) survived until the time of analysis at 3 dpf. In total, 86 of 190 surviving embryos were analyzed. Of those, 65 embryos were found to be mCherry-positive and 88 % of those (57 out of 65) expressed eYFP in OSNs (Figure 4.7c). This high frequency of expression is not significantly different ($p = 0.468$) from the frequency of expression observed for the original non-mutated $p_{1.2kb}$OR101-1::EYFP-pA construct (81%). Thus, mutation of the proximal upstream O/E-like site did not result in a drop of activity of the OR101-1 gene promoter, suggesting that either this candidate O/E-like site does not play a functional role in expression control of the OR101-1 gene, or that the effect of the mutation might be compensated by additional cis-activating motifs located upstream and/or downstream of the TSS, including the intron or untranslated regions.

In addition to the overall frequency of transgene-expressing embryos, the average number of transgene-positive OSNs was analyzed for $p_{1.2kb\text{-}\Delta O/E}$OR101-1::EYFP-pA amd compared to the efficiency of the non-mutated construct. On average, 7.88±0.81 mCherry and 5.14±0.66 eYFP expressing OSNs per OE could be observed for $p_{1.2kb\text{-}\Delta O/E}$OR101-1::EYFP-pA-injected embryos (n=65), which results in a ratio of 0.63±0.041 eYFP-/mCherry-positive OSNs. Statistical analysis showed that the difference between the relative proportion of eYFP cells (eYFP/mCherry) for $p_{1.2kb\text{-}\Delta O/E}$OR101-1::EYFP-pA and $p_{1.2kb}$OR101-1::EYFP-pA (0.76±0.049, n=101) is not significant (p=0.073), similar to the normalized number of transgene-expressing embryos for these constructs (88 % and 81 %, respectively). However, in the more stringent analysis in which only fish that expressed both transgenes (mCherry and eYFP) were analyzed (refer to Section 4.2.2.1 for detail), a

significant difference was found ($p<0.001$). When eYFP negative fish were excluded from the analysis, the normalized ratio of eYFP cells (eYFP-/mCherry-positive cells) for $p_{1.2kb-\Delta O/E}$OR101-1::EYFP-pA increased to 0.724±0.033 (n=57), which is significantly lower than the 0.94±0.040 ratio observed for $p_{1.2kb}$OR101-1::EYFP-pA ($n = 82$, $p < 0.001$). Thus, the same genomic upstream sequence drives the expression of a reporter gene in a lower number of OSNs when only two nucleotides within the predicted O/E-like site are mutated. However, unlike the difference in cell number, the difference in penetrance did not reach statistical significance between the two constructs.

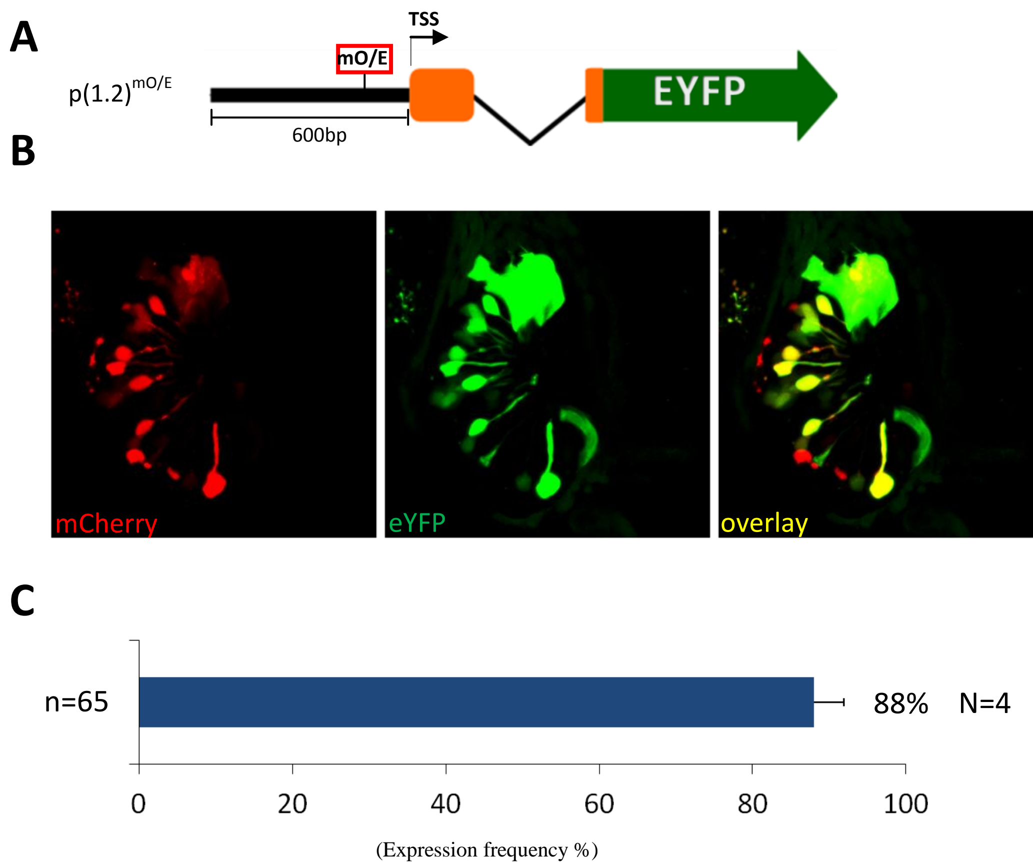


Figure 4.7. Expression analysis of the $p_{1.2kb-\Delta O/E}$OR101-1::EYFP-pA construct. (A) Overview of the construct. (B) Projection of a confocal z-stacks of a 3 dpf OE. (C) Expression frequency of the construct at 3 dpf relative to $p_{1.3kb}$OMP::mCherry-pA. n: mCherry (+) embryos. N: experiments.

In summary, the proximal upstream O/E-like site, which is localized at -88 bp relative to the OR101-1 TSS was analyzed to understand its possible role as a *cis*-regulatory motif that might control gene expression using a mutagenic approach. The analysis of transgene expression at the level of transgene-expressing embryos and OSNs per OE revealed that the candidate site is largely dispensable for high penetrance of expression (88 %). However, a slight and significant decrease was observed in the overall number of transgene-expressing OSNs for the mutated $p_{1.2kb\text{-}\Delta O/E}$OR101-1::EYFP-pA construct ($p < 0.001$; Figure 4.15), under some conditions. This suggests that the proximal upstream O/E-like site might function synergistically with additional *cis*-regulatory motifs to control OR101-1 gene expression. Those, as of yet unknown sequences might be contained in the genomic upstream sequence or in intronic and untranslated sequences.

4.2.4. Analysis of Sequences Downstream of Transcription Start Site

Up to here, only sequences upstream of the OR101-1 TSS were subjected to expression analysis to uncover essential sequences that promote transgene expression in the context of the 1.2 kb constructs. However, all of the constructs analyzed so far, also contain 588bp downstream of the TSS, which include 196 bp of a 5'-noncoding exon, a 345 bp 5'-intron and additional 47 bp of noncoding sequence in front of the CDS in exon 2 (Figure 4.2a). Thus, because only a modest decrease of promoter efficiency was observed after successive shortenings of the sequence upstream of TSS (from 81% to 75%) and because the candidate proximal upstream O/E-site did not contribute to high expression (Figure 4.7), noncoding sequences downstream of the TSS were included in the analysis.

In a first construct, all sequences downstream of the TSS were removed and the newly generated construct, $p_{0.6kb\text{-}\Delta UTR\text{-}\Delta intron}$OR101-1::EYFP-pA, only included the 600 bp of sequence upstream of the TSS (Figure 4.8a). The $p_{0.6kb\text{-}\Delta UTR\text{-}\Delta intron}$OR101-1::EYFP-pA construct was co-injected with the $p_{1.3kb}$OMP::mCherry-pA control plasmid into a total of 232 one cell-stage zebrafish oocytes (three independent experiments: 98, 75 and 59, respectively) of which 123 embryos (65, 24 and 34, respectively) survived until the time of

analysis at 3 dpf. In total, 120 of the 123 surviving embryos were analyzed. Of those, 84 embryos were found to be mCherry positive and 42 % of those (35 out of 84) expressed eYFP in OSNs (Figure 4.8c). This result shows that deletion of all the sequences downstream of TSS caused a 50% decrease in the penetrance of transgene efficiency, which is 81 % for the full construct ($p_{1.2kb}$OR101-1::EYFP-pA). However, since the deleted region included all the sequences downstream of TSS, this result does not indicate whether the observed phenotype occurred as an effect of deletion of the UTR or 5'-intronic sequence. To address the role of intronic and untranslated sequences separately, a new construct which includes only the 5'-untranslated sequence but not the 5'-intron was generated (Figure 4.9a). This construct, $p_{0.8kb\text{-}\Delta intron}$OR101-1::EGFP-pA, was co-injected with the OMP promoter reference plasmid into 225 one-cell stage zebrafish oocytes (three independent experiments: 47, 85 and 93, respectively) of which 147 embryos (44, 40 and 63, respectively) survived until the time of analysis at 3 dpf. In total, 84 of the 147 surviving embryos were analyzed. Of those, 66 embryos were found to be mCherry-positive and 45 % of those (30 out of 66) expressed eGFP in OSNs (Figure 4.9c). Thus, removal of the 5'-intronic sequence resulted in a similar 50% decrease in penetrance of transgene efficiency for this construct compared to the full one (81 %, $p_{1.2kb}$OR101-1::EYFP-pA). Interestingly, the decrease in efficiency is similar to that of the $p_{0.6kb\text{-}\Delta UTR\text{-}\Delta intron}$OR101-1::EYFP-pA construct (42 %), suggesting that relevant regulatory sites might be located within the intronic sequence but not the UTR.

4.2.4.1. Analysis of Transgene Expressing Cells in Embryos. Similar to previous constructs, the average number of transgene-expressing OSNs was analyzed as it may constitute a more sensitive measure of transgene efficiency. On average, 8±0.62 mCherry and 1.30±0.25 eYFP expressing OSNs could be observed in the OE of $p_{0.6kb\text{-}\Delta UTR\text{-}\Delta intron}$OR101-1::EYFP-pA-injected embryos (n = 84), which results in a ratio of 0.14±0.024 eYFP-/mCherry-positive OSNs. Similar results were observed for $p_{0.8kb\text{-}\Delta intron}$OR101-1::EGFP-pA-injected embryos with an average of 8.59±0.84 mCherry- and 1.76±0.39 eYFP-expressing OSNs (n = 66), yielding a ratio of 0.16±0.029 eYFP-/mCherry-positive OSNs. Statistical analysis showed that the difference between the relative proportion of eYFP cells (eYFP/mCherry) for these two constructs is not significant (p=0.592). However, these two results are extremely lower than that of $p_{1.2kb}$OR101-1::EYFP-pA

(0.76±0.049) and the observed differences are highly significant ($p < 10^{-16}$), which is also reflected by the difference in the normalized number of expressing embryos for these constructs (42%, 45% and 81%, respectively). Thus, a comparison of these three constructs strongly suggests that 5'-intronic sequence might contain strong cis-activating motifs to regulate OR101-1 gene expression.

When the cell number analysis was applied only to fish that expressed both transgenes (mCherry and eYFP), the relative ratio of eYFP cells (eYFP-/mCherry-positive cells) increased for $p_{0.6kb\text{-}\Delta UTR\text{-}\Delta intron}$OR101-1::EYFP-pA (0.33±0.040, n = 35) and $p_{0.8kb\text{-}\Delta intron}$OR101-1::EGFP-pA (0.34±0.044, n = 30) but both results remained statistically different from $p_{1.2kb}$OR101-1::EYFP-pA (0.94±0.040, n = 82, $p < 10^{-12}$).

4.2.4.2. Bioinformatics Analysis of 5'-Intronic Sequences. The observed effect of removal of the 5'-intronic sequence on transgene efficiency suggests that regulatory sequences within the intron contribute significantly (36 %) to the strong promoter activity of the 1.2 kb genomic upstream sequence. This suggests that 5'-intronic sequence might contain additional cis-regulatory motifs to control OR101-1 gene expression. Thus, the 1.2 kb genomic sequence, including sequences downstream of the TSS was screened again for the presence of O/E-like sites (Figure 4.10a). Interestingly, an additional O/E-like site was identified within 394 bp downstream of TSS, along with the previously identified upstream proximal site that was tested by mutagenesis (Figure 4.10b).

In the bioinformatic sequence scan to pinpoint O/E-like sites only the mouse definition of the motif was used because a zebrafish-specific motif has not been reported so far. This raised the suspicion that the recognition motif for O/E proteins might have undergone evolutionary changes and would only partially be conserved in mammals compared to fish. To circumvent this problem, a phylogenetic footprinting was performed on the candidate OMP gene promoter from 14 different fish species. The analysis revealed a slightly different definition of the O/E-binding site in fish. Using the two available sequence matrices for the murine and human and the newly identified fish motif, the 1.2 kb

promoter was scanned again for the occurrence of these motifs. Interestingly, the proximal upstream O/E like site that was identified by the MA0154.1 Jaspar matrix was not recognized by the human- (MA0154.2) or fish-specific matrix. Interestingly, an additional candidate O/E-like site was identified in the 5'-UTR (Figure 4.11). This site, "CCTTTAGGGAA", is located within 5'-intronic sequence. Therefore, exclusion of this site may be the reason for the dramatic decrease in normalized number of transgene-expressing cells and embryos for $p_{0.6kb\text{-}\Delta UTR\text{-}\Delta intron}$OR101-1::EYFP-pA and $p_{0.8kb\text{-}\Delta intron}$OR101-1::EGFP-pA constructs.

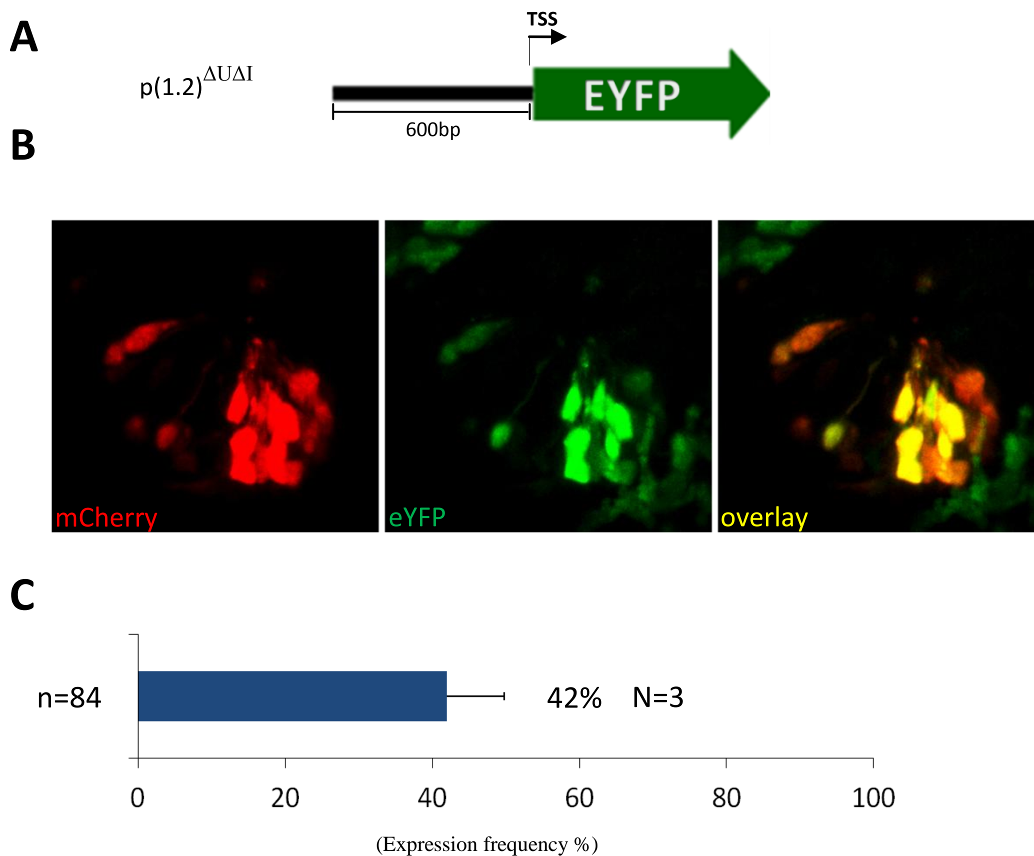


Figure 4.8. Expression analysis of the $p_{0.6kb\text{-}\Delta UTR\text{-}\Delta intron}$OR101-1::EYFP-pA construct. (A) Overview of the construct. (B) Projection of a confocal z-stacks of a 3 dpf OE. (C) Expression frequency of the construct at 3 dpf relative to $p_{1.3kb}$OMP::mCherry-pA. n: mCherry (+) embryos. N: experiments.

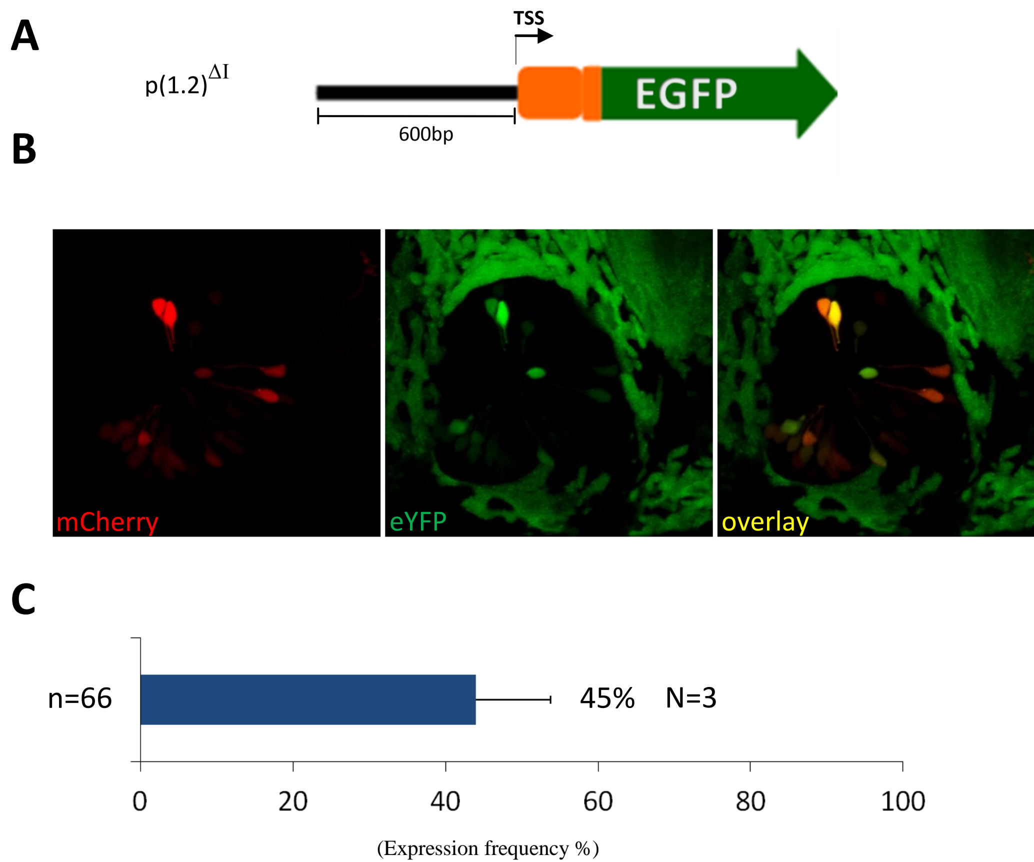
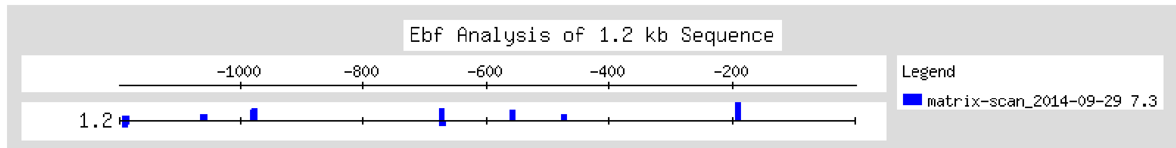


Figure 4.9. Expression analysis of the $p_{0.8kb\text{-}\Delta intron}$OR101-1::EGFP-pA construct. (A) Overview of the construct. (B) Projection of a confocal z-stacks of a 3 dpf OE. (C) Expression frequency of the construct at 3 dpf relative to $p_{1.3kb}$OMP::mCherry-pA. n: mCherry (+) embryos. N: experiments.

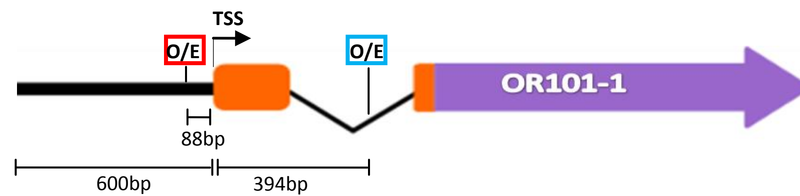
In conclusion, the deletion of 5'-intronic sequences from the 3'end of 1.2 kb promoter construct caused almost 50% decrease in transgene efficiency (from 81% to 45%, Figure 4.9). This highly significant effect which was observed both in cells and embryos strongly suggested a cis-regulatory role for 5'-intronic sequence on control of OR101-1 expression. Indeed, an additional O/E-like site was identified within the 5'-intronic sequences via bioinformatics analysis with motif-based approach. When the results are reconsidered with this and previously identified candidate motifs, it implies a regulatory model in which cis-activating motifs being located at upstream and downstream of TSS combinatorily act on to control OR101-1 gene expression. However, it cannot be ruled out

if splicing *per se* would increase transgene efficiency as both constructs tested here would not be spliced.

A



B



C

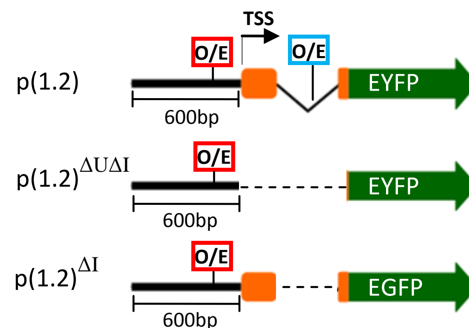


Figure 4.10. Motif-based analysis of downstream sequences of OR101-1 TSS. (A) Scan of 1.2 kb of sequences upstream of OR101-1 coding region for Olf1/Ebf1-like site (EBF). (B -C) Representation of the previously identified (red) and new candidate motif (blue) in gene and constructs.

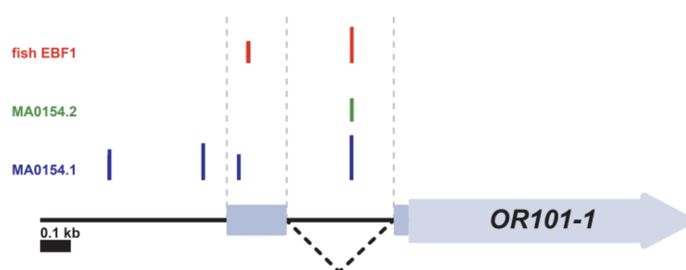


Figure 4.11. Scan of the 1.2 kb promoter region for the motifs representing O/E-binding sites identified in human (green), murine (blue) and fish (red).

4.2.5. Analysis of Expression with Olf1/Ebf1 Gene Knock Down

Following the identification of O/E-binding sites in the proximal upstream and 5'-UTR of 1.2 kb promoter and the analysis of proximal O/E site with mutagenesis, the potential role of these candidate sites on the high expression efficiency from 1.2 kb promoter was also studied in the context of bound protein,- Olf1/Ebf1. Olf1 protein belongs to repeated helix-loop-helix transcription factor family, which was first identified from a rat olfactory cDNA library by genetic selection in yeast and its mouse homolog was shown to be expressed in olfactory epithelia (Wang and Reed, 1993). It was also cloned by an independent group and characterized as Early B-Cell Factor (Ebf) being functional in B-cell development (Hagman *et al.*, 1993). Mice lacking functional Ebf1 protein were reported to display a profound B-cell deficit, however, the morphology and gene expression was normal in olfactory epithelium (Lin and Grosschedl, 1995). Later, two other members of mouse Olf1/EBF transcription factor family were identified (Wang *et al.*, 1997) and mice with null mutations of either Ebf2 or Ebf3 exhibited defect in axonal projection of olfactory neurons to dorsal olfactory bulb, suggesting a role for these proteins in olfactory receptor neuron projection (Wang *et al.*, 2004).

When zebrafish genome was searched for Ebf1 gene in *Ensembl* database, two zebrafish ortholog occurred for the gene, - Ebf1a and Ebf1b. To determine which one of them is primarily expressed in olfactory tissue; previously obtained RNA-Seq Data was used, which shows the expression levels of genes in olfactory tissue normalized to brain

tissue. According to this data (Table 4.1), Ebf1b is primarily expressed in zebrafish olfactory tissue. Thus, Ebf1b protein was included in the analysis. In order to knock down Ebf1b protein and analyze its effect on 1.2 kb promoter activity, the strategy of Morpholino antisense oligos (MO) was employed, which proved high efficacy in zebrafish (Draper *et al.*, 2001). Initially, one translation- and one splice site-blocking morpholino were designed for Ebf1b protein (Table 3.2). For splice site-blocking MO, exon5-intron5 boundary was targeted to exclude exon5 in the final protein, which was shown to encode critical residues for EBf1b protein binding (Hagman *et al.*, 1995).

Table 4.1. Expression levels of Ebf1a and Ebf1b genes in olfactory tissue at log_2 scale, which were normalized to brain tissue.

Gene	**Brain**	**Nose**	**Brain/Nose($_{log2}$)**
Ebf1a	3,617704	2,842503	-0,34791
Ebf1b	3,148081	16,46867	2,387179

To begin with, 1.2 kb promoter construct ($p_{1.2kb}$OR101-1::EYFP-pA) was co-injected with the OMP promoter reference plasmid into a total of 179 one cell-stage zebrafish oocytes in a single experiment. 59 of those 179 injected embryos were kept as control. Then, the rest of embryos (120) were additionally injected with Ebf1b-transblocking-MO, 8 ng per embryo at four cell-stages. 20 of 120 MO-injected- and 38 of 59 only 1.2 kb construct-injected embryos (control) survived until the time of analysis at 3 dpf. In total, 17 of 20 surviving MO-injected embryos were analyzed. Of those, 9 embryos were found to be mCherry-positive and all those (9 of 9) expressed eYFP in OSNs. For the control case, 19 of 38 surviving embryos were analyzed. Of those, 12 embryos were found to be mCherry-positive and 9 of them expressed eYFP in OSNs (Figure 4.12a). According to these results, transgene efficiencies for 1.2 kb promoter construct are similar for MO-injected and non-injected embryos. However, the survival ratio was too lower for MO-injected fish (17 %) compared to the high ratio for the control fish (64 %). The delivery of overdose of morpholino might be the reason for the observed phenotype. To avoid the presumed toxicity, the amount of injected MO was decreased to 4 ng per embryo in next

experiment. In this analysis, 1.2 kb promoter construct ($p_{1.2kb}$OR101-1::EYFP-pA) was co-injected with the OMP promoter reference plasmid into 53 one cell-stage zebrafish oocytes. After the save 20 of 53-injected embryos as control, the rest of embryos (33) were additionally injected with Ebf1b-Transblocking-MO, 4 ng per embryo at four cell-stages. 14 of 33 MO-injected- and 13 of 20 only 1.2 kb construct-injected embryos (control) survived until the time of analysis at 3 dpf. Survival ratios were closer for both fish (42 % for MO-injected and 65% for control) compared to previous experiment. All 14 surviving MO-injected-embryos were analyzed. Of those, 9 embryos were found to be mCherry-positive and all those (9 of 9) expressed eYFP in OSNs. All 13 surviving embryos were analyzed for the control fish. Of those, 11 embryos were found to be mCherry-positive and 10 of them expressed eYFP in OSNs (Figure 4.12b). According to these results, similar transgene efficiencies were observed for 1.2 kb promoter in MO-injected and non-injected embryos. However, the activity of injected MO within these analyzes were not verified by control experiments. When Ebf1b proteins were assumed to be efficiently knocked down in MO-injected embryos, these results suggests their abolishment did not cause any effect on transgene efficiency or the potential effect was compensated by functional redundancy from other members of Ebf family. However, these preliminary results are not enough quantitatively to calculate standard errors, yet (only two experiments with different conc. of MO); therefore additional experiments are needed to be done. Most importantly, the working concentration of Ebf1b morpholino per embryo should be determined to reach high efficacy. Moreover, as previously mentioned in analyses of other transgenic constructs, $p_{1.3kb}$OMP::mCherry-pA was also used in MO-injections as a positive control for injection efficiency and normalization of results. However, since OMP gene also has O/E-binding site in its proximal promoter region like other olfactory specific genes (Kudrycki *et al.*, 1993), transgene efficiency of $p_{1.3kb}$OMP::mCherry-pA would also be expected to decrease with the knock down of Ebf1b protein, impeding clear observation of the potential effect on 1.2kb promoter activity. Therefore, a different control plasmid would be more reliable for better analysis.

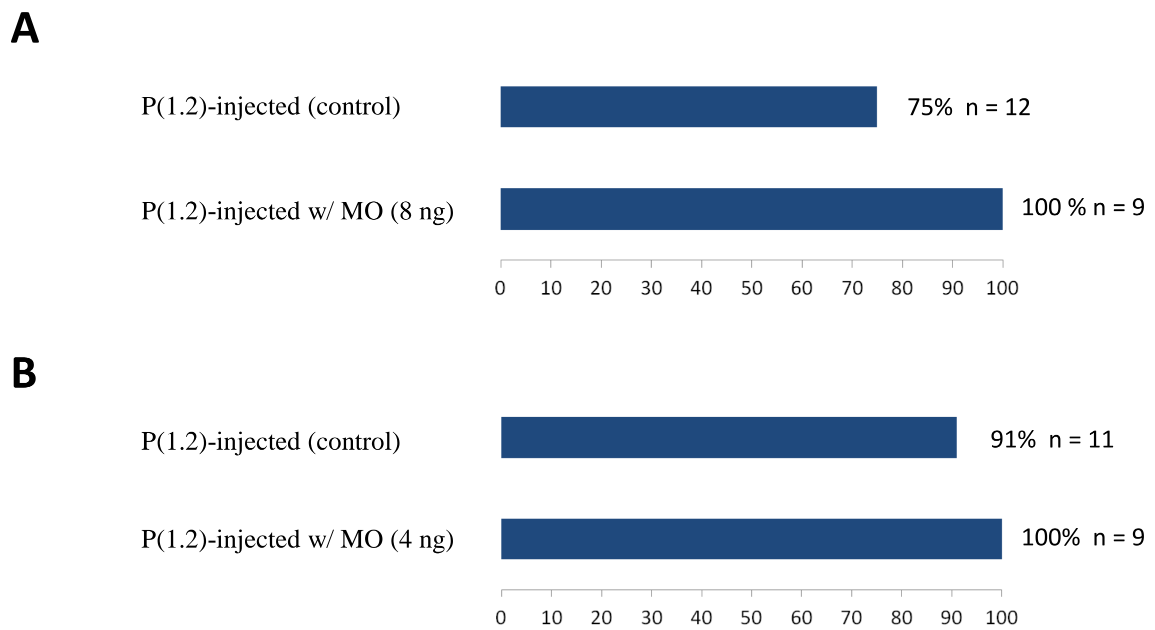


Figure 4.12. Expression analysis of the $p_{1.2kb}$OR101-1::EYFP-pA construct with co-injection of Ebf1b translation blocking morpholino. (A) Results of 8 ng MO-injection per embryo. (B) Results of 4 ng MO-injection per embryo (in single experiments, respectively).

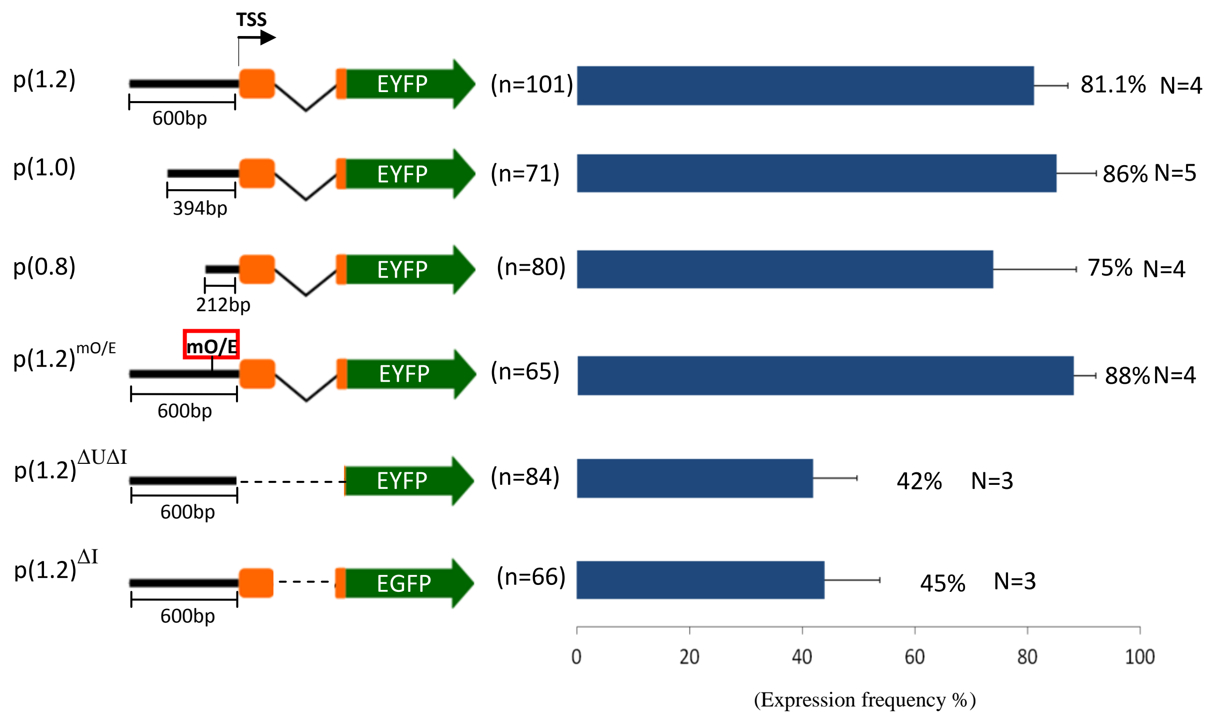


Figure 4.13. Summary of expression analysis of transgenic constructs throughout Section 4.2. (Left) Overview of the transgenic constructs. (Right) Penetrance of expression in embryos at 3 dpf. n: mCherry positive embryos analyzed. N: experiments.

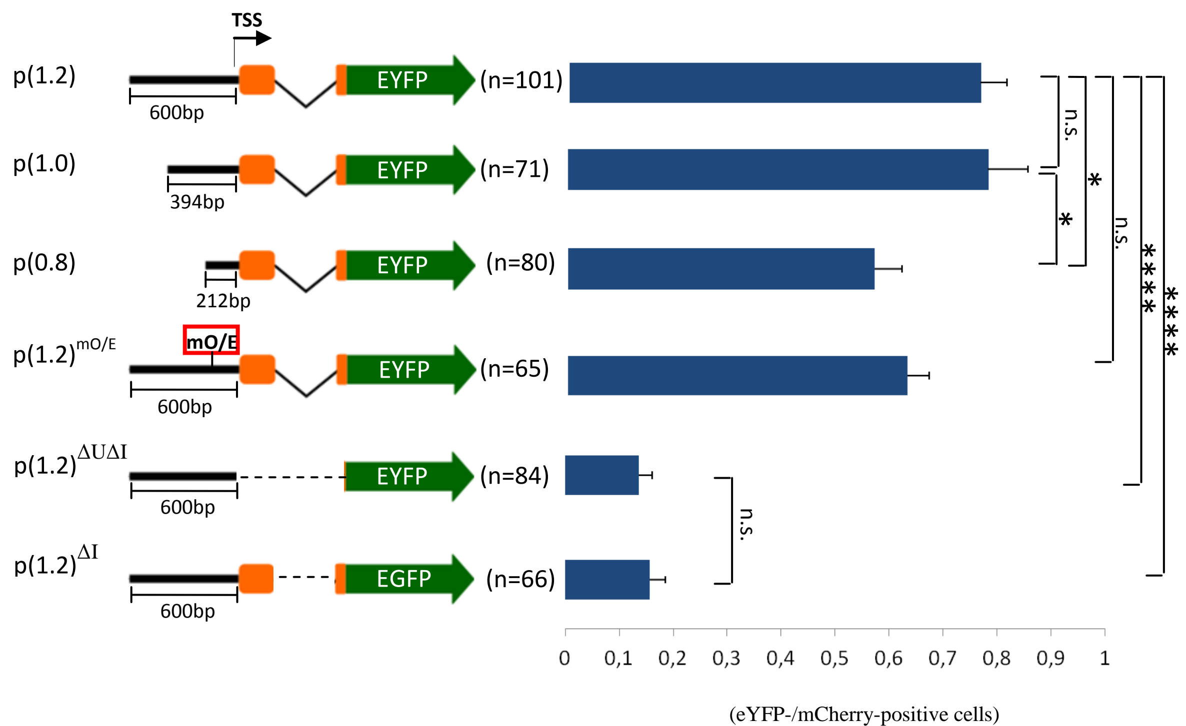


Figure 4.14. Summary of cell number analysis of transgenic constructs throughout Section 4.2. (Left) Overview of the constructs. (Right) Average ratio of eYFP-expressing cells(eYFP/mCherry) at 3 dpf. *$p < 0.01$, ****$p < 10^{-16}$ (Student's t test). n.s., not significant. n: mCherry positive embryos.

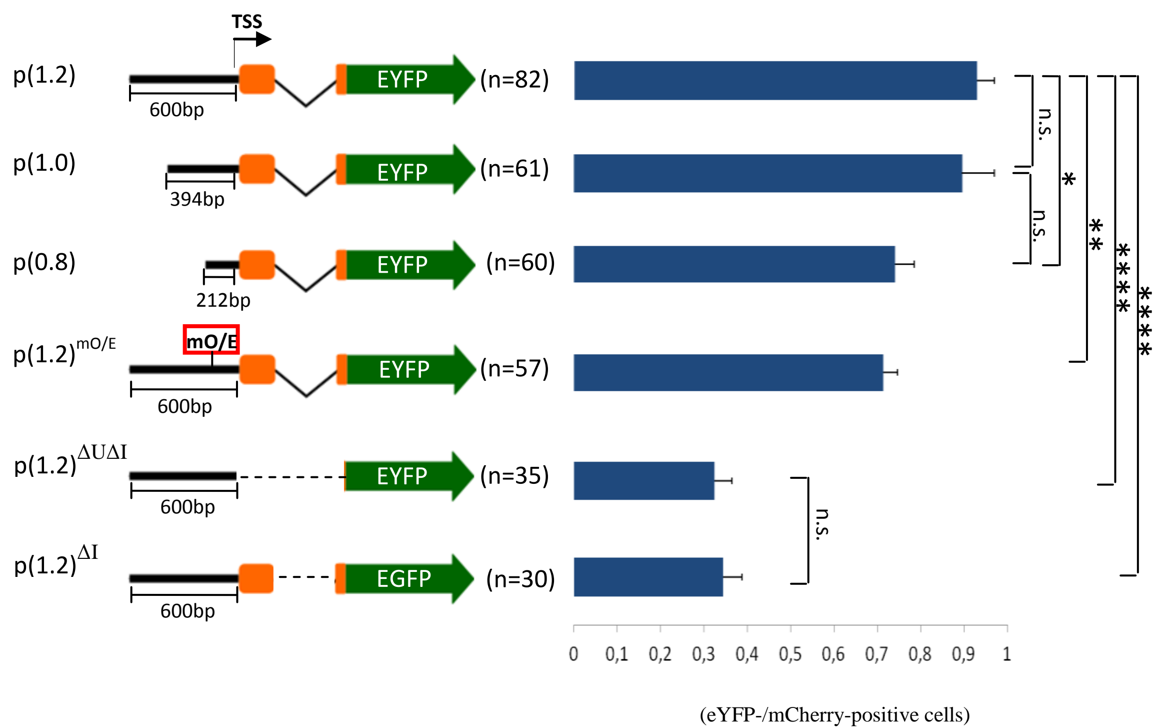


Figure 4.15. Summary of alternative cell number analysis of transgenic constructs throughout Section 4.2. (Left) Overview of the constructs. (Right) Average ratio of eYFP-expressing cells(eYFP/mCherry) at 3 dpf. *$p < 0.01$, **$p < 0.001$, ****$p < 10^{-12}$ (Student's t test). n.s., not significant. n: eYFP and mCherry positive embryos.

5. DISCUSSION

Two basic principles govern pattern formation in the vertebrate olfactory system. Firstly, each mature OSN expresses only one OR gene throughout its lifespan (Malnic *et al.*, 1999) and this OR gene is actively transcribed from only one allele (Chess *et al.*, 1994). Secondly, OSNs expressing the same OR gene project their axons to the same topographically fixed glomeruli in the OB (Mombaerts *et al.*, 1996; Wang *et al.*, 1998). Thus, OR expression in OSNs does not only determine their odorant specificity but also instructs their axons to target proper synaptic partners in the brain. However, it is not fully understood how OSNs choose one OR gene for expression from a very large OR repertoire. Recent progress has been made with the identification of LSD1, which appears to release individual OR gene loci from epigenetic silencing (Lyons *et al.*, 2013). Yet, even though the process of OR gene choice has historically been referred to as stochastic (Qasba and Reed, 1998; Serizawa *et al*, 2000), aspects of OR expression are non-random. For instance, expression of each OR is restricted to a narrow expression domain or zone and does not cover the entire sensory epithelium. In addition, ORs differ widely in the number of OSNs by which they are expressed. Thus, even though the LSD1 mechanism may explain monogenic / monoallelic expression, the interaction of LSD1 with OR gene loci appears to be guided by additional factors or mechanisms. Both distal *cis*-acting elements (Serizawa *et al.*, 2003; Fuss *et al.*, 2007; Nishizumi *et al.*, 2007) and proximal promoter elements have been implicated in the control of OR gene expression (Vassalli *et al.*, 2002; Rothman *et al.*, 2005) and might be relevant to further pattern OR expression across the OE.

5.1. Short-range vs long-range control of a zebrafish OR gene

To better understand the role of proximal promoter sequences in OR gene expression, regulation of the zebrafish OR101-1 gene by proximal sequence elements was studied. A previous promoter bashing study located several regulatory sequences within the first 3.5 upstream of the OR101-1 coding sequence using a transient transgenic assay.

It was found that the proximal 1.2 kb of sequence upstream of the OR101-1 coding region drives expression of fluorescence reporter proteins in zebrafish OSNs with high efficiency (61.3 %; Söğünmez, 2012). The 1.2 kb construct consists of 600 bp of genomic sequence upstream of the TSS, 196 bp of a 5'-noncoding exon, 345 bp of 5'-intronic sequence, and 47 bp of untranslated sequence directly followed by the coding frame (Figure 4.2A-B). Because of its high efficiency in transgenic assays, the 1.2 kb proximal promoter region most likely includes all relevant positive regulatory sites for OR101-1 gene expression.

Indeed, transgenic constructs comprising longer fragments of sequences upstream of the OR101-1 TSS were less well expressed. Thus, sequences located between 2 and 2.5 kb upstream of the OR101-1 coding sequence severely reduced the function of the promoter and acted as negative regulators (Söğünmez, 2012; Sancer 2015). It remains uncertain at this point, whether additional regulatory sequences could be located further upstream of the locus. Nevertheless, the zebrafish OR101-1 gene appears to be regulated by proximal positive and distal negative regulatory elements. The aim of the study presented here was to pinpoint positive regulatory sites and to understand the contribution of intronic and untranslated sequences to promoter function.

5.2. The basic OR101-1 promoter is contained within 212 bp upstream of the TSS

To do so, the promoter sequence of the basic $p_{1.2kb}$OR101-1::EYFP-pA construct was dissected by promoter bashing. In this study, expression efficiency of this construct was slightly higher (81 %) than the 65% observed by Söğünmez (2012). This difference may be explained by slight, but unavoidable, differences in the delivery of the DNA to zebrafish oocytes. Confounding factors might be small differences in DNA concentration or purity, as well as differences in the genetic background of the animals. Nevertheless, 81% transgenic efficiency allows for a rather wide dynamic range to detect even small differences in the efficiency of differently sized constructs.

The 1.2 kb promoter construct contains both, sequences upstream and downstream of the OR101-1 TSS. The sequences downstream comprise an intron and untranslated exonic sequences, which might contribute to regulation of expression (Furger *et al*., 2002; Nott *et al*., 2003). To narrow down the core proximal promoter sequences the construct was successively shortened from its 5'-end. Interestingly, removal of 388 bp from the 5'-end did not result in a dramatic loss of transgene efficiency. The transgenic construct included only the proximal 212 bp sequence upstream of OR101-1 TSS and still showed efficiency as high as 75%. Thus, positive regulatory sequences could be narrowed down to a sufficient core of 212 bp upstream of the TSS with only a modest decrease in promoter efficiency of 6%.

However, when the average numbers of transgene-expressing OSNs were compared for these constructs, a slight decrease was observed for the shortest construct $p_{0.8kb}$OR101-1::EYFP-pA, which reached significance under some conditions ($p<0.01$; Figure 4.14 and Figure 4.15). These observations suggested that the first 212 bp sequence upstream of TSS is most relevant for high penetrance of transgene expression and further upstream sequences might affect the transgene efficiency at the level of positive cells.

5.3. The OR101-1 core promoter contains O/E and homeodomain binding motifs

With the impulse of these results, the first 2 kb of sequence upstream of the OR101-1 TSS was subjected to bioinformatics analysis to find out if it contains any relevant cis-regulatory motifs that may account for the observed high promoter activity. Within the 600 bps upstream of TSS, which are included in the highly efficient 1.2 kb promoter, a single Olf1/Ebf1-like site and 3 homeodomain binding sites could be identified. The single O/E site is located 88 bp, while the most proximal homeodomain-like site can be found 215 bp upstream of the TSS (Figure 4.6A).

Interestingly, the most proximal homeodomain-like site is located at the breakpoint between constructs $p_{0.8kb}$OR101-1::EYFP-pA and $p_{1.0kb}$OR101-1::EYFP-pA. The exclusion

of this site may account for the slight but significant reduction in the number of transgene expressing cells for construct $p_{0.8kb}$OR101-1::EYFP-pA ($p<0.01$; Figure 4.14). Indeed, homeodomain-binding sites were reported to modulate the expression probability of a murine OR gene in OSNs (Vassalli *et al.*, 2011). However, a rather high frequency of transgene expression was observed both in cells and embryos for the $p_{0.8kb}$OR101-1::EYFP-pA construct that lacks the homeodomain-binding site. Thus, the site does not appear to be essential for the control of OR101-1 expression.

However, a caveat concerning the validity of the bioinformatic analysis of candidate binding sites is that the exact binding motifs for many fish-derived transcription factors are not known. On a similar note, there are no homologues of the OR101-1 gene locus in other fish species a comparison of which might have pointed towards evolutionary conserved regulatory sites. Therefore, the bioinformatic analysis had to be restricted to transcription factor binding sites that are well defined in other vertebrates and under the assumption that the motif is conserved in zebrafish.

5.4. The proximal O/E site upstream of the TSS is dispensable for promoter function

Since the transgenic construct with the shortest upstream sequence maintained high promoter efficiency but only contained a single candidate O/E-like site at position -88 bp, its possible role as a *cis*-regulatory motif was tested with site-directed mutagenesis. The limited bioinformatics analysis could not identify additional sequences resembling O/E-like sites within the 600 bp of promoter sequence that was tested and the next O/E-like site is located 1kb upstream of the TSS.

Yet, a loss of function mutation within the candidate O/E-site at position -88 showed that it is largely dispensable for high penetrance of expression, as the mutated construct was still expressed at 88% efficiency. However, a slight decrease in the number of transgene-expressing OSNs could be observed for the construct carrying the mutated site and which reached significance under some conditions ($p<0.001$; Figure 4.15). This

observation suggests that the proximal upstream O/E-like site could function synergistically with additional *cis*-regulatory motifs located within sequences upstream and/or downstream of the TSS. Similar observations have been made for the mouse M71 promoter (Rothman et al., 2005), where redundant O/E sites contributed to M71 expression. However, at least one O/E-like site had to be preserved in the M71 promoter and reduced transgenic constructs with similar point mutations that rendered the single O/E site functionless showed no expression.

The fact that mutation of the proximal upstream O/E motif did not prevent high efficiency of transgene expression suggests that additional O/E sites might be located further downstream of the TSS. All constructs for which the promoter efficiencies have been discussed so far contained 588 bps downstream of OR101-1 TSS, including 196 bp of a 5'-noncoding exon, a 345 bp 5'-intron and additional 47 bp of noncoding sequence in front of the coding sequence in exon 2 (Figure 4.2A).

5.5. The first intron of OR101-1 is critical for gene expression

To test if sequences downstream of the TSS contribute to promoter function, a transgenic construct was generated in which all the sequences downstream of TSS were deleted (Figure 4.8A). Thus, the construct effectively included only 600 bp of sequence upstream of the TSS. Interestingly, a severe reduction in efficiency, down to 42 % penetrance of expression, could be observed when sequences downstream of the TSS were removed (Figure 4.8C). However, in this construct, all sequences downstream of the TSS were excluded and it raises the question, whether relevant sequences are contained within the intron or the UTR sequences.

A similar reduction, down to 45 % penetrance of expression, was observed when only the intronic sequence was removed, suggesting that relevant regulatory sites may reside in the 5'-intron but not the UTR. To understand if the 5'-intronic sequence indeed containes any *cis*-activatory site, another bioinformatics search was conducted in which the

sequences downstream of the TSS were screened for the presence of O/E-like sites (Figure 4.10A). Interestingly, an additional O/E-like site could be identified 394 bp downstream of TSS, thus, residing within 5'-intronic sequence (Figure 4.10B).

However, as previously mentioned, only the mouse definition of the O/E-like motif was known and could be used to scan zebrafish sequences. But this fuelled the question if the recognition motif for O/E proteins is conserved in fish and if it may be conserved more among fish as compared to mammals. To circumvent this problem, a phylogenetic footprinting was performed on the candidate OMP gene promoter from 14 different fish species, since the O/E-like sites were initially identified and characterized in proximal promoter of OMP gene (Kudrycki *et al.*, 1993). The O/E binding site being obtained in this analysis was slightly different from that defined in mammals. Then, the 1.2 kb promoter was scanned by using the two available sequence matrices for the murine (MA0154.1) and human (MA0154.2) and the newly identified fish motif. Interestingly, the proximal O/E-like site that was identified by the murine matrix and shown by mutagenesis to be largely dispensable for high penetrance of expression (88 %) was not recognized by the human- (MA0154.2) or fish-specific matrix (Figure 4.11). Thus, the previously identified motif at position -88 might not be functional in fish, which could explain the observed phenotype for the mutated $p_{1.2kb\text{-}\Delta O/E}$OR101-1::EYFP-pA construct.

However, interestingly, the O/E-like site which was identified within the 5'-intronic sequence by the murine matrix, was also recognized by the newly formed fish-specific matrix (Figure 4.11). Therefore, the exclusion of this site in $p_{0.6kb\text{-}\Delta UTR\text{-}\Delta intron}$OR101-1::EYFP-pA and $p_{0.8kb\text{-}\Delta intron}$OR101-1::EGFP-pA could be the reason for their low transgene efficiencies observed in embryos (42 % and 45 %) and cells ($p < 10^{-16}$; Figure 4.14 and 4.15), compared to the full construct $p_{1.2kb}$OR101-1::EYFP-pA (81 %).

5.6. Regulatory sites in introns or regulation by intron splicing?

The results presented above suggest that critical regulatory sites may be located within the 5'-intron as removal of the intronic sequence or removal of the entire sequence downstream of the TSS resulted in a 50% reduction of transgene efficiency. This raises the question whether sequences contained within the intron or splicing of the intron itself contributes to improved expression. It is well known that splicing increases both nuclear export (Luo and Reed 1999; Zhou *et al*., 2000; Le Hir *et al*., 2001) as well as translation in the cytoplasm (Braddock *et al*. 1994; Matsumoto *et al*., 1998; Nott *et al*., 2003). Thus, different amounts of proteins made from the transgenic construct could be directly related to the observed efficiency of the transgenic construct, as low translation would result in weak reporter gene expression. However, in the analysis of the intronless construct, cells strongly expressing the reporter gene could be observed, although their proportion was lower in the population, compared to the intron-containing constructs.

However, the effect of splicing on protein translation can be quite variable and depend strongly on the promoter and the gene that is tested (Buchman and Berg 1988; Bourdon *et al*. 2001). The effect of intron splicing on OR gene promoter function is not known. An interesting way to test whether the absence of splicing causes the reduction in transgene expression would be to mutate the candidate O/E site without affecting relevant splice sites in the intron, or to re-introduce the O/E site upstream of the intronless construct.

Interestingly, both O/E-like and homeodomain-binding sites were identified within the second intron of a murine OR gene (MOR23; Vassalli et al., 2002). In that study, not any significant difference was reported to be observed in the expression levels of intron2-deleted MOR23 transgenes in OSNs. Actually, a MOR23 transgene including a deletion of all the 5'-upstream sequences up to a breakpoint being located upstream of the identified O/E-like and homeodomain-binding sites in intron 2 was referred to be expressed in OSNs which could converge their axons to the Olfactory Bulb as an unpublished data. However, since it was previously reported that MOR23 gene can be alternatively transcribed from a

TSS in exon2 in testicular cells (Asai *et al.*, 1996), the relevant intronic sequence was suggested to function as an alternate promoter (Vassalli et al., 2002). Similarly, it could be suggested that the 5'-intronic sequence being located downstream of OR101-1 TSS and that contains a candidate O/E-like site might function as an alternate promoter. However, arguing against this, our results suggest that this 5'-intronic sequence might constitute a critical part of the proximal promoter of zebrafish OR101-1 gene. In the analyzes of 600 bps of sequences upstream of OR101-1 TSS contained within highly efficient 1.2 kb promoter (81 %), the deletion of the first 388 bp from the 5'end did not result in any drastic decrease in promoter efficiency (75 % ; $p_{0.8kb}$OR101-1::EYFP-pA). Within the only remaining 212 bps of sequence of upstream of OR101-1 TSS contained in this deleted construct ($p_{0.8kb}$OR101-1::EYFP-pA), a single candidate O/E-like (-88bp) was identified. However, the lack of mutation of this proximal O/E-site did not generate any dramatic effect to the high efficiency of 1.2 kb promoter. When both of these results are assessed with the drastic decrease (50 %) observed in promoter efficiency of 1.2 kb construct due to the deletion of the 5'-intronic sequence, this suggests that the candidate O/E-like site identified within 5'-intronic sequence might be a critical part of the proximal promoter of OR101-1 gene.

5.7. No effect of O/E sites or no effect of Olf1/Ebf1?

The presence of O/E-like sites in the promoter sequence suggests that the Olf1/Ebf1 proteins may be critically involved in transcriptional regulation of zebrafish OR 101-1 gene. To analyze the functional roles of Olf1/Ebf1 proteins, Morpholino Antisense Oligos were utilized as a knock down strategy. However, the occurrence of genomic duplication in fish after the phylogenetic divergence from mammals generated gene redundancy in the zebrafish genome (Nowak *et al.*, 1997; Postlethwait *et al.*, 2000). Thus, there are many zebrafish orthologs of a particular mammalian gene, which makes its functional analysis difficult and complicated in zebrafish (Yoshihara, 2009). For the mammalian Ebf1 gene, there are two orthologs (Ebf1a and Ebf1b) in zebrafish. RNA-Seq data showed that Ebf1b is primarily expressed in olfactory tissue in zebrafish whereas Ebf1a is expressed primarily in brain (Table 4.1.); therefore Ebf1b was selected as an initial target to knock down with

MO. Initially, the transgene efficiency of 1.2 kb promoter construct ($p_{1.2kb}$OR101-1::EYFP-pA) was analyzed within the embryos co-injected with the Ebf1b-translation-blocking MO. In the preliminary results obtained for the tested concentrations of MO, no difference in transgene efficiency could be observed between the MO-injected and control embryos (Figure 4.12).

Although not verified with control experiments, if it was assumed that Ebf1b proteins could be efficiently knocked down in these experiments, these preliminary results would suggest that the lack of Ebf1b proteins does not generate any significant effect on transgene expression. However, the functional redundancy by the activity of other zebrafish orthologs of Ebf1b protein could also be the reason for the observed phenotype. In order to elucidate if Ebf1 proteins have any functional role in regulation OR101-1 gene expression, experimental conditions still need to be optimized. In addition, like in all previous analyzes with transgenic constructs, OMP promoter reference plasmid was used in these MO-injections as a positive control for injection efficiency. However, proximal OMP promoter also contains an O/E-like site; therefore another reference plasmid which does not contain any O/E-like site should be used to clearly observe the effect of lack of Ebf1b-proteins in expression of the transgene.

In summary, the overall analysis of sequences upstream and downstream of OR101-1 TSS within 1.2 kb proximal promoter suggest that the candidate upstream and downstream *cis*-regulatory motifs might function synergistically to control OR101-1 gene expression. However, these motifs might differ in function. Along with the *cis*-regulatory motifs whose functional activity is critical to promote OR101-1 expression, activity of the other motifs might only be restricted in regulation of the probability of this expression. To find out the exact role of one of these motifs in control of OR101-1 gene expression, the O/E-like site identified within downstream of OR101-1 TSS is currently being tested by site-directed mutagenesis in our lab.

APPENDIX A: EQUIPMENT

Table 6.1. Equipment.

4 °C Room	Birikim Elektrik, Turkey
Autoclaves	Astell Scientific, UK
Centrifuge	Eppendorf, Germany (5417R)
Confocal Microscope	Leica SP5-AOBS, USA
Electronic Balance	Sartorius, Germany (TE412)
Electrophoresis Supplies	Bio-Rad Labs, USA (ReadySub-Cell GT Cells)
Fluorescence Microscope	Leica Microsystems, USA (MZ16FA)
Freezer 1 -20 °C	Arçelik, Turkey
Freezer 2 -80 °C	Thermo Electron Corp., USA (Farma 723)
Gel Documentation	Bio-Rad Labs, USA (GelDoc XR)
Glass Bottles	Isolab, Germany
Incubator 1	Weiss Gallenkamp, UK
Incubator 2	Nuve, Turkey
Incubating Shaker	Thermo Electron Corp., USA
Micropipetters	Eppendorf, Germany (Research)
Microwave Oven	Vestel, Turkey
Microinjector	Eppendorf, Germany (FemtoJet)
Luminometer	Fluroskan Ascent Fl (Thermo Scientific)

Table 6.1. Equipment (cont.).

Refrigerator	Arçelik, Turkey
Softwares	Vector NTI (Invitrogen, USA)
Thermal Cyclers	Bio-Rad Labs, USA (C1000)
Vortex	Scientific Industries, USA

APPENDIX B: SUPPLIES

Table 6.2. List of Supplies.

1 kb DNA Ladder	New England Biolabs, U.S.A. (N3232)
100 bp DNA Ladder	New England Biolabs, U.S.A. (N3231)
5X GoTaq Flexi Buffer	Clontech, U.S.A. (639201)
Advantage 2 Polymerase Mix	Promega, U.S.A. (M890A)
BamHI	New England Biolabs, U.S.A. (R0136 L)
Bovine Serum Albumin	New England Biolabs, U.S.A. (B9001)
DMEM / F12 medium	Gibco, U.S.A. (11880).
EcoRI	New England Biolabs, U.S.A. (R0101 M)
EcoRV	New England Biolabs, U.S.A. (R0195 L)
Ethanol Absolute	Sigma-Aldrich, U.S.A. (34870)
Ethidium Bromide	Sigma Life Sciences, U.S.A. (E1510-1 ml)
EDTA Disodium Salt	Sigma-Aldrich., U.S.A. (E5134 - 1 kg).
Glycerol	Sigma-Aldrich, U.S.A. (G5516-500 ml)
GoTaq Flexi DNA Polymerase	Promega, U.S.A. (M830B)
KpnI	New England Biolabs, U.S.A. (R0558 L)

Table 6.2. List of Supplies (cont.).

LB Agar	Sigma Life Sciences, U.S.A. (SL08394)
LB Broth	Sigma-Aldrich, U.S.A. (L7658- 1 kg)
Magnesium Chloride, 25 mM	Promega, U.S.A. (A3511)
Magnesium Sulfate	Sigma-Aldrich, U.S.A. (M7506)
NcoI	New England Biolabs, U.S.A. (R0193 L)
NotI	New England Biolabs, U.S.A. (R0189 L)
pGEM®-T Easy Vector System	Promega, U.S.A. (A1360)
Potassium Chloride	Sigma-Aldrich, U.S.A. (P9541)
PstI	New England Biolabs, U.S.A. (R0140 L)
SalI	New England Biolabs, U.S.A. (R0138 L)
SeaKem® Agarose	Cambrex, U.S.A. (50004)
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 (R0133 L)
SphI	New England Biolabs, U.S.A (R0182 L)

Table 6.2. List of Supplies (cont.).

T4 DNA Ligase	New England Biolabs, U.S.A (M0202L)
Trizma® Base	Sigma-Aldrich, U.S.A. (T6066)
XhoI	New England Biolabs, U.S.A. (R0146 L)
X-treme® Transfection Reagent	Roche, Germany (04709691001)

APPENDIX C: RESULTS FOR INDIVIDUAL EXPERIMENTS

Table 6.3. Number of Transgenic Cells.

		p101-1-1.2kb_eYFP_pA			
survival ratio	fish	mCherry	eYFP	eYFP/mCherry	
67/84	F1	1	1	1	
	F2	6	3	0,5	
	F3	4	0	0	
	F4	3	3	1	
	F5	3	3	1	
	F6	4	1	0,25	
	F7	2	0	0	
	F8	1	1	1	
	F9	1	1	1	
	F10	3	2	0,666666667	
	F11	2	2	1	
	F12	2	2	1	
	F13	2	2	1	
	F14	1	0	0	
23/39	F15	3	2	0,666666667	
	F16	12	12	1	
	F17	10	10	1	
	F18	16	14	0,875	
	F19	6	5	0,833333333	
	F20	5	7	1,4	
	F21	2	4	2	
	F22	6	5	0,833333333	
	F23	5	5	1	
	F24	16	16	1	
	F25	11	14	1,272727273	
	F26	6	6	1	
	F27	5	6	1,2	
	F28	3	3	1	
	F29	2	2	1	
	F30	6	9	1,5	
	F31	1	1	1	
58/135	F32	3	3	1	
	F33	2	1	0,5	
	F34	2	2	1	
	F35	4	4	1	
	F36	1	1	1	
	F37	3	2	0,666666667	
	F38	3	2	0,666666667	
	F39	2	2	1	
	F40	2	2	1	
	F41	1	0	0	
	F42	2	0	0	
	F43	1	1	1	

Table 6.3. Number of Transgenic Cells (cont.).

	F44	1	1	1	
	F45	2	1	0,5	
	F46	1	0	0	
	F47	5	2	0,4	
	F48	2	1	0,5	
	F49	1	0	0	
	F50	2	2	1	
	F51	1	1	1	
	F52	4	4	1	
	F53	1	1	1	
	F54	1	1	1	
	F55	2	1	0,5	
	F56	1	0	0	
	F57	3	3	1	
	F58	1	1	1	
	F59	2	2	1	
	F60	7	7	1	
	F61	2	2	1	
61/97	F62	8	6	0,75	
	F63	11	6	0,545454545	
	F64	12	7	0,583333333	
	F65	10	8	0,8	
	F66	15	8	0,533333333	
	F67	2	6	3	
	F68	3	4	1,333333333	
	F69	3	3	1	
	F70	1	1	1	
	F71	2	0	0	
	F72	2	2	1	
	F73	9	6	0,666666667	
	F74	6	10	1,666666667	
	F75	3	3	1	
	F76	8	6	0,75	
	F77	4	5	1,25	
	F78	1	0	0	
	F79	5	0	0	
	F80	3	0	0	
	F81	4	3	0,75	
	F82	2	0	0	
	F83	5	3	0,6	
	F84	7	8	1,142857143	
	F85	6	4	0,666666667	
	F86	4	0	0	
	F87	5	3	0,6	
	F88	3	0	0	
	F89	3	0	0	
	F90	1	1	1	
	F91	1	1	1	
	F92	3	0	0	
	F93	1	0	0	
	F94	5	5	1	

Table 6.3. Number of Transgenic Cells (cont.).

F95	3	3	1
F96	6	4	0,666666667
F97	4	4	1
F98	3	0	0
F99	5	2	0,4
F100	5	4	0,8
F101	4	2	0,5
eYFP	/mCherry	fish:	82/101
avg:	3,99009901	3,168317	0,759762762
SD:	3,339146746	3,319847	0,491433573
SEM:	0,33225752	0,330337	0,048899468

Table 6.3. Number of Transgenic Cells (cont.).

		p(1.0_OR101-1_eYFP_pA)			
survival ratio	fish	mCherry	eYFP	eYFP/mCherry	
16/50	F1	4	4	1	
	F2	1	1	1	
	F3	1	0	0	
	F4	5	3	0,6	
	F5	5	2	0,4	
	F6	2	0	0	
	F7	1	1	1	
	F8	3	2	0,666667	
	F9	2	2	1	
	F10	3	0	0	
	F11	3	3	1	
10/?	F12	12	7	0,583333	
	F13	10	3	0,3	
	F14	2	2	1	
	F15	10	10	1	
	F16	3	3	1	
	F17	2	2	1	
	F18	8	8	1	
	F19	6	6	1	
	F20	4	4	1	
19/?	F21	4	4	1	
	F22	1	1	1	
	F23	1	0	0	
	F24	2	2	1	
	F25	3	4	1,333333	
	F26	1	1	1	
	F27	2	2	1	
	F28	1	1	1	
	F29	1	1	1	
	F30	2	2	1	
64/?	F31	1	1	1	
	F32	1	0	0	
	F33	3	1	0,333333	
	F34	1	1	1	
	F35	3	0	0	
	F36	1	1	1	
	F37	3	3	1	
	F38	3	3	1	
	F39	4	3	0,75	
	F40	1	5	5	
	F41	1	0	0	
	F42	5	3	0,6	
	F43	4	4	1	
	F44	2	0	0	
	F45	2	2	1	
	F46	8	8	1	
	F47	1	0	0	
	F48	1	1	1	
	F49	2	0	0	
	F50	4	3	0,75	

Table 6.3. Number of Transgenic Cells (cont.).

	F51	2	2	1	
	F52	7	4	0,571429	
	F53	5	2	0,4	
	F54	1	1	1	
	F55	4	4	1	
	F56	1	1	1	
	F57	4	2	0,5	
	F58	4	4	1	
	F59	2	1	0,5	
	F60	2	1	0,5	
	F61	2	2	1	
99/159	F62	2	2	1	
	F63	10	6	0,6	
	F64	3	1	0,333333	
	F65	3	3	1	
	F66	7	6	0,857143	
	F67	1	1	1	
	F68	3	2	0,666667	
	F69	16	3	0,1875	
	F70	16	4	0,25	
	F71	6	4	0,666667	
	eYFP	/mCherry	fish:	61/71	
	avg:	3,619718	2,478873	0,779569	
	SD:	3,283579	2,103461	0,633007	
	SEM:	0,389689	0,249635	0,075124	

Table 6.3. Number of Transgenic Cells (cont.).

		p(0.8_OR101-1_eYFP_pA)			
survival ratio	fish	mCherry	eYFP	eYFP/mCherry	
18/60	F1	4	4	1	
	F2	2	0	0	
	F3	3	0	0	
	F4	3	0	0	
	F5	4	0	0	
	F6	3	3	1	
	F7	9	0	0	
	F8	1	0	0	
58/90	F9	6	3	0,5	
	F10	1	2	2	
	F11	3	1	0,333333	
	F12	2	0	0	
	F13	2	0	0	
	F14	1	1	1	
	F15	3	2	0,666667	
	F16	2	2	1	
	F17	1	1	1	
	F18	8	3	0,375	
	F19	9	8	0,888889	
	F20	8	4	0,5	
	F21	12	12	1	
	F22	2	0	0	
	F23	6	5	0,833333	
	F24	11	10	0,909091	
	F25	14	10	0,714286	
	F26	4	4	1	
	F27	2	1	0,5	
	F28	2	1	0,5	
	F29	3	3	1	
	F30	4	4	1	
	F31	4	2	0,5	
	F32	3	3	1	
	F33	2	2	1	
	F34	3	1	0,333333	
	F35	2	2	1	
	F36	1	1	1	
	F37	2	2	1	
	F38	1	1	1	
	F39	10	10	1	
	F40	5	3	0,6	
	F41	2	2	1	
46/162	F42	2	2	1	
	F43	4	2	0,5	
	F44	9	9	1	
	F45	4	4	1	
	F46	1	1	1	
	F47	1	0	0	
	F48	2	0	0	
	F49	1	1	1	
	F50	1	0	0	

Table 6.3. Number of Transgenic Cells (cont.).

	F51	1	1	1	
	F52	1	1	1	
	F53	1	1	1	
	F54	1	0	0	
	F55	5	6	1,2	
57/83	F56	4	1	0,25	
	F57	7	1	0,142857	
	F58	5	1	0,2	
	F59	5	3	0,6	
	F60	14	8	0,571429	
	F61	4	1	0,25	
	F62	5	2	0,4	
	F63	5	2	0,4	
	F64	9	3	0,333333	
	F65	8	3	0,375	
	F66	5	0	0	
	F67	10	8	0,8	
	F68	5	1	0,2	
	F69	8	0	0	
	F70	4	4	1	
	F71	4	1	0,25	
	F72	2	0	0	
	F73	2	0	0	
	F74	3	0	0	
	F75	6	2	0,333333	
	F76	8	1	0,125	
	F77	7	0	0	
	F78	4	4	1	
	F79	2	0	0	
	F80	7	5	0,714286	
	eYFP	/mCherry	fish:	60/80	
	avg:	4,3375	2,4	0,55999	
	SD:	3,185981	2,772377	0,445293	
	SEM:	0,356204	0,309961	0,049785	

Table 6.3. Number of Transgenic Cells (cont.).

		p1.2_Δ_UTR_intron_eYFP_pA			
survival ratio	**fish**	**mCherry**	**eYFP**	**eYFP/mCherry**	
65/98	F1	1	0	0	
	F2	15	0	0	
	F3	1	0	0	
	F4	9	0	0	
	F5	6	5	0,833333	
	F6	17	0	0	
	F7	4	1	0,25	
	F8	3	0	0	
	F9	1	0	0	
	F10	2	0	0	
	F11	15	5	0,333333	
	F12	17	1	0,058824	
	F13	7	0	0	
	F14	9	2	0,222222	
	F15	2	0	0	
	F16	4	0	0	
	F17	5	0	0	
	F18	1	0	0	
	F19	4	2	0,5	
	F20	2	0	0	
	F21	8	0	0	
	F22	3	0	0	
	F23	5	0	0	
	F24	2	0	0	
	F25	7	1	0,142857	
	F26	6	1	0,166667	
	F27	3	0	0	
	F28	5	0	0	
	F29	15	0	0	
	F30	5	3	0,6	
	F31	13	0	0	
	F32	3	0	0	
	F33	1	0	0	
	F34	1	0	0	
	F35	2	0	0	
	F36	4	0	0	
	F37	4	1	0,25	
	F38	6	1	0,166667	
24/75	F39	5	0	0	
	F40	10	1	0,1	
	F41	18	0	0	
	F42	7	1	0,142857	
	F43	13	0	0	
	F44	9	0	0	
	F45	5	0	0	
	F46	11	0	0	
	F47	4	1	0,25	
	F48	7	1	0,142857	
	F49	11	0	0	
	F50	7	1	0,142857	

Table 6.3. Number of Transgenic Cells (cont.).

	F51	2	0	0	
	F52	6	4	0,666667	
	F53	26	7	0,269231	
	F54	2	0	0	
	F55	5	5	1	
	F56	10	3	0,3	
	F57	3	0	0	
	F58	7	5	0,714286	
34/59	F59	7	4	0,571429	
	F60	13	1	0,076923	
	F61	7	0	0	
	F62	5	0	0	
	F63	9	0	0	
	F64	12	0	0	
	F65	8	0	0	
	F66	3	0	0	
	F67	19	10	0,526316	
	F68	16	5	0,3125	
	F69	2	1	0,5	
	F70	17	3	0,176471	
	F71	10	1	0,1	
	F72	12	0	0	
	F73	19	10	0,526316	
	F74	7	0	0	
	F75	20	8	0,4	
	F76	3	0	0	
	F77	15	4	0,266667	
	F78	20	1	0,05	
	F79	17	1	0,058824	
	F80	11	3	0,272727	
	F81	2	0	0	
	F82	15	5	0,333333	
	F83	6	0	0	
	F84	11	0	0	
	eYFP	/mCherry	fish:	35/84	
	avg:	8	1,297619	0,136002	
	SD:	5,741416	2,274954	0,221287	
	SEM:	0,62644	0,248218	0,024144	

Table 6.3. Number of Transgenic Cells (cont.).

		p1.2-Δ_INTRON_eGFP_pA			
survival ratio	fish	mCherry	eYFP	eYFP/mCherry	
44/47	F1	5	0	0	
	F2	2	0	0	
	F3	8	0	0	
	F4	9	0	0	
	F5	23	3	0,130435	
	F6	5	1	0,2	
	F7	8	0	0	
	F8	11	0	0	
	F9	9	3	0,333333	
	F10	4	0	0	
	F11	12	0	0	
	F12	3	0	0	
	F13	12	0	0	
	F14	2	0	0	
	F15	3	0	0	
	F16	3	0	0	
	F17	4	0	0	
	F18	2	0	0	
	F19	2	1	0,5	
	F20	1	1	1	
	F21	2	0	0	
	F22	8	1	0,125	
	F23	2	1	0,5	
40/85	F24	9	0	0	
	F25	6	0	0	
	F26	4	0	0	
	F27	5	0	0	
	F28	11	5	0,454545	
	F29	12	5	0,416667	
	F30	11	2	0,181818	
	F31	2	0	0	
	F32	6	6	1	
	F33	5	0	0	
	F34	5	0	0	
	F35	7	1	0,142857	
	F36	5	1	0,2	
	F37	18	3	0,166667	
	F38	4	0	0	
	F39	4	0	0	
	F40	4	0	0	
	F41	22	11	0,5	
	F42	17	3	0,176471	
	F43	19	7	0,368421	
63/93	F44	28	8	0,285714	
	F45	18	2	0,111111	
	F46	7	1	0,142857	
	F47	24	5	0,208333	
	F48	17	1	0,058824	
	F49	8	0	0	
	F50	5	0	0	

Table 6.3. Number of Transgenic Cells (cont.).

F51	6	0	0
F52	10	0	0
F53	17	4	0,235294
F54	6	0	0
F55	14	9	0,642857
F56	6	2	0,333333
F57	8	1	0,125
F58	15	8	0,533333
F59	32	17	0,53125
F60	3	0	0
F61	4	2	0,5
F62	1	0	0
F63	6	0	0
F64	5	1	0,2
F65	3	0	0
F66	8	0	0
eYFP	/mCherry	fish:	30/66
avg:	8,590909	1,757576	0,156123
SD:	6,845498	3,172287	0,235557
SEM:	0,842623	0,390482	0,028995

Table 6.3. Number of Transgenic Cells (cont.).

survival ratio	fish	p(1.2_OR101-1_mO/E_eYFP_pA) mCherry	eYFP	eYFP/mCherry
42/54	F1	2	1	0,5
	F2	7	7	1
	F3	3	2	0,666667
	F4	13	9	0,692308
	F5	27	24	0,888889
	F6	4	4	1
	F7	10	6	0,6
	F8	3	0	0
	F9	7	6	0,857143
	F10	7	4	0,571429
	F11	1	0	0
	F12	16	12	0,75
	F13	6	3	0,5
	F14	2	2	1
	F15	9	7	0,777778
	F16	1	1	1
51/58	F17	22	16	0,727273
	F18	3	0	0
	F19	25	19	0,76
	F20	21	14	0,666667
	F21	12	8	0,666667
	F22	2	2	1
	F23	6	3	0,5
	F24	5	3	0,6
	F25	3	3	1
	F26	9	4	0,444444
	F27	10	10	1
	F28	8	1	0,125
	F29	13	10	0,769231
	F30	10	7	0,7
	F31	9	5	0,555556
	F32	14	6	0,428571
	F33	12	2	0,166667
	F34	18	14	0,777778
	F35	5	5	1
36/45	F36	6	4	0,666667
	F37	16	14	0,875
	F38	6	3	0,5
	F39	1	0	0
	F40	8	7	0,875
	F41	1	1	1
	F42	9	6	0,666667
	F43	2	2	1
	F44	12	0	0
	F45	9	4	0,444444
	F46	22	18	0,818182
	F47	9	4	0,444444
	F48	10	2	0,2
	F49	4	3	0,75
	F50	4	1	0,25

Table 6.3. Number of Transgenic Cells (cont.).

	F51	4	4	1	
	F52	23	17	0,73913	
61/68	F53	2	2	1	
	F54	3	3	1	
	F55	1	1	1	
	F56	4	0	0	
	F57	1	0	0	
	F58	1	1	1	
	F59	5	3	0,6	
	F60	2	2	1	
	F61	6	5	0,833333	
	F62	2	0	0	
	F63	6	2	0,333333	
	F64	1	1	1	
	F65	7	4	0,571429	
	eYFP	/mCherry	fish:	57/65	
	avg:	7,876923	5,138462	0,634765	
	SD:	6,522815	5,341105	0,332502	
	SEM:	0,809056	0,662483	0,041242	

Table 6.3. Number of Transgenic Cells (cont.).

		p101-1-1.2kb_eYFP_pA			Exp. Rate
		O/E-TB-MO			
survival ratio	fish	mCherry	eYFP	eYFP/mCherry	
14/33	F1	ND	ND	ND	
	F2	ND	ND	ND	
with MO	F3	4	2	0,5	
~ 4 ng	F4	23	24	1,043478261	
	F5	ND	ND	ND	
EXP1	F6	ND	ND	ND	
	F7	13	11	0,846153846	
	F8	ND	ND	ND	
	F9	ND	ND	ND	100
13/20	F1	ND	ND	ND	
	F2	ND	ND	ND	
only plasmid	F3	4	0	0	
	F4	ND	ND	ND	
Control	F5	ND	ND	ND	
	F6	ND	ND	ND	
	F7	40	37	0,925	
	F8	9	5	0,555555556	
	F9	31	22	0,709677419	
	F10	ND	ND	ND	
	F11	ND	ND	ND	91
36/77	F1	12	13	1,083333333	
	F2	11	10	0,909090909	
with MO	F3	7	3	0,428571429	
~ 4 ng	F4	2	1	0,5	
	F5	7	5	0,714285714	
Exp2	F6	16	16	1	
	F7	14	18	1,285714286	
	F8	4	1	0,25	
	F9	10	6	0,6	
	F10	12	9	0,75	
	F11	19	0	0	
	F12	4	1	0,25	
	F13	19	18	0,947368421	
NOT ANY	CONTROL	for Exp2	xxxxxxxxx	xxxxxxxxxx	

Table 6.3. Number of Transgenic Cells (cont.).

20/120	F1	3	1	0,333333333	
	F2	2	1	0,5	
with MO	F3	1	1	1	
~ 8 ng	F4	5	5	1	
	F5	10	4	0,4	
EXP3	F6	6	5	0,833333333	
	F7	5	4	0,8	
	F8	1	1	1	
	F9	1	1	1	100
38/59	F1	4	3	0,75	
	F2	4	3	0,75	
only plasmid	F3	8	0	0	
	F4	ND	ND	ND	
Control	F5	2	0	0	
	F6	13	7	0,538461538	
	F7	2	0	0	
	F8	6	4	0,666666667	
	F9	8	4	0,5	
	F10	ND	ND	ND	
	F11	ND	ND	ND	
	F12	12	9	0,75	75

REFERENCES

Ahuja, G., S. B. Nia, V. Zapilko, V. Shiriagin, D. Kowatschew, Y. Oka, and S. I. Korsching, 2014, "Kappe Neurons, A Novel Population of Olfactory Sensory Neurons.", *Scientific Reports*, Vol. 4, , pp. 4037.

Alioto, T. S., and J. Ngai, 2005, "The Odorant Receptor Repertoire of Teleost Fish.", *BMC Genomics*, Vol. 6, , pp. 173.

Asai, H., H. Kasai, Y. Matsuda, N. Yamazaki, F. Nagawa, H. Sakano, and A. Tsuboi, 1996, "Genomic Structure and Transcription of a Murine Odorant Receptor Gene: Differential Initiation of Transcription in the Olfactory and Testicular Cells.", *Biochemical and Biophysical Research Communications*, Vol. 221, No. 2, pp. 240–247.

Barth, A. L., N. J. Justice, and J. Ngai, 1996, "Asynchronous Onset of Odorant Receptor Expression in the Developing Zebrafish Olfactory System.", *Neuron*, Vol. 16, No. 1, pp. 23–34.

Barth, A. L., J. C. Dugas, and J. Ngai, 1997, "Noncoordinate Expression of Odorant Receptor Genes Tightly Linked in the Zebrafish Genome", *Neuron*, Vol. 19, No. 2, pp. 359–369.

Belluscio, L., G. Koentges, R. Axel, and C. Dulac, 1999, "A Map of Pheromone Receptor Activation in the Mammalian Brain.", *Cell*, Vol. 97, No. 2, pp. 209–220.

Bourdon, V., A. Harvey, and D. M. Lonsdale, 2001, "Introns and Their Positions Affect the Translational Activity of mRNA in Plant Cells.", *EMBO Reports*, Vol. 2, No. 5, pp. 394–398.

Bozza, T., A. Vassalli, S. Fuss, J. J. Zhang, B. Weiland, R. Pacifico, P. Feinstein, and P. Mombaerts, 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*, Vol. 61, No. 2, pp. 220–233.

Braddock, M., M. Muckenthaler, M. R. White, A. M. Thorburn, J. Sommerville, A. J. Kingsman, and S. M. Kingsman, 1994, "Intron-Less RNA Injected into the Nucleus of Xenopus Oocytes Accesses a Regulated Translation Control Pathway.", *Nucleic Acids Research*, Vol. 22, No. 24, pp. 5255–5264.

Buchmant, A. R., and P. Berg, 1988, "Comparison of Intron-Dependent and Intron-Independent Gene Expression", *Molecular and Cellular Biology*, Vol. 8, No. 10, pp. 4395–4405.

Buck, L., and R. Axel, 1991, "A Novel Multigene Family May Encode Odorant Receptors: A Molecular Basis for Odor Recognition.", *Cell*, Vol. 65, No. 1, pp. 175–187.

Buck, L. B., 2000, "The Molecular Architecture of Odor and Pheromone Sensing in Mammals.", *Cell*, Vol. 100, No. 6, pp. 611–618.

Buck, L. B., 2004, "Olfactory Receptors and Odor Coding in Mammals.", *Nutrition Reviews*, Vol. 62, No. 11 Pt 2, pp. S184–S188; discussion S224–S241.

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., I. Simon, H. Cedar, and R. Axel, 1994, "Allelic Inactivation Regulates Olfactory Receptor Gene Expression", *Cell*, Vol. 78, No. 5, pp. 823–834.

Clowney, E. J., A. Magklara, B. M. Colquitt, N. Pathak, R. P. Lane, and S. Lomvardas, 2011, "High-Throughput Mapping of the Promoters of the Mouse Olfactory Receptor Genes Reveals a New Type of Mammalian Promoter and Provides Insight into Olfactory Receptor Gene Regulation", *Genome Research*, Vol. 21, No. 8, pp. 1249–1259.

Dalton, R. P., D. B. Lyons, and S. Lomvardas, 2013, "Co-Opting the Unfolded Protein Response to Elicit Olfactory Receptor Feedback", *Cell*, Vol. 155, No. 2, pp. 321–332.

Dillon, N., T. Trimborn, J. Strouboulis, P. Fraser, and F. Grosveld, 1997, "The Effect of Distance on Long-Range Chromatin Interactions.", *Molecular Cell*, Vol. 1, No. 1, pp. 131–139.

Dhallan, R. S., Y. King-Wai, K. A. Schrader, and R. R. Reed, 1990, "Primary Structure and Functional Expression of a Cyclic Nucleotide-Activated Channel from Olfactory Neurons", *Nature*, Vol. 347, No. 6289, pp. 184–187.

Draper, B. W., P. A. Morcos, and C. B. Kimmel, 2001, "Inhibition of Zebrafish fgf8 Pre-mRNA Splicing with Morpholino Oligos: A Quantifiable Method for Gene Knockdown", *Genesis*, Vol. 30, No. 3, pp. 154–156.

Dulac, C., and R. Axel, 1995, "A Novel Family of Genes Encoding Putative Pheromone Receptors in Mammals.", *Cell*, Vol. 83, No. 2, pp. 195–206.

Feinstein, P., and P. Mombaerts, 2004, "A Contextual Model for Axonal Sorting into Glomeruli in the Mouse Olfactory System", *Cell*, Vol. 117, No. 6, pp. 817–831.

Furger, A., J. M. O'Sullivan, A. Binnie, B. A. Lee, and N. J. Proudfoot, 2002, "Promoter Proximal Splice Sites Enhance Transcription", *Genes and Development*, Vol. 16, No. 21, pp. 2792–2799.

Fuss, S. H., M. Omura, and P. Mombaerts, 2007, "Local and Cis Effects of the H Element on Expression of Odorant Receptor Genes in Mouse", *Cell*, Vol. 130, No. 2, pp. 373–384.

Fuss, S. H., Y. Zhu, and P. Mombaerts, 2013, "Odorant Receptor Gene Choice and Axonal Wiring in Mice with Deletion Mutations in the Odorant Receptor Gene SR1", *Molecular and Cellular Neuroscience*, Vol. 56, pp. 212–224.

Glusman, G., I. Yanai, I. Rubin, and D. Lancet, 2001, "The Complete Human Olfactory Subgenome", *Genome Research*, Vol. 11, No. 5, pp. 685–702.

Hagman, J., C. Belanger, a. Travis, C. W. Turck, and R. Grosschedl, 1993, "Cloning and Functional Characterization of Early B-Cell Factor, a Regulator of Lymphocyte-Specific Gene Expression", *Genes and Development*, Vol. 7, No. 5, pp. 760–773.

Hagman, J., M. J. Gutch, H. Lin, and R. Grosschedl, 1995, "EBF Contains a Novel Zinc Coordination Motif and Multiple Dimerization and Transcriptional Activation Domains.", *EMBO Journal*, Vol. 14, No. 12, pp. 2907–2916.

Hansen, A., S., Rolen, K., Anderson, Y., Morita, J., Caprio, and T., Finger, 2003, "Correlation between Olfactory Receptor Cell Type and Function in the Channel Catfish.", *Journal of Neuroscience*, Vol. 23, No. 28, pp. 9328-9339.

Hansen, A., K. Anderson, and T. E. Finger, 2004, "Differential Distribution of Olfactory Receptor Neurons in Goldfish: Structural and Molecular Correlates.", *Journal of Comparative Neurology,* Vol. 477, No. 4, pp. 347–359.

Herrada, G., and C. Dulac, 1997, "A Novel Family of Putative Pheromone Receptors in Mammals with a Topographically Organized and Sexually Dimorphic Distribution", *Cell*, Vol. 90, No. 4, pp. 763–773.

Hirota, J., and P. Mombaerts, 2004, "The LIM-Homeodomain Protein Lhx2 Is Required for Complete Development of Mouse Olfactory Sensory Neurons.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 101, No. 23, pp. 8751–8755.

Hirota, J., M. Omura, and P. Mombaerts, 2007, "Differential Impact of Lhx2 Deficiency on Expression of Class I and Class II Odorant Receptor Genes in Mouse", *Molecular and Cellular Neuroscience*, Vol. 34, No. 4, pp. 679–688.

Hoppe, R., H. Breer, and J. Strotmann, 2006, "Promoter Motifs of Olfactory Receptor Genes Expressed in Distinct Topographic Patterns", *Genomics*, Vol. 87, No. 6, pp. 711–723.

Hoppe, R., H. Frank, H. Breer, and J. Strotmann, 2003, "The Clustered Olfactory Receptor Gene Family 262: Genomic Organization, Promotor Elements, and Interacting Transcription Factors", *Genome Research*, Vol. 13, No. 12, pp. 2674–2685.

Imai, T., M. Suzuki, and H. Sakano, 2006, "Odorant Receptor-Derived cAMP Signals Direct Axonal Targeting.", *Science (New York, N.Y.)*, Vol. 314, No. 5799, pp. 657–661.

Ishii, T., S. Serizawa, A. Kohda, H. Nakatani, T. Shiroishi, K. Okumura, Y. Iwakura, F. Nagawa, A. Tsuboi, and H. Sakano, 2001, "Monoallelic Expression of the Odorant Receptor Gene and Axonal Projection of Olfactory Sensory Neurons.", *Genes to Cells*, Vol. 6, No. 1, pp. 71–78.

Jones, D. T., and R. R. Reed, 1989, "Golf: An Olfactory Neuron Specific-G Protein Involved in Odorant Signal Transduction.", *Science,* Vol. 244, No. 4906, pp. 790–795.

Kardash, E., 2012, "Current Methods in Zebrafish Research.", *MATER METHODS*, Vol. 2, No. 109.

Khan, M., E. Vaes, and P. Mombaerts, 2011, "Regulation of the Probability of Mouse Odorant Receptor Gene Choice.", *Cell*, Vol. 147, No. 4, pp. 907–921.

Kubick, S., J. Strotmann, I. Andreini, and H. Breer, 1997, "Subfamily of Olfactory Receptors Characterized by Unique Structural Features and Expression Patterns.", *Journal of Neurochemistry*, Vol. 69, No. 2, pp. 465–475.

Kudrycki, K., C. Stein-Izsak, C. Behn, M. Grillo, R. Akeson, and F. L. Margolis, 1993, "Olf-1-Binding Site: Characterization of an Olfactory Neuron-Specific Promoter Motif.", *Molecular and Cellular Biology*, Vol. 13, No. 5, pp. 3002–3014.

Le Hir, H., D. Gatfield, E. Izaurralde, and M. J. Moore, 2001, "The Exon-Exon Junction Complex Provides a Binding Platform for Factors Involved in mRNA Export and Nonsense-Mediated mRNA Decay", *EMBO Journal*, Vol. 20, No. 17, pp. 4987–4997.

Lewcock, J. W., and R. R. Reed, 2004, "A Feedback Mechanism Regulates Monoallelic Odorant Receptor Expression.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 101, No. 4, pp. 1069–1074.

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.

Lomvardas, S., G. Barnea, D. J. Pisapia, M. Mendelsohn, J. Kirkland, and R. Axel, 2006, "Interchromosomal Interactions and Olfactory Receptor Choice", *Cell*, Vol. 126, No. 2, pp. 403–413.

Luo, M. J., and R. Reed, 1999, "Splicing Is Required for Rapid and Efficient mRNA Export in Metazoans.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 96, No. 26, pp. 14937–14942.

Lyons, D. B., W. E. Allen, T. Goh, L. Tsai, G. Barnea, and S. Lomvardas, 2013, "An Epigenetic Trap Stabilizes Singular Olfactory Receptor Expression", *Cell*, Vol. 154, No. 2, pp. 325–336.

Magklara, A., A. Yen, B. M. Colquitt, E. J. Clowney, W. Allen, E. Markenscoff-Papadimitriou, Z. A. Evans, P. Kheradpour, G. Mountoufaris, … S. Lomvardas, 2011, "An Epigenetic Signature for Monoallelic Olfactory Receptor Expression", *Cell*, Vol. 145, No. 4, pp. 555–570.

Malnic, B., D. C. Gonzalez-Kristeller, and L. M. Gutiyama, 2010, "Odorant Receptors", *The Neurobiology of Olfaction*, pp. 181–202.

Malnic, B., J. Hirono, T. Sato, and L. B. Buck, 1999, "Combinatorial Receptor Codes for Odors.", *Cell*, Vol. 96, No. 5, pp. 713–723.

Markenscoff-Papadimitriou, E., W. E. Allen, B. M. Colquitt, T. Goh, K. K. Murphy, K. Monahan, C. P. Mosley, N. Ahituv, and S. Lomvardas, 2014, "Enhancer Interaction Networks as a Means for Singular Olfactory Receptor Expression", *Cell*, Vol. 159, No. 3, pp. 543–557.

Matsumoto, K., K. M. Wassarman, and A. P. Wolffe, 1998, "Nuclear History of a Pre-mRNA Determines the Translational Activity of Cytoplasmic mRNA", *EMBO Journal*, Vol. 17, No. 7, pp. 2107–2121.

McIntyre, J. C., S. C. Bose, A. J. Stromberg, and T. S. McClintock, 2008, "Emx2 Stimulates Odorant Receptor Gene Expression", *Chemical Senses*, Vol. 33, No. 9, pp. 825–837.

Michaloski, J. S., P. A. F. Galante, and B. Malnic, 2006, "Identification of Potential Regulatory Motifs in Odorant Receptor Genes by Analysis of Promoter Sequences.", *Genome Research*, Vol. 16, No. 9, pp. 1091–1098.

Michaloski, J. S., P. A. F. Galante, M. H. Nagai, L. Armelin-Correa, M. S. Chien, H. Matsunami, and B. Malnic, 2011, "Common Promoter Elements in Odorant and Vomeronasal Receptor Genes", *PLoS ONE*, Vol. 6, No. 12.

Miyamichi, K., S. Serizawa, H. M. Kimura, and H. Sakano, 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.", *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, Vol. 25, No. 14, pp. 3586–3592.

Mombaerts, P., F. Wang, C. Dulac, S. K. Chao, A. Nemes, M. Mendelsohn, J. Edmondson, and R. Axel, 1996, "Visualizing an Olfactory Sensory Map", *Cell*, Vol. 87, No. 4, pp. 675–686.

Mori, T., M. Sakai, I. Matsuoka, and K. Kurihara, 2000, "Analysis of Promoter Activity of 5'-Upstream Regions of Zebrafish Olfactory Receptor Genes", *Biological and Pharmaceutical Bulletin*, Vol. 23, no. 2, pp. 165–173.

Nathans, J., C. M. Davenport, I. H. Maumenee, R. A. Lewis, J. F. Hejtmancik, M. Lirr, E. Lovrien, R. Weleber, B. Bachynski, F. Zwas, R. Klingaman, and G. Fishman, 1989, "Molecular Genetics of Human Blue Cone Monochromacy.", *Science,* Vol. 245, no. 4920, pp. 831–838.

Nédélec, S., I. Foucher, I. Brunet, C. Bouillot, A. Prochiantz, and A. Trembleau, 2004, "Emx2 Homeodomain Transcription Factor Interacts with Eukaryotic Translation Initiation Factor 4E (eIF4E) in the Axons of Olfactory Sensory Neurons.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 101, No. 29, pp. 10815–10820.

Ngai, J., M. M. Dowling, L. Buck, R. Axel, and A. Chess, 1993, "The Family of Genes Encoding Odorant Receptors in the Channel Catfish.", *Cell*, Vol. 72, No. 5, pp. 657–666.

Nguyen, M. Q., Z. Zhou, C. A. Marks, N. J. P. Ryba, and L. Belluscio, 2007, "Prominent Roles for Odorant Receptor Coding Sequences in Allelic Exclusion", *Cell*, Vol. 131, No. 5, pp. 1009–1017.

Niimura, Y., and M. Nei, 2007, "Extensive Gains and Losses of Olfactory Receptor Genes in Mammalian Evolution", *PLoS ONE*, Vol. 2, No. 8.

Nishizumi, H., K. Kumasaka, N. Inoue, A. Nakashima, and H. Sakano, 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*, Vol. 104, No. 50, pp. 20067–20072.

Nott, A., S. H. Meislin, and M. J. Moore, 2003, "A Quantitative Analysis of Intron Effects on Mammalian Gene Expression.", *RNA (New York, N.Y.)*, Vol. 9, No. 5, pp. 607–617.

Nowak, M. A., M. C. Boerlijst, J. Cooke, and J. M. Smith, 1997, "Evolution of Genetic Redundancy.", *Nature*, Vol. 388, No. 6638, pp. 167–171.

Oka, Y., L. R. Saraiva, and S. I. Korsching, 2012, "Crypt Neurons Express a Single v1r-Related Ora Gene", *Chemical Senses*, Vol. 37, No. 3, pp. 219–227.

Postlethwait, J. H., I. G. Woods, P. Ngo-Hazelett, Y. L. Yan, P. D. Kelly, F. Chu, H. Huang, A. Hill-Force, and W. S. Talbot, 2000, "Zebrafish Comparative Genomics and the Origins of Vertebrate Chromosomes", *Genome Research*, Vol. 10, No. 12, pp. 1890–1902.

Qasba, P., and R. R. Reed, 1998, "Tissue and Zonal-Specific Expression of an Olfactory Receptor Transgene.", *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, Vol. 18, No. 1, pp. 227–236.

Ressler, K. J., S. L. Sullivan, and L. B. Buck, 1993, "A Zonal Organization of Odorant Receptor Gene Expression in the Olfactory Epithelium", *Cell*, Vol. 73, No. 3, pp. 597–609.

Rodriguez, I., 2013, "Singular Expression of Olfactory Receptor Genes.", *Cell*, Vol. 155, No. 2, pp. 274–7.

Rodriguez, I., P. Feinstein, and P. Mombaerts, 1999, "Variable Patterns of Axonal Projections of Sensory Neurons in the Mouse Vomeronasal System.", *Cell*, Vol. 97, No. 2, pp. 199–208.

Rothman, A., P. Feinstein, J. Hirota, and P. Mombaerts, 2005, "The Promoter of the Mouse Odorant Receptor Gene M71", *Molecular and Cellular Neuroscience*, Vol. 28, No. 3, pp. 535–546.

Sambrook, J., and D. W. Russell, 2001, Molecular Cloning: A Laboratory Manual. New York, USA, Cold Spring Harbor Laboratory.

Sancer, G., 2015, Regulation of Olfactory Receptor Gene Choice by Negative Regulatory Elements. M.Sc. Thesis. Boğaziçi University.

Sato, Y., N. Miyasaka, and Y. Yoshihara, 2005, "Mutually Exclusive Glomerular Innervation by Two Distinct Types of Olfactory Sensory Neurons Revealed in Transgenic Zebrafish.", *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, Vol. 25, No. 20, pp. 4889–4897.

Sato, Y., N. Miyasaka, and Y. Yoshihara, 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*, Vol. 27, No. 7, pp. 1606–1615.

Serizawa, S., T. Ishii, H. Nakatani, A. Tsuboi, F. Nagawa, M. Asano, K. Sudo, J. Sakagami, H. Sakano, … H. Sakano, 2000, "Mutually Exclusive Expression of Odorant Receptor Transgenes.", *Nature Neuroscience*, Vol. 3, No. 7, pp. 687–693.

Serizawa, S., K. Miyamichi, H. Nakatani, M. Suzuki, M. Saito, Y. Yoshihara, and H. Sakano, 2003, "Negative Feedback Regulation Ensures the One Receptor-One Olfactory Neuron Rule in Mouse.", *Science (New York, N.Y.)*, Vol. 302, No. 5653, pp. 2088–2094.

Shykind, B. M., S. C. Rohani, S. O'Donnell, A. Nemes, M. Mendelsohn, Y. Sun, R. Axel, and G. Barnea, 2004, "Gene Switching and the Stability of Odorant Receptor Gene Choice", *Cell*, Vol. 117, No. 6, pp. 801–815.

Söğünmez, N., 2012, Transgenic Analysis of the Zebrafish OR101-1 Gene Promoter. M.Sc. Thesis. Boğaziçi University.

Strotmann, J., S. Conzelmann, A. Beck, P. Feinstein, H. Breer, and P. Mombaerts, 2000, "Local Permutations in the Glomerular Array of the Mouse Olfactory Bulb.", *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, Vol. 20, No. 18, pp. 6927–6938.

Tastekin, İ., 2012, The Role of the Odorant Receptor Coding Sequence in the Regulation of Odorant Receptor Transgene Expression In Zebrafish. M. Sc. Thesis. Boğaziçi University.

Tian, H., and M. Ma, 2004, "Molecular Organization of the Olfactory Septal Organ.", *Journal of Neuroscience*, Vol. 24, no. 38, pp. 8383–8390.

Tinaztepe, E., 2009, Class Restriction in Odorant Receptor Gene Regulation. Boğaziçi University. M.Sc. Thesis. Boğaziçi University.

Tsuboi, A, S. Yoshihara, N. Yamazaki, H. Kasai, H. Asai-Tsuboi, M. Komatsu, S. Serizawa, T. Ishii, Y. Matsuda, … H. Sakano, 1999, "Olfactory Neurons Expressing Closely Linked and Homologous Odorant Receptor Genes Tend to Project Their Axons to Neighboring Glomeruli on the Olfactory Bulb.", *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, Vol. 19, No. 19, pp. 8409–8418.

Tsuboi, A., T. Miyazaki, T. Imai, and H. Sakano, 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*, Vol. 23, No. 6, pp. 1436–1444.

Uzel, K., 2014, Validation of the OR103-1/5 Intergenic Sequence as a Molecular Tool to Promote Bi-Cistronic Translation in Zebrafish. M.Sc. Thesis. Boğaziçi University.

Vassalli, A., P. Feinstein, and P. Mombaerts, 2011, "Homeodomain Binding Motifs Modulate the Probability of Odorant Receptor Gene Choice in Transgenic Mice", *Molecular and Cellular Neuroscience*, Vol. 46, No. 2, pp. 381–396.

Vassalli, A., A. Rothman, P. Feinstein, M. Zapotocky, and P. Mombaerts, 2002, "Minigenes Impart Odorant Receptor-Specific Axon Guidance in the Olfactory Bulb", *Neuron*, Vol. 35, No. 4, pp. 681–696.

Vassar, R., J. Ngai, and R. Axel, 1993, "Spatial Segregation of Odorant Receptor Expression in the Mammalian Olfactory Epithelium.", *Cell*, Vol. 74, pp. 309–318.

Wang, F., A. Nemes, M. Mendelsohn, and R. Axel, 1998, "Odorant Receptors Govern the Formation of a Precise Topographic Map", *Cell*, Vol. 93, No. 1, pp. 47–60.

Wang, M. M., and R. R. Reed, 1993, "Molecular Cloning of the Olfactory Neuronal Transcription Factor Olf-1 by Genetic Selection in Yeast.", *Nature*, Vol. 364, No. 6433, pp. 121–126.

Wang, S. S., J. W. Lewcock, P. Feinstein, P. Mombaerts, and R. R. Reed, 2004, "Genetic Disruptions of O/E2 and O/E3 Genes Reveal Involvement in Olfactory Receptor Neuron Projection.", *Development (Cambridge, England)*, Vol. 131, No. 6, pp. 1377–1388.

Wang, S. S., R. Y. Tsai, and R. R. Reed, 1997, "The Characterization of the Olf-1/EBF-like HLH Transcription Factor Family: Implications in Olfactory Gene Regulation and Neuronal Development.", *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, Vol. 17, No. 11, pp. 4149–4158.

Wang, Y., J. P. Macke, S. L. Merbs, D. J. Zack, B. Klaunberg, J. Bennett, J. Gearhart, and J. Nathans, 1992, "A Locus Control Region Adjacent to the Human Red and Green Visual Pigment Genes", *Neuron*, Vol. 9, No. 3, pp. 429–440.

Westerfield, M., 2007, "THE ZEBRAFISH BOOK; A Guide for the Laboratory Use of Zebrafish (Danio Rerio), 5th Edition.", Eugene: University of Oregon Press.

Weth, F., W. Nadler, and S. Korsching, 1996, "Nested Expression Domains for Odorant Receptors in Zebrafish Olfactory Epithelium.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 93, No. 23, pp. 13321–13326.

Yoshihara, Y., 2009, “Molecular Genetic Dissection of the Zebrafish Olfactory System.”, *Chemosensory Systems in Mammals, Fishes, and Insects*, Vol. 1, pp. 1– 24.

Zhang, X., and S. Firestein, 2002, "The Olfactory Receptor Gene Superfamily of the Mouse.", *Nature Neuroscience*, Vol. 5, No. 2, pp. 124–133.

Zhang, X., M. Rogers, H. Tian, X. Zhang, D.-J. Zou, J. Liu, M. Ma, G. M. Shepherd, and S. J. Firestein, 2004, "High-Throughput Microarray Detection of Olfactory Receptor Gene Expression in the Mouse.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 101, No. 39, pp. 14168–14173.

Zhang, X., X. Zhang, and S. Firestein, 2007, "Comparative Genomics of Odorant and Pheromone Receptor Genes in Rodents", *Genomics*, Vol. 89, No. 4, pp. 441–450.

Zhou, Z., M. J. Luo, K. Straesser, J. Katahira, E. Hurt, and R. Reed, 2000, "The Protein Aly Links Pre-Messenger-RNA Splicing to Nuclear Export in Metazoans.", *Nature*, Vol. 407, No. 6802, pp. 401–405.