TRANSGENIC ANALYSIS OF THE ZEBRAFISH OR101-1 GENE PROMOTER

by

Nuray Söğünmez

B.S., Molecular Biology and Genetics, Boğaziçi University, 2010

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

Graduate Program in Molecular Biology and Genetics
Boğaziçi University

2012

ACKNOWLEDGEMENTS

I would like to thank toStefan Fuss for his encouragement, support, and his guidance on our experiments. I am really lucky that I started to my academic career in Fish lab, and I appreciate that he believed in both me and my project.His valuable ideas made our lab to improve strikingly in such a short time.

I would like to thank to my project partner Xalid Bayramlı, for trusting me without hesitating and sharing his knowledge with me. I would like to thank to İbrahim Taştekin, for being our live-pubmed, and being a perfect friend. Also I would like to thank to my previous lab mate, Murat Atasoy, for always helping me to calm down when I panicked, and for helping me in my experiments, and in my life. I also want to thank to Kerem Uzel, for his shows in the lab, to Mustafa Uğur Baş, for his attention and care, to Büşra Çoban, for her kindness and for being a second female student in our Lab., andto Enes Karabacak, for his help.

I would like to thank to all current and former Fly Lab members, Prof. Arzu Çelik, Sercan Sayın, Mustafa Talay, Ece Terzioğlu-Kara, Güner Kaçmaz, Burak Tepe, Arzu Arat, Sandra Rode, Gamze Akgün, and Duygu Koldere and other undergrad.s for their nice accompaniments.

I would like to thank to Stefan Köstler for spending his days and nights to help me, and for his patience, his care, his support, his reliability, his creativity, and his understanding. He is one of the bests that one can have, and I am very lucky to have such a good person in my life.

I would like to thank AK Lab members for their generosities, especially to Emine Dindar for her support and friendship, and Tuncay Şekerfor his ideas.

I would like to thank to İY Lab members, Zeynep Özcan, Nehir Banaz, Gizem Gül, and Tijen Bergin, and Gen-Reg member Ulduz Sobhi Afshar, for making my days melodic and cheerful.

I would like to thank to Hikmet Budak and Melda Kantar for helping me with my RACE experiments, and Emre Yakşi lab.members for showing me the miracles of electrophysiology.

I would like to thank to my friends Remzi Artar, Erkan Küçük, Burcu Ayhan, and YusufMeydancı for being with me whenever I needed them, and changing my mood when I was frustrated.

I would like to thank to my dear family, Nurten Söğünmez, Ahmet Ali Söğünmez and Turgay Söğünmez for supporting my all decisions, and always encouraging me to be the best whatever I started to do. Their supports are really important to me

ABSTRACT

TRANSGENIC ANALYSIS OF THE ZEBRAFISH OR101-1 GENE PROMOTER

The repertoire of olfactory receptors constitutes the largest gene families in humans and in other vertebrates. Typically, olfactory sensory neurons only express a single member of the large genomic repertoire and only one allele of it. The mechanisms by which olfactory sensory neurons choose to express an olfactory receptor gene for expression and by which other receptors are prevented from being expressed in the same cell are largely unknown. It was shown that translation of the olfactory receptor protein and accumulation in the cell membrane somehow prevents expression of other receptors. However, far less is known about how a sensory neuron initiates expression of a receptor. Regulatory elements located in distal or proximal regions of the receptor coding sequences might affect gene expression, either by selecting a single receptor gene from a larger gene cluster and activating its expression, or by the interaction of specific transcription factors with proximal promoter regions of a gene. Here studies on the transcriptional regulation of a model olfactory receptor, OR101-1, are presented in zebrafish. Four complementary approaches were used in this study to gain insight into OR101-1 gene expression:a BAC transgenesic approachwas utilized to demonstrate the OR101-1 expression profile, 5'-RLM RACE was performed to identify the TSS of the OR101-1 gene, and a promoter bashing approachwas employed to pinpoint specific sequences that control OR101-1 expression in a proximal promoter assay. To complement these studies, basic bioinformatic analysiswas performedon sequences that affect OR101-1 expression to determine candidate binding sites located within the OR101-1 promoter sequence. By using an array of different transgenic constructs and scoring the efficiencies of their expression in a transient transgenic assay, a 500bp region could be identified that is located between 2kb– 2.5kb upstream of the OR101-1 coding sequence, which, when present in transgenic constructs, decreases the efficiency of expression. Partial deletion of either the 5'- or 3'-half of this sequence had intermediate effects, suggesting that multiple binding sites for negative regulators might be present. Bioinformatics reveals that the candidate inhibitory sequence contains two occurrences of a zbtb binding motif, located in either half of the sequence. The results suggest that olfactory receptor gene expression is also controlled by, as of yet undefined, inhibitory mechanisms that might have important implications for the specificity of receptor gene expression.

ÖZET

ZEBRABALIĞI OR 101-1 GEN PROMOTÖRÜNÜN TRANSGENİK ANALİZİ

Koku reseptörü gen repertuarı insanlarda ve diğer omurgalılardaki en geniş gen ailelerinden birini oluşturmaktadır. Tipik olarak koku duyu nöronları geniş genomik repertuardaki tek bir geni ve onun tek alelini tanımlar.Bir hücre içinde koku reseptörü seçimini ve tanımlanmasını sağlayan ve diğer reseptörlerin ekspresyonunu durduran mekanizmalar henüz çok fazla bilinmemektedir.Koku reseptör proteininin translasyonunun ve hücre zarında birikmesinin bir şekilde diğer reseptörlerin ekspresyonunu engellediğin gösterilmiştir.Fakat, bir duyu nöronunun nasıl bir reseptörün ekspresyonunu başlattığı ile ilgili bilinenler daha da azdır.Reseptör kodlama sekanslarının distal ve proksimal bölgelerinde bulunan regulator elementlerin, ya geniş gen kümesinden bir reseptörü seçip active ederek, ya da genlerin proksimal promoter bölgelerindeki belirli transkripsiyon faktörleriyle etkileşerek gen ekspresyonunu etkileyebileceği öne sürülmektedir.Burada, zebrabalığındaki model koku reseptörü olan OR101-1 genini kullanarak, transkripsiyonel regülasyon konusundaki çalışmalarımsunulacaktır. Çalışmalarımda birbirini tamamlayan dört yaklaşım kullanarak OR101-1 gen ekspresyonunu derinlemesine incelenmiş: BAC transgenesis yaklaşımıyla OR101-1 ekspresyon profilini bulunmuş, 5'-RLM RACE metoduyla OR101-1 geninin transkripsiyon başlangıç noktasını belirlenmiş, ve 'promoter bashing' yaklaşımıyla OR101-1 gen ekspresyonunu kontrol eden spesifik dizileri proksimal promotör çalışmalarıyla işaretlenmiştir. Bu çalışmaları desteklemek için, OR101-1 ekspresyonunu etkileyen diziler üzerinde biyoinformatik analizler yaparak, OR101-1 promotör dizisinde bulunan aday bağlanma noktaları bulunmuştur. Geçici transgenyöntemiyle, bir dizi farklı transgenik yapılar kullanarak ve onların ekspresyon verimleriniskorlayarak, OR101-1 kodlama sekansının 2kb-2.5kb yukarısındaki 500bp bölgeyi, be bu bölgenin bulunduğu transgenik yapıların OR101-1 ekspresyonunu azalttığı keşfedilmiştir. Bu bölgenin 5' ve 3' yarısı kısmi olarak silindiğinde ise elde edilen ortalama ekspresyon seviyeleri, bu bölgede negatif regulatörlerin bağlanabileceği birden fazla bölgenin olabileceğini işaret etmiştir.Biyoinformatik analizlerde ise bu aday inhibitor bölgenin iki yarısında iki zbtb protein bağlanma motifinin bulunduğu görülmüştür.Bütün bu sonuçlar, koku reseptör geni ekspresyonunun henüz daha tam anlamıyla belirlenemeyen inhibitor mekanizmalarla ve bunların sağladığı reseptör gen ekspresyonu spesifitesiyle de control edilebileceğini göstermektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ ABBREVIATIONS

AOS	Accessory Olfactory System
BAC	Bacterial Artificial Chromosome
Bp	Base Pair
BSA	Bovine Serum Albumin
BTB-ZF	Bab-ttk-BR-C-Zinc Finger
cDNA	Complementary Deoxyribonucleic Acid
CIP	Calf Intestinal Phosphatase
DNA	Deoxyribonucleic Acid
FPR	Formyl Peptide Receptor
GFP	Green Fluorescent Protein
GPCR	G-protein-coupled Receptor
ISH	In Situ Hybridization
Kbp	Kilo base Pair
LCR	Locus Control Region
LH	Left Homology
mRNA	Messenger Ribonucleic Acid
MOR	Mouse Olfactory Receptor
MOS	Main Olfactory System
OB	Olfactory Bulb
OE	Olfactory Epithelium
OMP	Olfactory Marker Protein
OR	Olfactory Receptor

OSN	Olfactory Sensory Neuron
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
pOMP	Olfactory Marker Protein Promoter
RACE	Rapid Amplification of cDNA Ends
RH	Right Homology
RLM	RNA-ligase Mediated
RNA	Ribonucleic Acid
RT-PCR	Reverse-transcription Polymerase Chain Reaction
TAP	Tobacco Acid Pyrophosphatase
TetO	Tet Operator
TTA	Tetracycline Transactivators
TF	Transcription Factor
TFBS	Transcription Factor Binding Site
TSS	Transcription Start Site
UTR	Untranslated Region
Upstr	Upstream
VR	Vomeronasal Receptor
VNE	Vomeronasal Epithelium
YFP	Yellow Fluorescent Protein
ZOR	Zebrafish Olfactory Receptor

1.INTRODUCTION

1.1. The Olfactory System

1.1.1. General AnatomyofVertebrateOlfactory Systems

The olfactory system serves a variety of important functions in the life of an organism: I acts as elicitor of emotional responses and aggression, it is critical for the recognition of conspecifics and social interactions, it is implicated in prey and predator recognition, it regulates neuroendocrine function, and it has a role in the perception of flavors during food selection (Vincent and Lledo, 2001). All vertebrates share common characteristics in the organization of the olfactory system, both anatomically and molecularly. A major subdivision of olfactorysystems can be made between the main olfactory system (MOS) and the accessory olfactory system (AOS), which serve different aspects of olfactory function(Breer *et al.*, 2006; Ma, 2007)(Figure 1.1).

The sensory neurons of the AOS are located inside the bony capsule of the vomer bone in the roof of the plate, the vomeronasal organ, which is located far away from the main olfactory system and not dependent on the airflow through the nose(Meredith, 1991). The vomeronasal organ detects mostly pheromonecues through specialized classes of vomeronasal receptors (VRs) that are enriched inthe microvilli of the vomeronasal sensory neurons (Dulac and Axel, 1995; Herrada and Dulac 1997; Matsunami and Buck, 1997; Ryba and Tirindelli, 1997; Oikawa*et al.,* 1998). The vomeronasal organ consists of two distinct neuronal populations that are located either in the apical or basal zone of the accessory olfactory epithelium (AOE), respectively(Mori*et al.,* 2000), and which express different types of VRs as well as signaling molecules.

V1Rs are seven-transmembrane receptors expressed by the neurons located inthe apical layer of the vomeronasal epithelium (VNE). On the other hand, V2R genes encode seven transmembrane like receptors with a long extracellular N-terminus (Herrada and Dulac, 1997; Matsunami and Buck, 1997; Ryba and Tirindelli, 1997)andare expressed in

the basal zone of the VNE. Formyl-peptide receptors (FPRs) are encoded by the neurons located in the basal layer of theVNE and these neurons have been shown to mediateolfactoryresponses to infections of conspecifics(Riviere*et al.*, 2009).

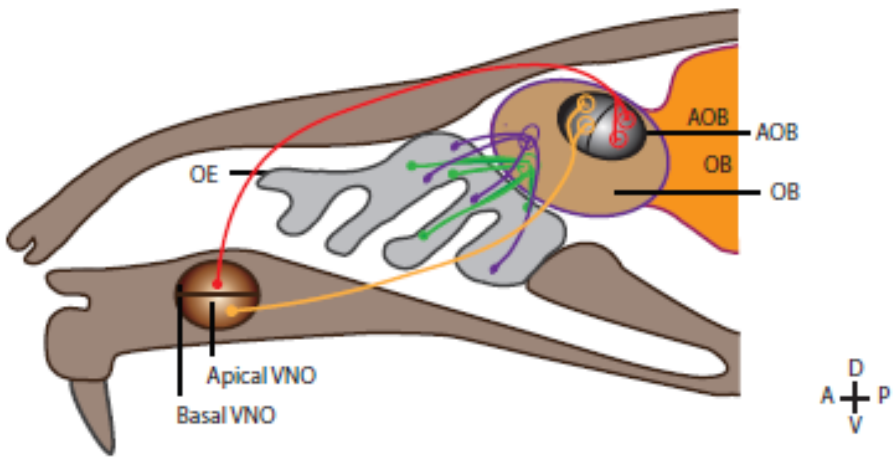


Figure 1.1.Structural Representation of the AOS and the MOS. In mice the VNO is spatially distinct from the MOE, and the neurons in the VNO project their axons to the AOB, located in the posterior-dorsal olfactory bulb (Derived from Brennan and Zufall, 2006;Pifferi *et al.*, 2009).

.

The main olfactory epithelium (MOE) lines the interior of the nasal cavity and is exposed to the major airflow through the nose. Historically, theMOEwas divided into four zones (from the dorso-medial zone I to the ventro-lateral zone IV) that were defined by the expression of distinct sets of OR genes (Ressler *et al.*, 1993; Vassar *et al.*, 1993; Sullivan *et al.*, 1995).In addition, the expression patterns of the OR37 (Strotmann *et al.*, 1992) and OR25 (Mori *et al.*, 2000) families indicated the presence of a fifth zone in the MOE. Recently, the zone concept has been challenged by the discovery that many OR genes are expressed in a continuous and overlapping expression domains along the dorso-ventral axis (Miyamichi *et al.*, 2006).

Typically, olfactory sensory neuronsexpressing the same OR gene, project their axons to a pair of defined glomeruli in the main olfactory bulb (MOB). The position of the

target glomeruli correlates with the zonal distribution of OSNs expressing the receptor in the olfactory epithelium (OE). There is a tendency that ORs with relatedspecificitiesare biased to be represented in the same zone (Mori *et al.*, 2000).

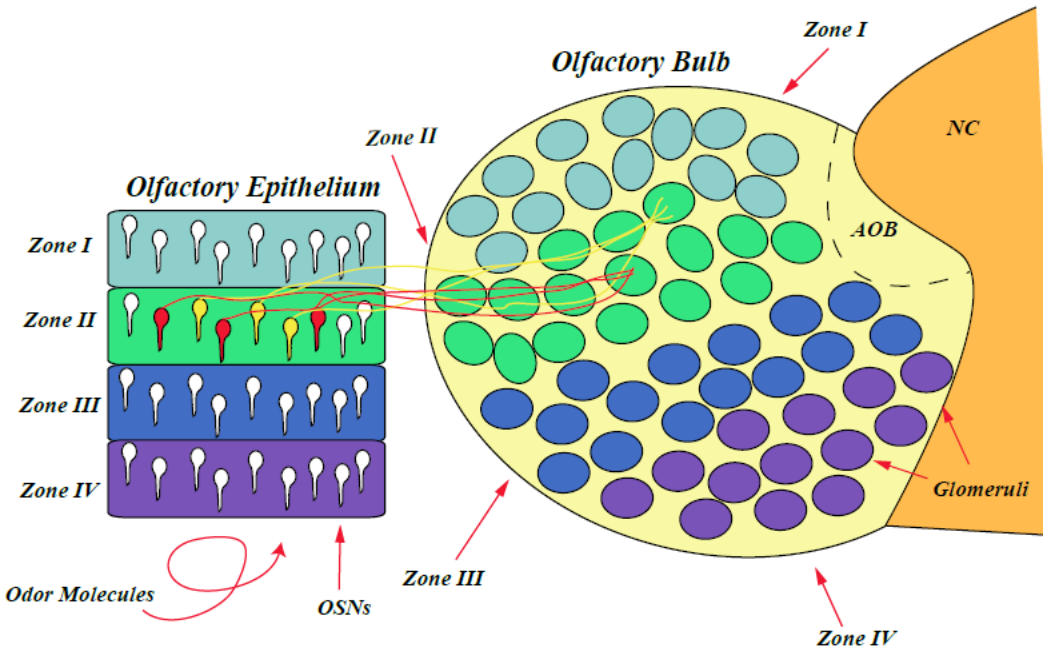


Figure 1.2. Zonal Expression and Signal Transduction in the MOS.The MOE is divided into four zones, which is also represented in the MOB. ORs expressing same receptor gene project their axons to the pair of glomeruli in approximately the same dorso-ventral position within the MOB (Derived From Mori, 1996).

Olfactory receptor neurons (ORNs) are the first-order neurons in the olfactory pathway. Each neuron extends a single dendrite towards the nasal cavity, directly interacting with the odor molecules through cilia at the tips of their dendrites, and projectsanaxon to the MOB of the forebrain. Here OSNs form synapses with mitral and tufted cells, which are the output neurons of the MOB and which relay odor signals to higher brain centers.

The relay of information from the MOB to the brain is highly regulated through modulatorinputfrom local cells in the MOB and by centrifugal projections from the brain. Target structures of projections from the MOB in the brain include, from rostral to caudal, the olfactory peduncle, pinform cortex, olfactory tubercle, enthorinal cortex, and amygdaloidal nuclei(Lotto and Price, 1994; Ressler *et al.*, 1994; Vassar *et al.*, 1994; Demir *et al.*, 1998; Wang *et al.*, 1998).

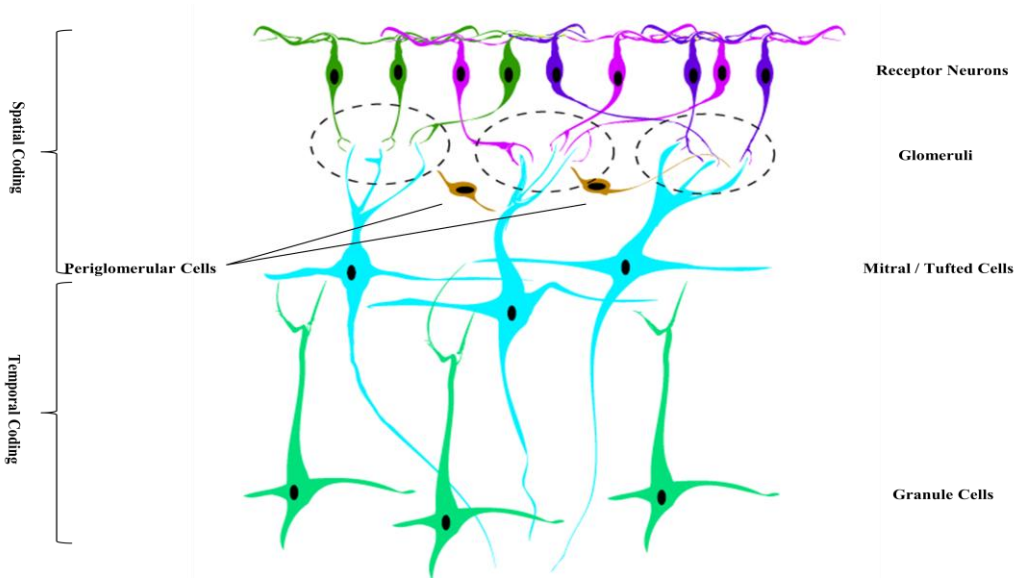


Figure 1.3.Synaptic Organization of the Main Olfactory Bulb. The axons of the OSNs expressing the same OR converge onto glomeruli in the OB, and mitral/tufted cells send signals from the MOB to the olfactory cortex after receiving inhibitory inputs from granule and periglomerular cells *(Derived from Brennan and Keverne, 1997).*

The vertebrate olfactory system is able to perceive and to discriminate an amazingly largenumberof odor molecules. Olfactory perception is initiated by the selective binding of odorant molecules to severalof many different ORs, which are Gprotein-coupled receptors, located on the surface of the ORNs (Buck and Axel, 1991; Krieger and Breer, 1999; Mombaerts, 1999). In the mouse about1400 mouse OR genes, including 20% pseudo genes, were identified from the Celera genome using a data mining strategy. These genes were classified into 228 gene families and are organized into 43 genomic clusters

distributed across the whole mouse genome, except for chromosome 12 and the Y chromosome (Zhang and Firestein, 2007). Humans on the other hand have lost about 60% of their functionalOR genes (Dugas and Ngai, 2001; Fuchs *et al.,* 2001; Zozulya *et al.,* 2001) whencompared to mice (Dugas and Ngai, 2001; Zhang and Firestein, 2002) andthus have a higher percentage of pseudo genes.

1.1.2. Zebrafish OR Anatomy and Physiology

The general anatomical and molecular organization of the olfactory system of the zebrafish (*Danio rerio*) is similar to that of other vertebrates. However, VR and OR receptor-expressing neurons are located within the same OE but project their axons to different sets of glomeruli, which occupy different regions of the OB. Anatomically, the zebrafish OE can be subdivided into two main parts: the midline raphe, which is surrounded by sensory lamellae (Hansen and Zeiske, 1998). Three distinct types of sensory neurons in the zebrafish OE have been reported, based on their morphological appearance and localizationalongthe apical-basal dimension of the OE (Hansen and Zeiske, 1998; Hansen and Finger, 2000; Hamdani and Doving, 2007). These cell types include ciliated cells expressing OR genes, microvillous cells expressing VR genes, and crypt cells expressing at least one V1R gene (Oka *et al.,* 2012). Ciliated cells reside predominantly in the basal epithelium and extend long dendrites to the surface of the epithelium, whereas crypt cells and microvillous cells are located more apically and have short or no dendrites, respectively (Hansen and Zeiske, 1998).

Zebrafish olfactory signal processing isalso similar to the processing described in rodents regarding the neuronal structures, the connectivity of axons and the glomerular organization. The glomerular pattern in zebrafish was identified by injecting the dye DiI into the nasal cavities of adult zebrafish and anterograde tracing of axon terminals in the ipsilateral OB (Baier and Korshing, 1994; Braubach *et al.,* 2012). Around 100 – 140 olfactory glomeruli could be identified invariantly in the adult zebrafish OB. About 20 - 27 of them are clearly distinguishable, whereas the remaining glomeruli are organized in tight clusters containing multiple glomeruli (Baier and Korshing, 1994; Braubach *et al.*, 2012).

1.1.3. Zebrafish, as a Model Organism

Given that the molecular and anatomical organization of the zebrafish olfactory system is very similar to that of higher vertebrates it constitutes an attractive and tractable model system to study olfactory processing and development. A distinct advantage over the mouse olfactory system is the reduced level of complexity, and the ease with which mutagenesis, transgenesis, gene knockdown, and transposon-mediated gene transfer can be conducted. In addition, the zebrafish system has distinct advantageous features, such as the transparency of embryos, external fertilization, and rapid embryonic development. Transgenes can be introduced into the genome by simple injection of DNA into fertilized oocytes. Much to experimental advantage, the size of the injected DNA can be quite large and up to 300kbcan successfully be integrated into the zebrafish genome using BAC transgenic approaches (Oka *et al.,* 2003; Yang *et al.,* 2006; Sato *et al.,* 2007).The discovery of the Tol2-transposable element (Kawakami and Shima, 1999; Kawakami *et al.,* 2004; Kwan *et al.,* 2007) and the*I-SceI* (Thermes *et al.* 2002) meganuclease significantly facilitated the stable introduction of transgenes into the zebrafish genome.

1.1.4. Structure of OR genes

ORs are seven-transmembrane domain containing G-protein-coupled receptors, expressed in ciliated neurons of the OE and they were first identified in the rat (Buck and Axel, 1991).The OR gene family is the largest gene family in any genome, comprising 143 genes in the zebrafish genome, around 1400 genes in the mouse and around 800 in human genomes (Mombaerts, 1999; Buck, 2000).The OR gene family appears to follow a birth-and-death model of evolution, where novel genes are formed by gene duplication and functional diversification of duplicated genes,while others are inactivated by becoming pseudo genes or deleted from the genome, depending on selective evolutionary pressure(Nei, 1969; Nei *et al.,* 1997). About 70% of human OR genes are pseudo genes, (Rouquier *et al.,* 2000; Dugas and Ngai, 2001), while this ratio is lower for fish and mouse with only 5-10 % pseudo genes (Rouquier *et al.,* 2000; Dugas and Ngai, 2001), probably because of their different evolutionary history.

Common features of typical vertebrate OR genes are single intronless ~ 1kb long coding sequences (Mombaerts,1999). Typically, the third, forth, and fifth transmembrane domains are highly variable when compared among many OR genes, implying that these regions are for the site of ligand binding (Buck and Axel, 1991; Buck, 2004).

1.1.5. Class Distinction in OR Genes

OR genes can be divided into two main subgroups, the class I and class II ORs, based on phylogenetic analysis of their protein similarities (Zhang and Firestein 2002; Alioto and Ngai, 2005; Niimura and Nei, 2005). Class I Ors are supposed to be evolutionary more ancient and are dominant in aquatic animals, whereas class II genes dramatically expanded in terrestrial vertebrates, constituting up to 90% of the mammalian OR gene repertoire.

In the mouse it has been reported that neurons expressing Class I OR genes are located exclusively in the most dorsal zone of the MOE and accordingly, they project their axons to the dorsal surface of the OB. On the other hand, neurons expressing class II OR genes are distributed across the entire MOE and project axons to glomeruli within the entire OB (Tsuboi *etal.,* 2006).There is experimental evidence, that the two classes of OR gens are expressed by distinct subpopulations of OSNs in the mouse, that are pre-determined for the expression of either class of OR gene (Bozza *et al.,* 2009).

Interestingly, almost all zebrafish OR genes, with the single exception of OR101-1, belong to class I Ors (Alioto and Ngai,2005; Niimura and Nei, 2005).A phylogenetic analysis of human and zebrafish OR genes reveals8 distinct phylogenetic clusters, based on their amino acid similarities. Among these branches, human Class I ORs cluster with the A branch of zebrafish OR genes, while the human Class II ORs cluster with only a single zebrafish OR homolog, OR101-1. The remaining zebrafish ORs are more distantly related to mammalian class I genes and can be divided into six phylogenetic branches(Alioto and Ngai,2005; Niimura and Nei, 2005).

The genomic organization of zebrafish OR genes shows that OR genes sharing sequence similarities of up to 60% are arranged incoherentclusters on the same chromosome in the zebrafish genome.

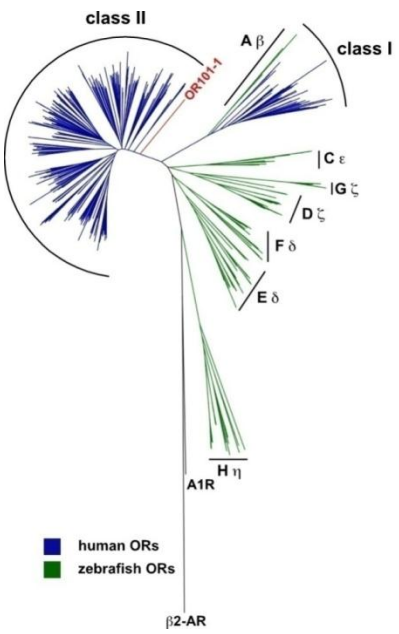


Figure 1.4.Phylogenetic Analysis of Amino acid Sequences from Human and Zebrafish OR Genes.Zebrafish ORs were subjected to phylogenetic analysis and compared to human OR genes. They can be categorized into 8 branches as A-H (Alioto and Ngai, 2005). All zebrafish ORs, except OR101-1, belong to the Class I, whereas OR101-1 belongs to the class II(Derived from Tınaztepe E., 2009).

1.1.6. Genomic Organization of OR101-1 Gene in Zebrafish

It was found by analyses of genomic data that the OR101-1 gene is located on the - strand of zebrafish chromosome 21 between the regions 42.178.376 – 42.179.326 (www.ensemble.org).The gene is 948bp in length and it locates next to a cluster of Class I ORs from the OR115 gene family. The first member of this gene family is a pseudo gene,

called OR115-3p, and locates in a reverse orientation downstream of OR 101-1 with 8kb distance.

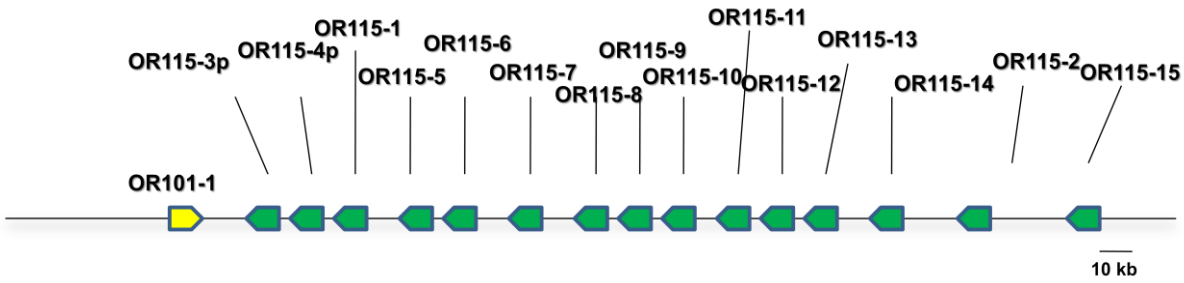


Figure 1.5.Genomic Organization of OR 101-1 Gene and OR115 Gene Family.The OR101-1 gene is located on chromosome 21 in reverse orientation to members of the OR115 gene family. OR 115 gene family consists of 13 full length OR genes and two pseudo genes. The closest OR 115 gene to OR 101-1 is OR115-3p pseudo gene, locates in the plus strand of zebrafish chromosome 21, and ~8kb downstream of OR 101-1 coding sequence.

1.1.7. OR101-1 Expression Pattern in Zebrafish Olfactory Epithelium

The OR101-1 gene is expressed in the (OE) of both adult and embryonic zebrafish, supportingthe fact that this gene is an OR gene. Further analysis of these results indicate that the receptors neurons expressing the OR101-1 gene are arranged in a concentric ring in the olfactory rosette in adult OE sections, comparable to other OR genes in zebrafish. Recent results from the laboratory showedthat OR101-1-expressing neurons might constitute a unique class of OSNs similar to the class I class II distinction in the mouse (Bayramli, personal communication).

1.1.8. Odorant Receptor (OR) Gene Expression Mechanism

1.1.8.1. Monogenic and Monoallelic Expression.It has been shown that OR genes are expressed in a monogenic fashion in OSNs (Malnic *et al.,* 1999; Serizawa *et al.,* 2000; Serizawa *et al.,* 2003), meaning that every single OSN expresses only a single gene of the large OR gene repertoire and only one allele of it(Chess *et al.,* 1994). The evidence comes from two experiments. In the first experiment mice from different backgrounds carrying polymorphic I7 alleles were crossed and single cells were used to conduct RT-PCR. The result showed that all of these individual cells did express the I7 alleles in a monoallelic fashion, and there was no preference for any of the parental alleles, in other words, they were expressed with equal frequencies (Chess *et al.,* 1994).

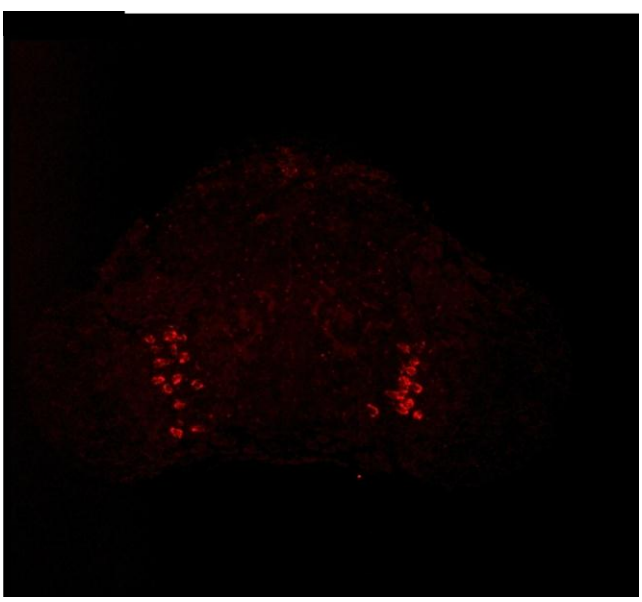


Figure 1.6. OR101-1 Whole Mount In-Situ Hybridization on 2dpf Zebrafish Embryo.In-situ hybridization and HNPP detection of OR101-1 expression on 2 dpf zebrafish embryos from the AB/AB wild-type strain. On average 12-14 cells are expressing the OR101-1 gene. (Bayramlı, unpublished)

In the second experiment, monoallelic expression of the M71 OR gene was demonstrated by tagging both parental alleles with different fluorescent proteins (Li *et al.,* 2004). Distinct OSN cell populations were observed which were positive for either one of the fluorescent proteins, in equal numbers.

Selection of a single gene among the large number of OR gene repertoire is not an easy task for an ORN, however, it was assumed that the genomic organization of OR genes as well as spatial and temporal expression patterns might have a role in the transcriptional regulations of the OR genes, as observed in other multigene families, such as Hox and β-globulin genes (Duboule, 1998). For instance Hox genes are also organized in clusters and it was found that the 3'- to 5'-organization of genes in the cluster resembles their anterior-posterior expression (Duboule,1998).

The 'one neuron-one receptor' rule describes the observation that each OSN only expresses a single OR gene from the large genomic repertoire. Even though there might be exceptions violating this rule (Rawson *et al.,* 2000; Sato *et al.,* 2007), the approximate expression frequency of each individual OR genes in mice is about 0.1% (Ressler *et al.,* 1993; Vassar *et al.,* 1993), given that there are around 1400 OR genes expressed in the mouse olfactory tissue, each cell should express on average only a single OR gene. In zebrafish the same approach has shown expression frequencies of 0.5–2%, which is consistent with the approximately 130 expressed OR genes (Chess *et al.,* 1992) and supports the idea that every OSN expresses only one OR in zebrafish as well.

1.1.8.2. Negative Feedback Mechanism and Second Choice.It appears that maintenance of expression form an OR locus depends on the protein product that is translated. There is accumulating evidence, that the presence of the OR protein in a sensory neuron prevents expression of other OR gens by a negative feedback mechanism (Qasba and Reed 1998; Serizawa *et al.,* 2003; Lewcock and Reed 2004; Li *et al*.,2004; Shykind *et al.,* 2004; Tsuboi *et al.,* 2006; Nguyen *et al.,*2007; Bozza *et al.*,2009).

This strongly resembles the situation in the immune system, where surface expression of B- and T-cell receptors prevents continuation of somatic recombination in lymphocytes (Nemazee *et al.,* 2002). Experiments in which frame shift-mutations were introduced into the coding sequence of OR genes or in which the entire coding sequence has been deleted, revealed that OSNs continue to express other OR genes during a process called ‘second choice’.

To better understand the feedback inhibition, additional experiments were conducted. The study claims that the repressor signal might act on the OR coding sequence at the DNA level but not at the OR promoters. To test this hypothesis OR proteins were expressed using the heterologous TetO promoter, which is active only when genetically encoded tetracycline transactivators (TTA) are present, which is expressed from a driver locus (Nguyen et al, 2009). Different transgenic mice were generated, expressing TTA under control of 3 different olfactory-specific promoters such as OMP, Gγ8 and the OR P2 promoters, and crossed with TetO-OR transgenic lines. Fluorescent reporter analysis revealed that neurons expressing the OR transgene under the TetO promoter never co-express endogenous ORs and vice versa, suggesting that expression of the OR protein is sufficient to repress other OR loci.

Likewise, transgenic constructs under control of the TetO promoter were repressed by expression of OR proteins from endogenous OR loci, suggesting that negative feedback signals act on the OR coding sequence rather than the OR promoter or the promoter-OR coding sequence fusion (Nguyen *et al.,* 2007).

1.1.9. Transcriptional Regulation of OR Genes

The transcriptional mechanisms behind the one neuron – one receptor rule are largely enigmatic. Studies revealed the presence of distal as well as the proximal elements regulating the OR gene expression either in an inhibitory or in an activatory manner (Serizawa *et al.* 2000; Vassalli *et al.,* 2002; Serizawa *et al.,*2003; Rothman *et al.,* 2005; Fuss *et al.,* 2007; Nishizumi *et al.,* 2007; Bozza*et al*., 2009; Khan *et al.,* 2011).

1.1.9.1. Regulation by Locus Control Regions.The presence of far distant sequences can have an important influence on the regulation of OR genes, a concept that is reminiscent of Locus Control Regions found in other systems. A YAC transgenesis approach using the*MOR28* gene cluster on chromosome14 *(Serizawa et al,.* 2000; Serizawa *et al.*,2003)revealed that the *MOR28, MOR10, MOR83, and MOR29A* genes require a 2.1kb DNA segment, called H-element, located about 85kbupstream of the*MOR28* transcription start site. It was hypothesized that the H-element also organizes expression of genes from the cluster by interacting with only one gene in the *MOR28* gene cluster at a time (Serizawa *et al*., 2003, 2004). Recently, other genomic elements with related function, equivalent to a locus control region, have been identified in the mouse (Khan *et al*., 2011), and zebrafish (Nishizumi *et al.,* 2009).

The evidence of zebrafish LCRs comes from the experiment where the promoter activation of OR111-1 and OR116-1 was analyzed. The coding sequences of these two genes were replaced with different fluorescent proteins inmodifiedtransgenic BAC constructs and the transient expressions were examined in zebrafish embryos. The comparison of expression patterns in different transgenic constructs bearing small deletions indicated that two LCRs/enhancers might be present in the analyzed OR gene cluster, which were termed as E15-1 and E15-2. However, these LCR/ enhancer sequences did not show any sequence similarity among each other or to the mouse H sequence. However, they share certain small sequence motifs, such as homeodomain-like sequences and OE-like sequences in the zebrafish enhancers (Nishizumi *et al.,* 2007).

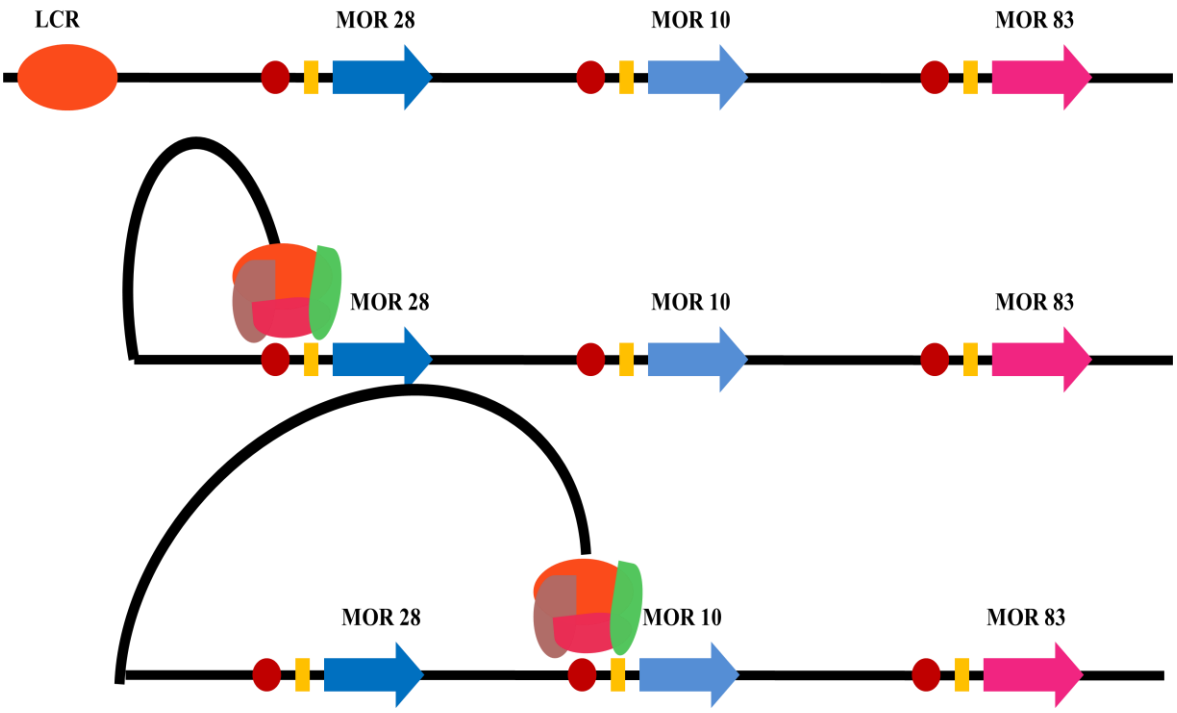


Figure 1.7.Proposed Model for LCR Mechanism of Action.LCRs effect gene expression by forming complexes with other regulatory proteins and interacting with promoter motives of targeted genes. In this figure, the mechanism of action of H-element is depicted. (Derived from Serizawa *et al.*,2003).

Yet another recently identified candidate for a cluster specific regulatory element is the P-element (Bozza *et al.*, 2009; Khan *et al.*, 2012). This element is 317bp in length and shows 70% similarity to the promoter of the closest neighboring gene P3. Recent studies show that P- and H-element share a common 13 nucleotide motif that is highly conserved and constitutesahomeodomain proteinbinding site (Vassalli *et al.*, 2011).

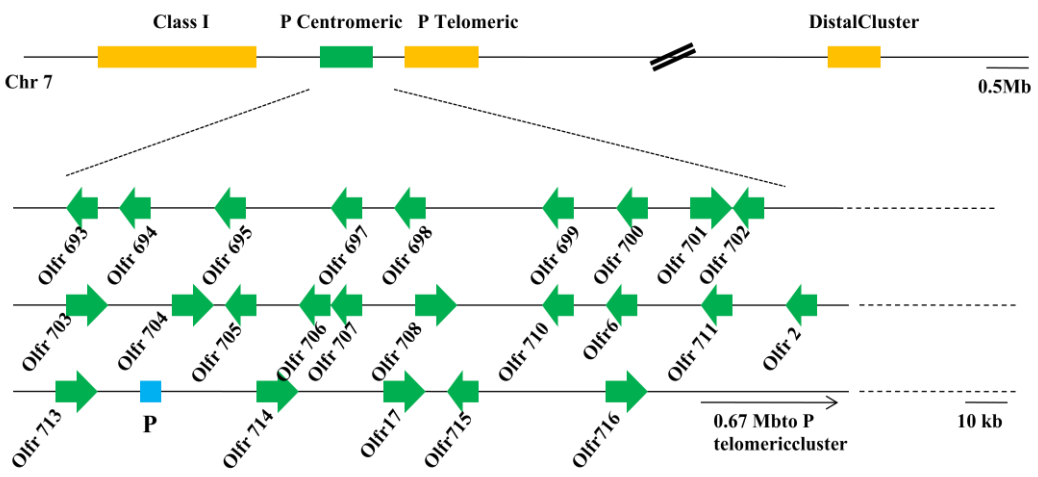


Figure 1.8.Genomic Localization of P Element.P element is located on the chromosome 7 in the mouse genome. Centromeric is to the left, telomeric to the right. The green arrows indicate the transcriptional orientation of OR genes (Derived from Khan et al., 2012).

1.1.9.2. Regulation by Proximal Promoter Elements.In addition to far distant genomic elements, conserved motifs have been identified in proximal promoter regions of OR genes (Touhara *et al.*, 1999; Bulger *et al.*, 2000; Hoppe *et al.*, 2000; Sosinsky *et al.*, 2000; Lane *et al.*, 2001). Interestingly, for some OR genes these short promoters are sufficient to recapitulate expression of the endogenous genes, as seen for the M71 and MOR23 genes. The relationship between those genes and genes controlled by locus control regions is not well understood. Thus, some OR genes might be controlled entirely by the short-range elements (Qasba and Reed, 1998), whereas others require long-range regulators (Serizawa *et al.*, 2000).

Genetically manipulated OR genes and minigenes comprising short genomic sequences were used to analyze OR gene promoter function (Vassalli *et al.*, 2002; Rothman *et al.*,2005; Vassalli *et al.*, 2011).Two mouse OR genes belonging to different gene clusters were used. Initially ~9kb genomic fragments, named as minigenes, of both MOR23 (Vassalli *et al.*, 2002) and M71 (Vassalli *et al.*, 201, Rothman *et al.*, 2005) were used to analyze the expression patterns. It was observed that transgene expression wasOSN specific, punctate, restricted to the correct zone, and likely singular (Vassalli *et al.*, 2002). The length of the minigene was modified to reveal the minimally required sequence that

drives OR gene expression in a fashion similar to the endogenous OR gene. Curiously, sequences containing only161bpafter the putative transcription start site (TSS)of the M71 OR gene were sufficient and no long distant enhancer was required (Rothman *et al.*, 2005).

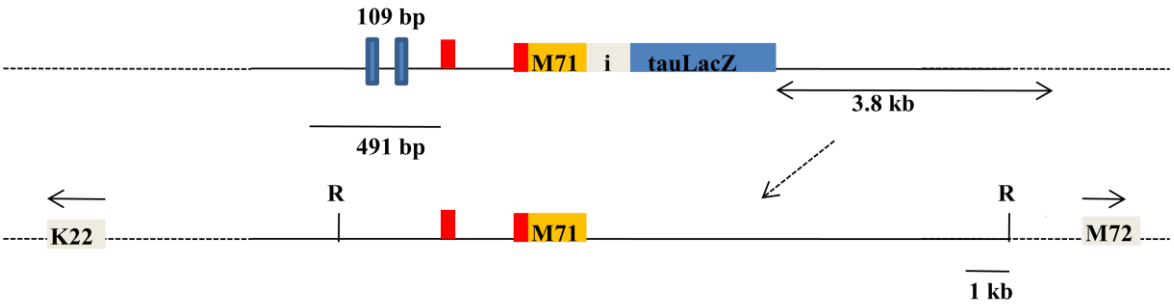


Figure1.9.Structure of Tg 491bp Deletion Construct from Whole Genomic Sequence.M71 gene locates between K22 and M72 locus. Red boxes represent the 5' UTR region. Blue boxes represent the Homeodomain and O/E-like Binding sites. The putative TSS is at 2.2kb upstream of the ATG (Vassalli *et al.*, 2002).

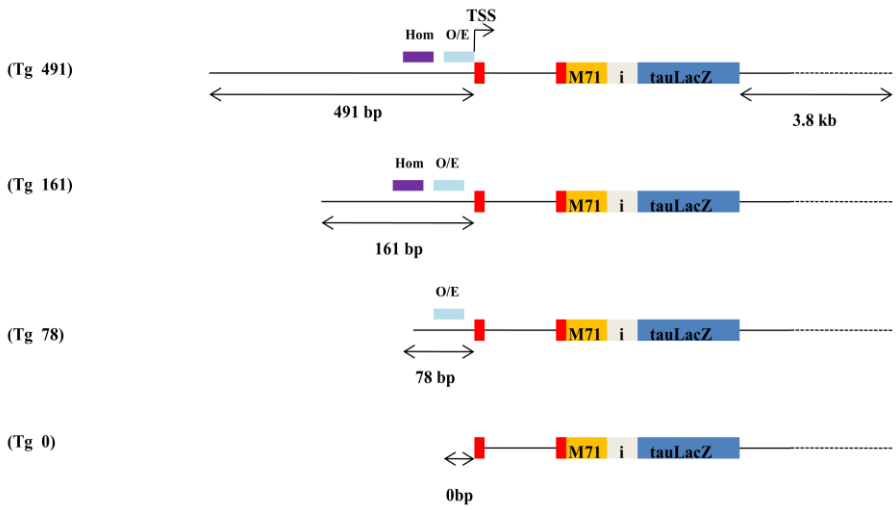


Figure 1.10.Proximal Promoter Analysis of M71 Gene.Transgenic deletion from Tg 491bp to Tg 0bpupstream of the TSS of M71 CDS. Homeodomain binding site and O/E-like site depicted as purple and light blue, respectively.

Moreover, sequence analyses revealed that *MOR23, M71* and other OR genes contain conserved homeodomain and O/E-like binding sites within these promoters, and in the analysis of a knockout mutation of the *lhx2* transcription factor encoding gene, which binds to the mouse M71 homeodomain site, it was shown that OR ClassII gene expression was completely abolished (Hirota and Mombaerts, 2004; Kolterud *et al.,* 2004), while ClassI OR genes were still expressed (Hirota *et al.,* 2007).

1.1.10. Bioinformatic Analysis Of OR Promoters and Potential TF-Binding Sites

Bioinformatics and experimental approaches have been used extensively to identify conserved sequence motifs within OR gene promoters (Clowney *et al.,*2011; Hoppe *et al.,* 2000; Sosinsky *et al.,* 2000; Lane *et al.,* 2001; Vassalli *et al.,* 2002; Hoppe *et al.,* 2003; Rothman *et al.,* 2005; Hoppe *et al.,* 2006; Michaloski *et al.,* 2006; Vassalli *et al.,* 2011).These studies showed that the candidate regions regulating specific OR expression typically locate at within short proximity to the TSSs of OR gens and thatORs sharing similar expression pattern have conserved structures and arrangements of these motifs (Strotmann *et al.,* 1994a*;* Kubick *et al.,* 1997; Strotmann*et al.,* 2000; Hoppe *et al.,* 2003).

There are ~1400 OR genes dispersed throughout the mouse genome (Young *et al.,* 2002; Zhang and Firestein, 2002; Godfrey *et al.,* 2004). One way of identification of potential regulatory cis-acting sequences in the genomic sequence is, to search for DNA sequence elements that are conserved in a large number of OR gene promoters. Promoter sequences, which are usually located proximal to and upstream of the TSS, were analyzed by 5'-RLM RACE method to obtain full length mRNA sequences for many OR genes, and the results were compared with their genomic counterparts (Michaloski *et al.,* 2006). The 5'-untranslated region (UTR) sequences can be highly variable for the OR genes, even though they belong to a specific subgroup. ORs in the mouse P2 OR cluster have 5'-UTRs that range from 1.7 to 9kb (Lane *et al.,* 2001).

It was shown, that a collection of random OR genes will have the same types of cis-regulatory elements, as homeodomain motifs and O/E like sites, suggesting that these

common promoter elements might play an important role in OR gene expression (Michaloski *et al.,* 2006). It is also possible that enhancers or LCRs interact with elements in one OR gene promoter to select that specific OR for expression. A striking fact shows ,that the H region, which works as an LCR and is located 75kb upstream of the MOR28 gene cluster (Serizawa *et al.*, 2003), also contains at least one set of homeodomain- and O/E-like sites (Hirota andMombaerts, 2004).

A comprehensive study was conducted on the mouse OR 37 subfamily, which comprises eight members. These genes are divided in two clusters, which is a unique phenomenon for OR subfamilies. The first cluster contains five genes with 90% sequence identity, and clusterII contains three genes with a lower sequence identity, which is 60%, compared to genes in cluster I. During the comparative analyses of these genes, it was shown that the promoter sequences are highly identical, and they were categorized to six common blocks based on their nucleotide content. Blocks I and IV are AT-rich, and block II and V contains short GA-rich sequences. The third block, block III and VI has a well conservedTCCCA motif. It was also identified that, these conserved regions act as potential binding sites for the NFY-type transcription factor, homeobox gene S8, and the zinc-finger transcription factor IK2 (Hoppe *et al.*, 2000).

The strongest candidates of regulatory motifs in OR gene promoters, however, are the *Olf-1*(O/E-like site) and homoeodomain binding sites, which were found to be conservedwithinthe promoter regions of all OR genes (Glusman *et al.,* 2000b; Sosinsky*et al.,* 2000; Vassalli *et al.,* 2002; Hoppe *et al.,* 2003). However olf-1 binding sites are also present in other the olfactory specific genes, such as *GnaI* (formerly known as *Golf*), adenylyl cyclase III (*AcIII*), olfactory cyclic nucleotide gated channel (*Cnga2*), and olfactory marker protein (*Omp*) (Kudrycki *et al.,* 1993; Wang *et al.,* 1993).These O/E-like binding sites can be further subdivided into four unique motif classesas M1, M2, M3, and M4 (Vassalli *et al.,* 2002; Rothman *et al.,* 2005; Michaloski *et al.,* 2006).

It was found, that one factor, O/E-2, can interact with structurally different binding sites, exclusively in the mouse OE (Hagman *et al.*, 1993; Wang *et al.,* 1997). However, the

presence of multiple copies of highly conserved ‘CCC’half side (Kudrycki *et al.,* 1993) of the ‘YTCCCYRGG GAR’ O/E-motif in the putative promoter regions of some clustered genes, shows the relevance of this motive for OR gene expression.

The identification of binding sites for distinct homeobox-type transcription factors (such as Ptx-1, Alx-3, and Lhx-2/LH-2) in OR37 promoters supports the idea, that distinct transcription factors might have important roles in pattern formation in the OE.

Starting from the identification of homeodomain binding site and O/E-like site in the OR gene promoters, a third transcription factor with a regulatory role in OSN development was found. It was reported, that this transcription factor, *Lhx2*, binds to the homeodomain site of the *M71*promoter or the H-element. The *Lhx2* deficient mice showed abnormalities in the OSN maturation process, in the states between terminally differentiated cell stage and immature neuronal stage, which indicates that *Lhx2* is important for late stage of OSN development (Hirota and Mombaerts, 2004; Kolterud *et al.*, 2004). Like Lhx2, another transcription factor, *Emx2*, was also identified on the M71 gene promoter (Hirota and Mombaerts, 2004). In mouse studies, it was shown that 75% of OR genes were not expressed in *Emx2* mutant mice, which were located on most chromosomes. However, Lhx2 mutation affected selectively Class II OR gene expression (Hirota *et al*., 2007). These results may indicate the presence of different regulatory mechanisms on Class I and Class II OR gene expression but also that ORs could be differentially regulated by patterned expression of transcription factors acting on related or identical sequence motifs within OR gene promoters.

Therefore, transcription factors, in combination with motifs located in far distance upstream of OR genes may represent essential components of cis-element-mediated regulation of OR gene expression.

So far only activators, such as Lhx, Emx, and Olf-1 were thought to affect gene regulation. However, a recent finding indicates, that mouse VRs and ORs also contain

conserved sequences, repressingVR and OR expression by interacting withBroad-Complex (BR-C), tramtrack (ttk), and brick a brack (bab) domain Zinc Finger (BTB-ZF) motifs (Godt *et al.,*1993; Zollman *et al.,* 1994; Michaloski *et al.,* 2012).

The mouse OR and V1R gene promoters were analyzed for common elements, and the existence of common motives were identified. They are present in 49–74% of the promoters, and are concentrated in the close proximity of TSSs of V1R genes. Two of these motifs share a consensus sequence (CNTCTGG) which is also present in the promoter regions of 40% of the 198 OR genes that were analyzed, and in a similar fashion in the TSSs of OR genes, as in the case of VR gene promoters. This sequence resembles closely to the sequence (CATCTGG) which is the binding site of the transcriptional repressor RP58, belongs to the BTB-ZF family (Aoki *et al.,* 1998). It was shown, that RP58 is related with condensed chromatin regions in the nucleus of IMR32 (human neuroblastoma) cells, indicating that its transcriptional repression activity might be important in regulating heterochromatin-mediated gene inactivation processes (Aoki *et al.,* 1998).

Among the Zbtb genes, it was found that Zbtb7b, and Zbtb7a are predominantly expressed in the mouse OE (Michaloski *et al.,* 2012). In zebrafish genome, however, there are 25 different Zbtb genes, and one of them, Zbtb10, is strongly expressed in the olfactory vesicle and in the OB (Gompel *et al.,* 2001).

2. PURPOSE

OR gene regulation is believed to be a combinatorial process that depends on the interactions of distal and proximal regulatorysequences which might constitute binding sites for specific TFs, as well as unknown inhibitory feedback signals that emanatefrom expressed receptors to maintain expression of a chosen OR gene locus. In this framework, it was aimed to identify sequences affecting OR gene expression that are located within sequences immediately upstream of a modelzebrafish OR gene, OR101-1. For this purpose, the promoter structure of the zebrafish OR101-1 gene was analyzed, because this gene is the only representative of the family of class II ORs in zebrafish. Thisunique situation provides distinct experimental and scientific advantages for the design of experiments targeting OR gene regulation, especially for promoter analyses. It was shown by deletion of the OR101-1 coding sequence from transgenic constructs that OR101-1 might be expressed by a single and specialized OSN cell type destined to the expression of class II ORs. This would constitute a selective experimental advantage because OR gene expression of a single gene could be analyzed without interference from the majority of other zebrafish ORs, which belong to the class I family. A variety of complementary approaches to study the regulation of OR101-1 gene expression were used, comprising BAC transgenesis and promoter bashing of short promoter transgenes. By scoring the efficiency of expression of the various transgenic constructs in OSNs in the nose, it was aimed to understand if and whichregulatory sites might be located in the OR101-1 promoter and how they affect expression of the gene.

3. MATERIALS AND METHODS

3.1. Materials

3.1.1. Fish

Adult and embryonic specimens of the AB/AB (ZL1) and AB/Tü (AB/TU TAB-14, ZL1438) inbred zebrafish strains were obtained from the Zebrafish International Resource Center (ZIRC), at the University of Oregon, Eugene, OR, USA and raised locally at the Bogazici University Life Sciences Center, Vivarium.

3.1.2. Equipment and Supplies

A list of equipment and supplies used can be found in Appendix A and B.

3.1.3. Solutions and Buffers

Solutions for standard molecular biology techniques were prepared according to Sambrook and Russell(1989) or were supplied with molecular biology kits and used according to the manufacturers' recommendations. Zebrafish specific solutions, such as embryo medium, were prepared according to Westerfield (2007).

3.2. Methods

3.2.1. Fish Growth and Breeding

AB/AB and AB/Tü strain zebrafish were kept in a dedicated fish room at Bogazici University Life Sciences Center Vivarium under constant light and temperature conditions, with a 14 hour light/10 hour dark cycle and at constant 28^{o}C. Fish were fed two times per day with flake food and once with live brine shrimp (Artemia sp.).

Matings to obtain fertilized oocytes for transgene injections were set up in 1.5 L breeding tanks in the evening of the day before harvesting. To obtain eggs used for injections male and female fish were separated by a divider in the same tank. The separators were removed in the morning and the eggs were collected immediately after fertilization. For regular breedings, male and female fish were put together in a tank without using separators.

3.2.2. Dissection of Olfactory Epithelial Tissue

For extraction of total RNA from olfactory epithelia, fish were euthanized in ice-water slurry for 10 minutes and the olfactory epithelia were dissected out in 1x ice cold PBS under 20x magnification ofa stereomicroscope. The dissection steps were as follows: first the head was separated from the body and then the lower jaw was separated from the head with surgical scissors. The tissue around the epithelia was removed by scraping around the olfactory rosette with forceps until the epithelia could be removed intact using fine forceps. The epithelia were stored overnight in TRIzol™ (Invitrogen) in 1.5 ml eppendorf microcenrifuge tubesat-80^{o}C.

3.2.3. Total RNA Isolation

A standard TRIzol™ (Invitrogen) protocol was performed to isolate total RNA of olfactory epithelia according to the manufacturer’ s instructions. The extracted and precipitated total RNA from 20 olfactory epitheliawasresuspendedina final volume of 20 µl of ultrapure water treated with Diethyl-pyrocarbonate (DEPC).

3.2.4. Polymerase Chain Reaction

PCR reactions were carried out usingGoTaq® Flexi DNA Polymerase (Promega) or Advantage® 2 Polymerase Mix (Clontech) according to manufacturer’s protocol. A maximum number of 30 cycles was used in a Standard PCR reaction and the template

DNA amounts were between 1ng and100 ng. A standard PCR reaction was set up as shown in Table 3.1 and a standard PCR program was run as shown in Table 3.2.

Table 3.1.Standard PCR components.

Reagent	1 Reaction
10X GoTaq Polymerase Buffer	3 µl
$MgCl_2$ (25µM)	1.8 µl
Forward Primer (10µM)	1.5 µl
Reverse Primer (10µM)	1.5 µl
dNTP (10µM)	0.5 µl
DNA polymerase	0.5 µl
Template	0.5 µl
dH_2O	20.7 µl
Total Volume	30 µl

Table 3.2.Standard PCR Program.

95°C	5 minutes
29 cycles of	
95°C	40 seconds
Annealing Temperature	40 seconds
72°C	1 min./kb
72°C	10 minutes
4°C	Forever

3.2.5. First-Strand cDNA Synthesis

First-Strand cDNA synthesis was performed using the SuperScript First-Strand cDNA Synthesis System for RT-PCR (Invitrogen) according to the manufacturer's manual. Shortly, 1 µg of total RNA extracted from olfactory epithelia was mixed with 1 µl of 10 mM dNTP mix 1µl of 0.5 µg/ml oligo (DT)primer and 6 µl of diethyl-polycarbonate (DEPC)-treated H2O. The mix was incubated at 65 ^{o}C for 5 minutes and placed on ice. 2 µl of 10X RT buffer 4 µl of 25mM MgCl2, 2 µl of 0.1M DTT and 1µl RNaseOUT™ was added and the reaction was kept on 42 ^{o}C for 2 minutes. Then, 1 µl of SuperScript™ II RT was added and the mix was incubated at 42 ^{o}C for 50 minutes. The reaction was stopped by incubating at 70^{o}C for 15 minutes. 1µl of RNase H was added and the reaction was incubated 20 minutes at 37^{o}C. The synthesized cDNA was then stored at -20^{o}C.

3.2.6. RNA Ligase Mediated Rapid Amplification of cDNA Ends

For 5'-RACE experiments, the FirstChoice® RLM-RACE Kit (Ambion) was used. The reaction steps were performed according to the manufacturer's protocol. Briefly, 1 µg of extracted total RNA of olfactory epithelia was treated with 2µl Calf Intestine Alkaline Phosphatase (CIP) for 1 hour to remove PO_4 from 5'-degraded mRNAs, rRNAs and tRNAs. CIP was removed by phenol: chloroform extraction followed by chloroform extraction and the sample was treated with 1µl Tobacco Acid Pyrophosphatase (TAP) to remove G-caps from full-length mRNA molecules in the sample. T4 RNA Ligase was then used to ligate a 5'-RACE adapter to full length mRNAs. The readily processed sample was stored at -80^{o}C.

2 µl of the sample was used to synthesize first-strand RLM-RACE cDNA by SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen) according to the manufacturer's manual. 1 µl of the first-strand cDNA was used as template for PCR with an outer primer specific for the 5'-RACE adapter and a gene specific primer for the OR101-1 gene. 35 cycles of PCR were performed and 1 µl of the product was used as template for a nestedPCR with an inner 5'-RACE adapter primer and a gene specific primer for the OR101-1 gene for 35 cycles. Advantage® DNA polymerase (Clontech) was

used in these PCR reactions with the supplied 10X buffer and 1 µl of 10mM dNTP mix. The products were cloned to pGEM®-T Easy vector (Promega), selected for size and sequenced.

3.2.7. Restriction EndonucleaseDigestion of DNA

For restriction endonuclease digestion of DNA, restriction enzymes from New England Biolabs, Fermentas or Promega were used. Digestion reactions were set up with 3 - 5 Units / µg of DNA in a final concentration of 1xof the supplied buffers, and 1x of BSA, if indicated by the manufacturer. The reactions were incubated at 37°C or appropriate temperature for up to 8 hours.

3.2.8. End polishing of 5'-overhangs and 3'-overhangs (blunting)

End polishing reactionswereperformed according to the manufacturer's recommendations. 1x NEB Buffer 2, 33µM dNTPs, 1unit of Klenow Large Fragment (NEB) /1ug of DNA and the DNA fragment to be blunted were combined in a reaction tube and incubated at 25°C for 15 minutes for 3-overhangs, or at 37°C for 30 minutes for 3'-recessive ends. The reaction was stopped by adding10mM EDTA and the enzyme was heat inactivated at 75°C for 20 minutes. The DNA was purified by regular PCR purification or gel purification methods.

3.2.9. Gel Electrophoresis and Gel Extraction of DNA

DNA samples were run on 1% agarose gels stained with EtBr. As a marker 1kb DNA Ladder (Roche) and/or 100bp (Roche) have been used in all gels.

Table 3.3.Oligonucleotide primers used in this thesis and their sequences (5'-3').

101 Pst Forward	ATCTGCAGCGCTCTTGCACTTGCCCTAATTGG
101 Sph Reverse	TAGCATGCGTACCACAGTATGTTGGGCCATTGG
OR101-1 Forward	ATGAACACCAGCGGCTCGGT
OR101-1 Reverse	GTGCCATGCTCATCCTTCTCA
OR101-1 Rec5-Forward	CGCCTGCACAACATCATCTAGTCG
OR101-1 Rec3-Reverse	CGCAAAACATTTCCTTCAGCTTTAAGAC
OR101-1Nco Forward	TGCCATGGACACCAGCGGCTCGGTG
OR101-1 Upstream Nco Reverse	ACCCATGGTGGTGATACACAGCCG
OR101-1 HindIII Forward	GAAAGCTATGAGAAGGATGAGCATGGC
OR101-1 NsiI Reverse	CAGCTGCATCTGTGTAATAACTGACGC
101-1 EcoRI Reverse	TTTCCAGATTGATGGAATTCATTCTAAGCGACAGAAGG
101-1 EcoRI Forward	CGCTTAGAATGAATTCCATCAATCTGGAAATTGCAATG

Table 3.3. Oligonucleotide primers used in this thesis and their sequences (5'-3')(Cont.).

Indigo Rec Forward	GGCGTTTCCGTTCTTCTTCGTC
Indigo Rec Reverse	GGCGCCTGATGCGGTATTTTCTCC
pIndigo tol2 Forward	TTCTCTGTTTTTGTCCTGGGAATGAACAATGGAAGTCCG AGCTCATCGCTCCCTGCTCGAGCCGGGCCCAAGTG
Indigo tol2 Reverse	CCCGCCAACACCCGCTGACGCGAACCCCTTGCGGCCGC ATATTATGATCCTCTAGATCAGATCT
OR101-1 3.5ups Sac1-Forward	ATGAGCTCCAGGGACTGGTTGCATGC
OR101-1 3.5 ups Spe1- Reverse	TAACTAGTCCTGTGGGTCAAGTCCTG
101-1 5RACE 1	CCGCGATGGCGTTTGACCGCTACGTAGC
101-1 5RACE 2	CGTGCTTCCCATCTTCTTCACCGTCTACG
101-1 GSP 5’RACE	CGCGGCCTGTCAGTCCCATCCAGAGAAC
OR101-1 Inhibitor Forward	GAACTAGTCTAATGTGCTGC
OR101-1 Inhibitor Reverse	CTGGAAGCATTGCTACCTAGG
GalK Forward	TGCGTTGGCAAACAGAGATTGTGTT

Table 3.3. Oligonucleotide primers used in this thesis and their sequences (5'-3') (Cont.).

GalK Reverse	TGAAACGTATGGGCGAGTTGATGG
Bactin-Forward	CTGGGATGACATGGAGAAGATCTG
Bactin-Reverse	CCTTGATGTCACGGACAATTTCTC
T7-HighTM	TAATACGACTCACTATAGGGCGAATTGG
M13-R-HighTM	GGAAACAGCTATGACCATGATTA
M13-Forward	CGCCAGGGTTTTCCCAGTCACGAC
GFP-Forward	GCGACGTAAACGGCCACAAGTT
GFP-Reverse	TACTTGTACAGCTCGTCCATGCCG
GFP 819 Forward	GAGAAGCGCGATCACATGGTCCTG
GFP 5’ Reverse	AACTTGTGGCCGTTTACGTCGC

3.2.10. Gel Extraction of DNA fragments from Agarose Gels

The QIAquick Gel Extraction Kit (QIAgen) was used to extract DNA fragments from agarose gels after electrophoretic separation. The samples were run on agarose gel at 80V for short runs or at 30V for overnight separations and the band of interest were cut with a scalpel under UV illumination. The gel slice was melted in 3 volumes of the supplied solubilization Buffer (QG) and centrifuged through a spin column. The column was washed once with wash buffer and DNA was eluted by elution buffer. The purity of the DNA band was analyzed by agarose gel electrophoresis and the concentration of the obtained DNA was measured using a NanoDrop Spectrophotometer.

3.2.11. Purification of DNA

For purification of PCR products and restriction enzyme reactions the High Pure PCR Purification Kit (Roche) was used according to the manufacturer's protocol.

3.2.12. Ligation of DNA to Vectors

DNA fragment to be cloned (insert) and the vector backbone were prepared by digestion with suitable restriction endonucleasesfollowed byDNA purifications as described above. 3 µl of the purified insert and vector was run on agarose gels side by side to estimate relative intensity and molar ratios of vector to insert. Using the sizes of the fragments and the estimated relative intensity of the samples, a 1:3 and/or 1:5 molar vector to insert ratio were calculated. Vector and insert DNAs were combined, not exceeding a total of 10 ng DNA, and H_2O was added to bring the reaction volume to 20 µl. 1 µl of T4 DNA Ligase (NEB) was added after addition of 1 µl of supplied 10X T4 Ligation Buffer. The reactions were kept at 16^oC in a water bath for overnight and transformed to competent cells.

For ligation of PCR products the pGEM®-T Easy Vector System (Promega) was used. 3.5 µl of PCR products was mixed with 0.5 µl of pGEM®-T Easy vector and 5 µl of supplied 2X Rapid Ligation Buffer. 1 µl of T4 DNA Ligase (Promega) was added and the reaction was incubated at 16 degrees overnight and transformed to competent cells.

3.2.13. Preparation of Competent Cells (Calcium Chloride method)

A single colony of the Top10 MRF'*E.Coli* strain was picked from an agarose plate and used to inoculated in a 100ml LB liquid culture. The culture was grown until the OD_{550}reached0.50-0.60. The bacterial culture was then chilled on ice for 15 minutes and centrifuged gently (4K) for 10 minutes at 4^oC. The supernatant was poured off and the pellet was drained by inverting the tube onto tissue paper. The remaining supernatant that was retained in the tube was used to tab the bacteria back into suspensionbeforetheywere

resuspended in 10ml 0.1M $CaCl_2$ / 100 ml culture. The mixture was stored on ice for 15 minutes and then centrifuged gently (4K) for 10 minutes at 4^oC. Supernatant was poured off and the tube was inverted onto a tissue paper to drain the pellet. The bacteria were tapped again with the remaining supernatant and resuspended with 4 ml 0.1 M $CaCl_2$, 15% glycerol / 100 ml culture. The mixture was kept on ice again for 15 minutes. 50µl aliquots were prepared and were immediately frozenin liquid nitrogen, and stored at -80^oC until used. To control the efficiency of competent cells, one vial was transformed with a circular plasmid of known concentration *(See transformation method)* and 100µl of the transformed bacteria was spread on an LB_{Amp} plates. The plate was kept at 37^oC overnight and on the following day the colony numbers were counted and calculated according to thefomula below:

Transformation Efficiency (CFU/µg) = (colonies per plate, CFU) * (total volume, µl) * (dilution applied if any) * (10^6 pg/g) / (volume plated, µl) * (pg DNA used)

3.2.14. Transformation of Plasmids to Competent Cells

50µl of competent cells were thawed on ice for 10 minutes and up to 10 µl of ligation reactions or plasmid DNA were added. The mixture was incubated 20 minutes on ice, incubated 1.5 minutes at 42^oC and 2 more minutes on ice. The reaction was then incubated at 37^oC in 400µL of LB for recovery and it then spread to ampicillin-containing LB agar plates and incubated overnight at 37^oC. Colony PCR reactionswereperformed by using suitable primers and positive colonies were selected based on the PCR results. Colonies were picked the next day and grown in 6 ml LB with 6µl ampicillin overnight. The next day, plasmids were isolated using the QIAprep Spin Miniprep Kit (Qiagen) and further analyzed by analytical digests.

3.2.15. Analytical Digests

In order to confirm successful cloning of DNA fragments, isolated plasmids were cut with suitable restriction enzymes (NEB, Fermentas or Promega) that would cut DNA into at least 2 distinct fragments and the reactions were run on 1% agarose gel stained with EtBr.

3.2.16. *GalK* Recombination

This recombination technique is used to achieve recombinant BAC plasmids carrying targeted mutationsusing the *galK* positive and counter-selection procedure. The BAC modificationsare performed byusing a modified bacterial strain, SW102, which contains the λ-prophage recombineering system but in whichthe galactose operon has been modified by a deletion of the galactokinase gene (*galK*). Importantly, the *galK* function can be supplied back*in trans* form a targeted integration of the *galK* gene, and the ability to grow on galactose as carbon source is restored. Thus*galK* selection can be used efficiently to follow the success of the recombineering process.

In a first round of recombination,a *galK* cassette, flanked with specific homology arms, is inserted into the BAC by homologous recombination. Successful recombinants will be able to grow on minimal media with galactose as the only carbon source, using the *galK* gene as a positive selection marker.

In a second round of recombination, the *galK* cassette is replaced by the modified target sequence using identical homology arms for recombination. This time *galK* can be used as a negative selection marker by selecting against thepresence of the *galK* cassette. 2-deoxy-galactose (DOG) is a non-toxic substance for bacteria, unless phosphorylated by functional *galK*. Phosphorylation by *galK* turns DOG into 2-deoxy-galactose-1-phosphate, a non-metabolizable and therefore toxic intermediate.

Based on the analysis of publicly available database, the D-KEY 206L8 BAC clone was selected for OR101-1 transgenic studies. The D-KEY206L8 BAC contains the OR101-1 gene and is symmetrically flanked by approximately 60kb of surrounding genmic

sequence. It was aimed to observe and to compare the expression of a fluorescently-tagged OR101-1 transgene and a OR101-1 deletion constructs *in vivo.*The differences in transgenic expression patterns and the number of transgene expressing sensory neurons when the transgene did or did not include a sequence coding for the OR101-1 gene were aimed to be analyzed. The analysis was performed by tagging the coding sequence with a enhanced yellow fluorescent protein (EYFP) via *galK* recombination technique to trace the transgenic expression. A related construct, in which the OR101-1 coding sequnce was replaced with a sequence coding for the yellow fluorescent protein Venus,which was made available by X. Bayramli.

3.2.17. Transformation of BAC into Electrocompetent Cells (SW102)

The SW102 E.coli strain and necessary plasmids for *galK*-mediated recombinationwere kindly provided byNCI-Frederik National Laboratory, US.BAC cloneDKEY-206L8, harboringthe OR101-1target gene and its flanking sequences was transformed into SW102 electrocompetent cells according to the protocol (Sorensen *et al*., 2005). 5ml o/n cultures were inoculated from SW102 frozen stocks. The o/n culture was diluted 1:50 in an autoclaved 50ml Erlenmayer flask with 25 ml LB. Cultures were incubated for 3-5 hours with chloramphenicol selection in a 32^{o}C shaking water bath until the density reached an OD_{600} of 0.55-0.6. Then bacteria were put into an ice/ water slurry to be cooled down for a few minutes and transferred into pre-cooled 15ml falcon tubes. The bacteria were centrifuged at 4^{o}C in a table top centrifuge for 15 minutes at 4500 RPM and the supernatant was discarded. The tube was briefly inverted on a paper towel and 1ml of pre-cooled dH_2O was added into the tube while keeping the tube in the ice/water slurry.

The pellet was re-suspended by gently swirling the tube in ice-water and when the pellet was re-suspended, the tube was filled up to 10 ml with ice cold dH_2O and the tube was inverted a couple of times for mixing. The bacteria were centrifugedagain for 15 minutes at 4500 RPM, the supernatant was discarded and the pellet was re-suspended in 1 ml ice cold dH_2O following a 9ml of ice cold dH_2O.Afterwards, the solution was centrifuged one more time for 15 minutes at 4500 RPM.

All supernatant was gently removed by inverting the tube on a paper towel. The competent cells were stored on ice, and 25µl of the freshly made electrocompetent cells, which were re-suspended with the remaining solution, were transferred to a pre-cooled 0.1cm cuvette. 1-5µg of BAC midi prep DNA was added into the cuvettes and transformed by electroporating with 1.80 kV for 6.1ms. The transformed bacteria were transferred to 1.5ml eppendorf tubes with 1ml LB medium, incubated at 32^{o}C for 1 h and plated on both ampicillin selective and chloramphenicol selective plates, and verification was done using standard PCR reactions, analytical digests, and direct DNA sequencing (Figure 3.1., and Figure 3.2.).

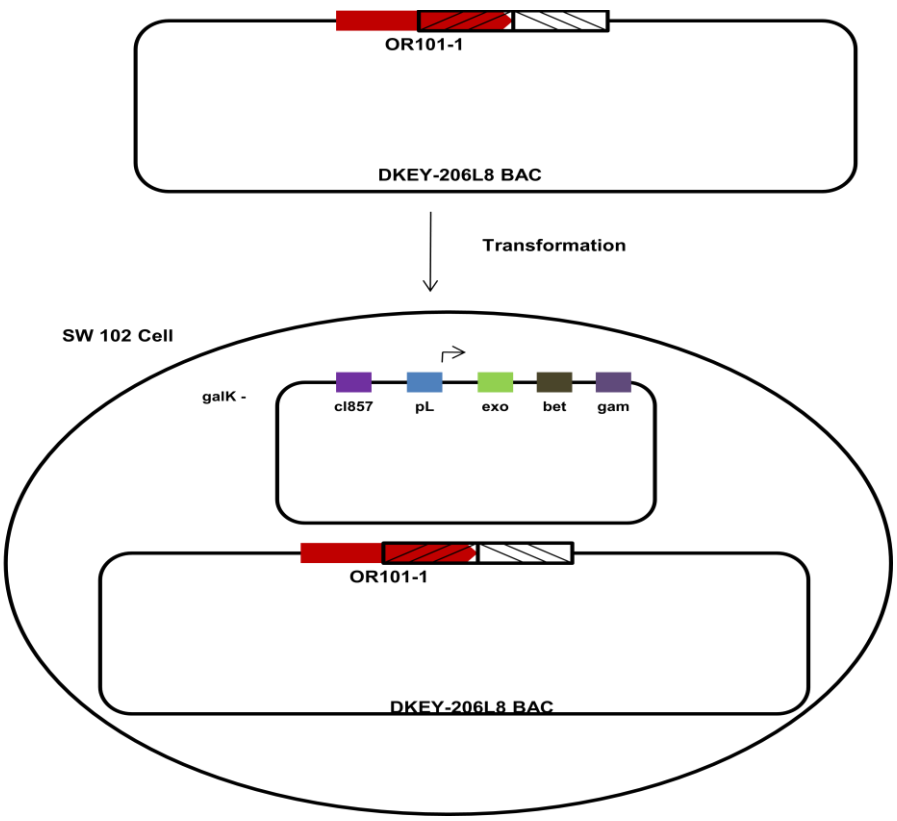


Figure 3.1. Transformation of DKEY-206L8 BAC Clone into SW102 Modified Bacterial Strain.This SW102 strain contains a λ-prophage recombineering system (colored boxes) , in which targeted BAC sequence will be transformed.

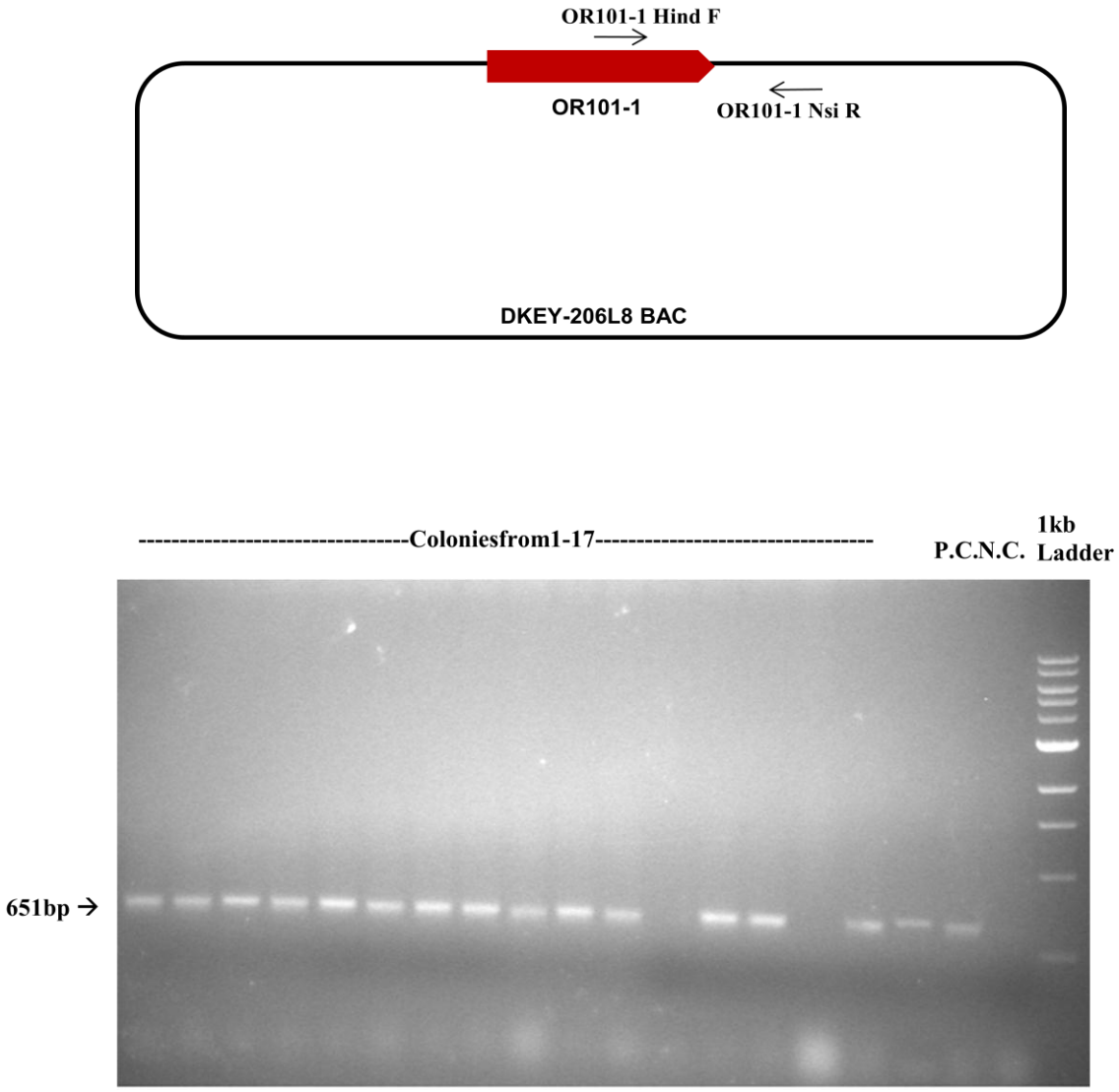


Figure 3.2. Colony PCR for Verification of BAC Transformation into SW102 Electrocompetant Cells. BAC transformation control with primers OR101-1 Hind F and OR101-1 Nsi Reverse. The expected band size was 651bp. First 17 lanes were tested colonies, Lane 18 includes a positive control, and lane 19 was negative control.15 out of 17 colonieswere succesfully transformed.

By using PCR-based screening for transformation efficiency,15 out of 17 SW102 colonies were succesfully transformed with the DKEY- 206L8 BAC clone. The transformation efficiency was ~90%. Among positive colonies, based on their brightness on the gel, colonies 5 and 13 were selected and inoculated for further processing.

3.2.18. Preparation of Plasmids Having Flanking Homology Arms

The BAC engineering is based on recombination using homology sequences flanking the target integration site. For the construction of homology arms 2 different vectors,pGEM®-T Easy, and a modified version of the pACSF vector harboring a sequence coding for a short T2A peptide sequence and an in frame fusion of the reporter enhanced yellow protein (EYFP), were used. T2A is a "self-cleaving" 2A peptide, which can be used to express two independent proteins from a single promoter (de Felipe and Ryan, 2004; Osborn *et al.,* 2005; Heras*et al.,* 2006). The 2A-like sequences exist in several viruses and can be used to mediate protein cleavage from a single open reading frame during translation. Through a ribosomal skipping mechanism, the 2A peptide prevents normal peptide bond formation, resulting in the translation of 2 protein fragments from the same mRNA (Donnelly *et al.,* 2001).

The left homology arm (LH) was first amplified with the primers 101_PstI_F and 101_EcoRI_R and the right homology arm (RH) was amplified with the primers 101_EcoRI_F and 101_SphI_R from the identified BAC, and cloned into the pGEM®-T Easy vector (*See in the Appendix C).* The left homology arm was digested out from pGEM®-T Easy with XhoI and SpeI restriction enzymes and cloned into pACSF vector, possessing T2AEYFP*(See in the Appendix C),* digested with XbaI and SpeI restriction enzymes. The right homology arm was digested with EcoRI and SphI restriction enzymes and cloned into the pACSF plasmid by using the same restriction sites. After the final construct was confirmed with PCRs, analytical digests, and sequencing, the T2AEYFP fragment was replaced by *galK* fragment by a sequential digestion of (LH) T2AEYFP (RH) fragment in pACSF plasmid with XhoI restriction enzyme, following by blunting XhoI overhang sequence and then a NotI digestion. ThepgalK vector was digested with EcoRV and NotI restriction enzymes for obtaining *galK* cassette to be cloned between LH and RH into the abovementioned vector(Figure 3.3.).

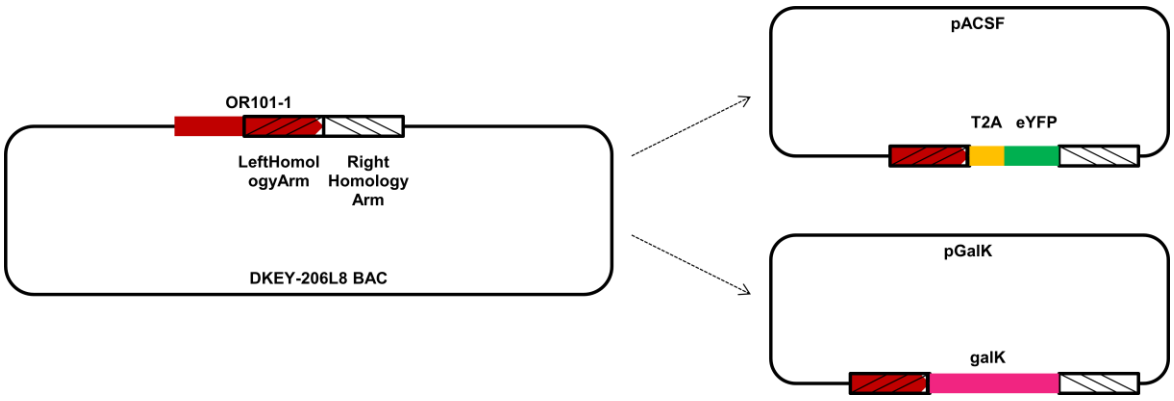


Figure 3.3.Construction of Recombination Plasmids.The 374bp Left homology arm was cloned into the pGEM®-T vector. Using standard restriction digests and cloning techniques, it was cloned into both pACSF vector and pgalK vector. The 543bp right homology arm was cloned using similar cloning methodology into both pACSF and pgalKvectors.

3.2.19. OR101-1 T2A EYFP Construction in BAC

5ml of modified SW102 cells were inoculated o/n at 32^{o}C in LB containing chloramphenicol (12.5µg/ml). The following day 500µl of culture was taken and used to inoculate in a 25ml LB liquid culture containingchloramphenicol until the OD reached0.55-0.6, in a 32^{o}C water bath. 10ml of the culture was used for heat induction in a 42^{o}C water bath for 15 minutes while continuously shaking. The remaining culture was left at 32^{o}C as an un-induced control. After heat induction, induced and un-induced bacteria were cooled down briefly and transferred to 15 ml falcon tubes. They were centrifuged at 4^{o}C for 20 minutes at4500 RPM. The supernatant was discarded and the pellet was resuspended with 1ml ice-cold dH_2O by swirling in ice water. Then dH_2O was added up to 10 ml and the centrifugation was repeated. The supernatant was removed by inverting the tubes briefly on a paper towel and approximately 50µl of the competent cells were kept on ice for transformation. 25µl of competent cells and 200 ng of plasmid DNA were combinedinpre-cooled 0.1 cm cuvettes for electroporation at 1.80 kV for 6.1ms. The electroporated cells were suspended into 1ml LB for 4h for recovery in a32^{o}C water bath, and after the recovery period they were washed twice with 1X M9 salt solution.

1ml culture was pelleted in an eppendorftube at 13200 RPM for 15 seconds and the supernatant was removed briefly with a pipette. The pellet was resuspended in 1ml 1XM9 salt solution, and pelleted again. This washing step was repeated once more and 100µl of pellet was plated onto M63 minimal media plates with galactose, leucine, biotin and chloramphenicol with serial dilutions (1:10, 1:100). The plates were incubated for 3 days at32^{o}C in an incubator. After 3 days single colonies were selected streaked onto MacConkey/galactose/chloramphenicol plates for verification of *galK*+ colonies. These plates were incubated overnight at32^{o}C in an incubator(Figure 3.4.).

Single, bright red (*galK positive*) colonies were inoculated overnight in 5ml LB containingchloramphenicol at 32^{o}C in a water bath. Electrocompetent cells were prepared for substitution with the target fragment, as mentioned above. 25ul of electrocompetent cells were electroporated with200ng of the targeting construct, allowed to recover in 10ml LB for 4 hours in ashaking water bath at 32^{o}C, and washed twice with 1XM9 salts as detailed above. This time the pellet was serially plated onto M63 minimal media plates containing glycerol, leucine, biotin, 2-deoxygalactose (DOG), and chloramphenicol and incubated in a 32^{o}C incubator for 3 days. Verification of positive colonies was performedbycolony PCR reactions and analytical digests (Figure 3.5.).

Based on the mentioned protocol above, the recombination reaction was carried out. First the (LH) *galK* (RH) in pgalK vector was transformed into the DKEY-206L8-positive SW102 clone, which was followed by streaking the colonies onto McConkey Agar plates for further verification. Then the recombination reaction performed between (LH) T2AEYFP (RH) in pACSFvector and the *galK*-modified version of the DKEY206L8-positive SW102 clone. As a result the final transgenic BAC clone containing the OR101-1 gene that is tagged with T2A-EYFPwas obtained (Figure 3.6.).

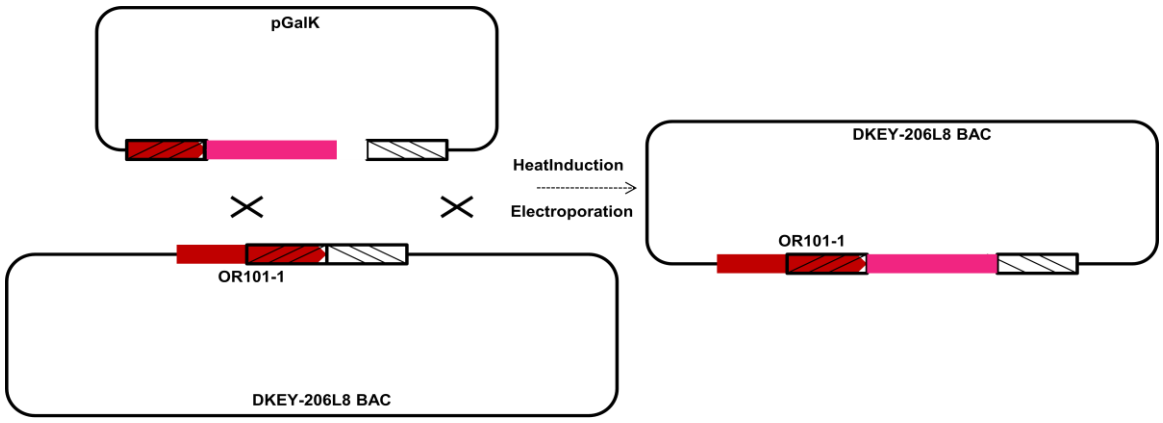


Figure 3.4. First Step of *galK* Recombination Technique.The linearized pgalK vector and *galK* cassette was electroporated into SW102. The selection for positive recombinants was done by spreading the bacteria onto M63 minimal media plates and then streaking colonies onto McConkey Agar plates containing galactose as carbon source.

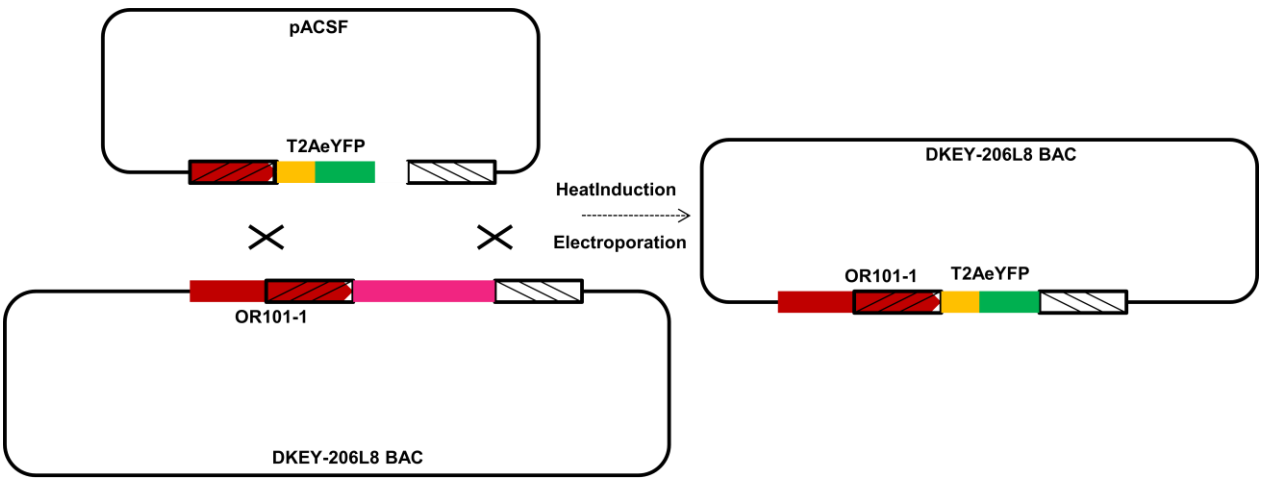


Figure 3.5.Second Step of *galK* Recombination Technique. Linearized pACSF-*galK* vector was used for recombination. The selection was done via spreading the bacteria M63 minimal media plates containing 2-DOGto eliminate *galK* positive colonies and to select for colonies in which the *galK* cassette was replaced by the OR101-1-T2A-EYFP target sequence.

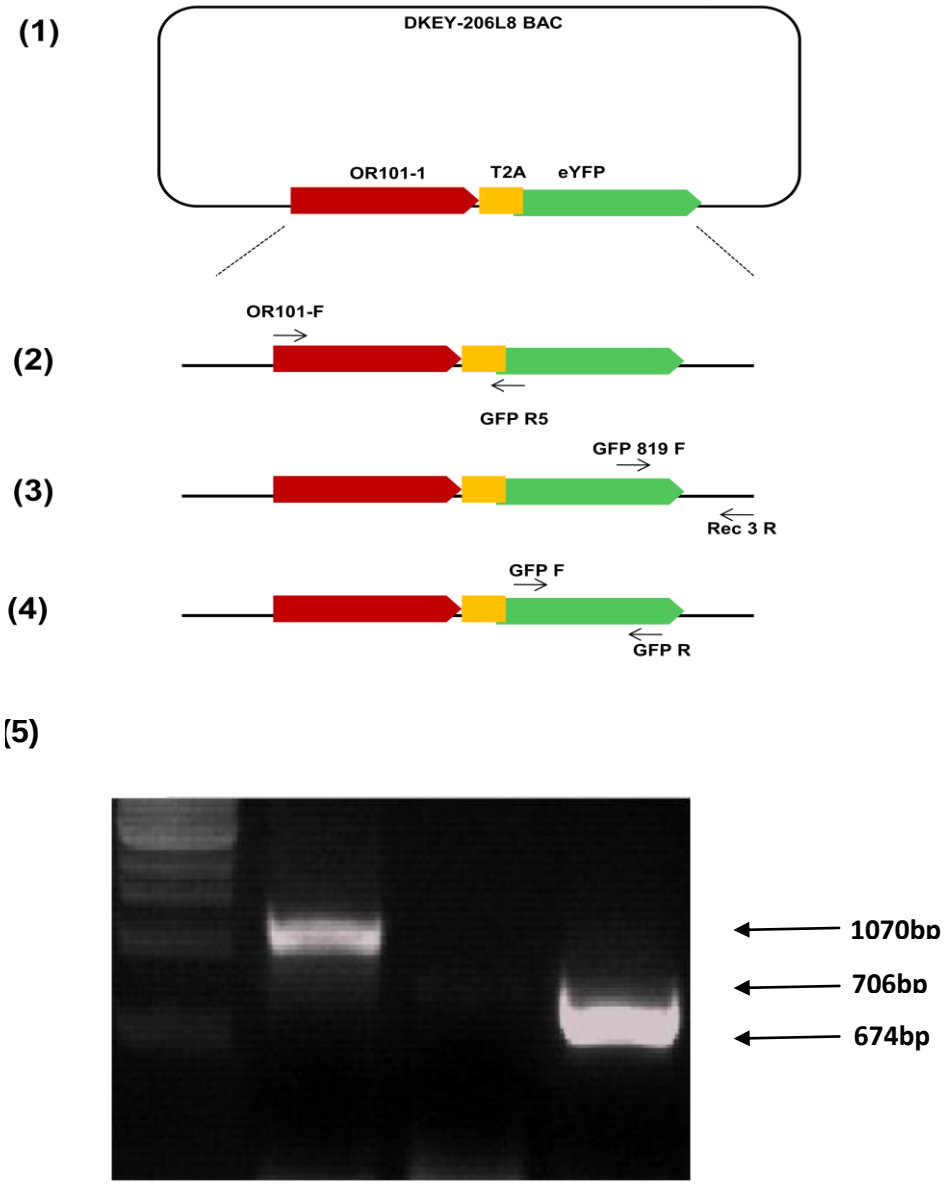


Figure 3.6. OR101-1 Tagged with EYFP in DKEY-206L8 BAC.**(1)**The scheme of DKEY 206L8 harboring the OR101-1 gene (red) tagged with T2A (orange) EYFP (green). PCR verification was done using a standard PCR reaction with the primers: **(2)** 101 F/GFP-R5, **(3)** GFP-819 F/Rec 3 R and **(4)** GFP F/GFP R. **(5)** expected band sizes were showed above.

GalK recombination is a relatively time consuming protocol but the average recombination efficiency of both steps was 50 %. In the first step, when the*galK* cassette was recombined into the BAC DKEY 206L8, 10 colonies were arbitrarily selected and streaked onto McConkey-Agar plates, and 7 out of 10 colonies were bright red, indicating that they were *galK* positives and could utilize the supplied galactose as carbon source. In

the second step, when the target sequence T2AEYFP, was introduced into the *galK*-positive BAC 206L8, colony PCR on 20 colonies was performed, and six of them were positive for the modification. The relatively low number of positive colonies in the second step might be because of some experimental mistakes, as insufficient washing or electroporation problems.

3.2.20. PiTol2-Amp Transformation

Sequences derived from the iTol2 transposon (tol sites) were introduced to the modified BAC clone via another round of recombination. The recombination construct to be used in the iTol2 recombination was prepared by PCR amplification with 75bp long Forward and Reverse primers, 50bp of which are complementary to the BAC backbone, pIndigo plasmid, and the remaining 25bp are complementary to iTol2 transposon end sequences. After DpnI treatment and Gel purification of the amplified PCR product, recombination was conducted and the toll sites were inserted into the BAC sequence with a technique similar to the one used during the second step of *galK* recombination. The final construct was verified using standard PCR reactions, analytical digests, and direct DNA sequencing(Figure 3.7., Figure 3.8., and Figure 3.9.).

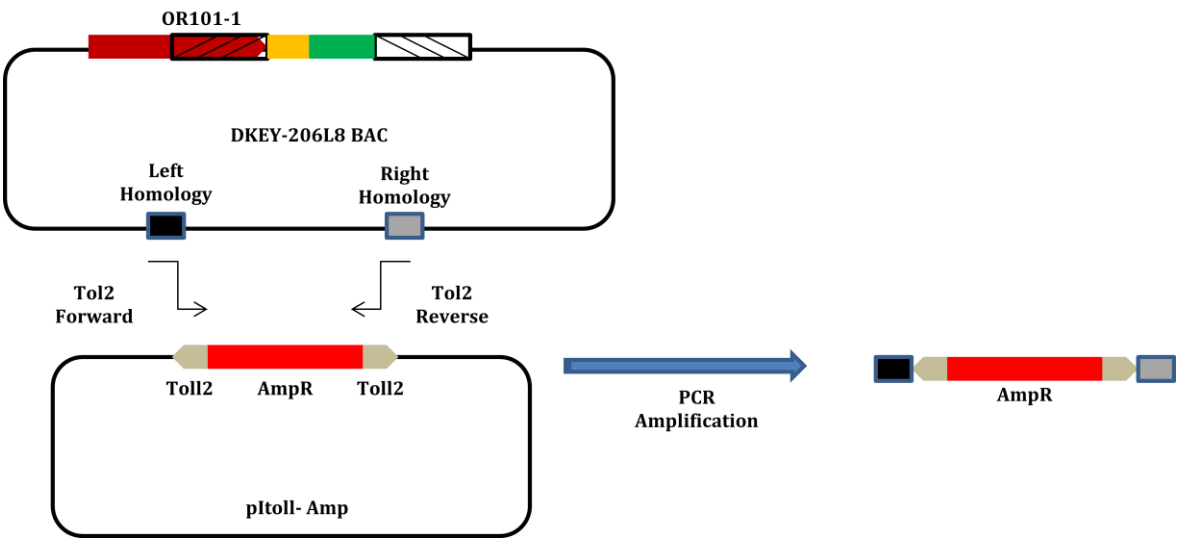


Figure 3.7.Preparation of the PCR Product to be used in the Recombination.Special primers named as Indigo_tol2_Forward and Indigo_tol2_Reverse were designed to amplify the AmpR gene with tol2 sites. PCR product was used for iTol2 recombination procedure.

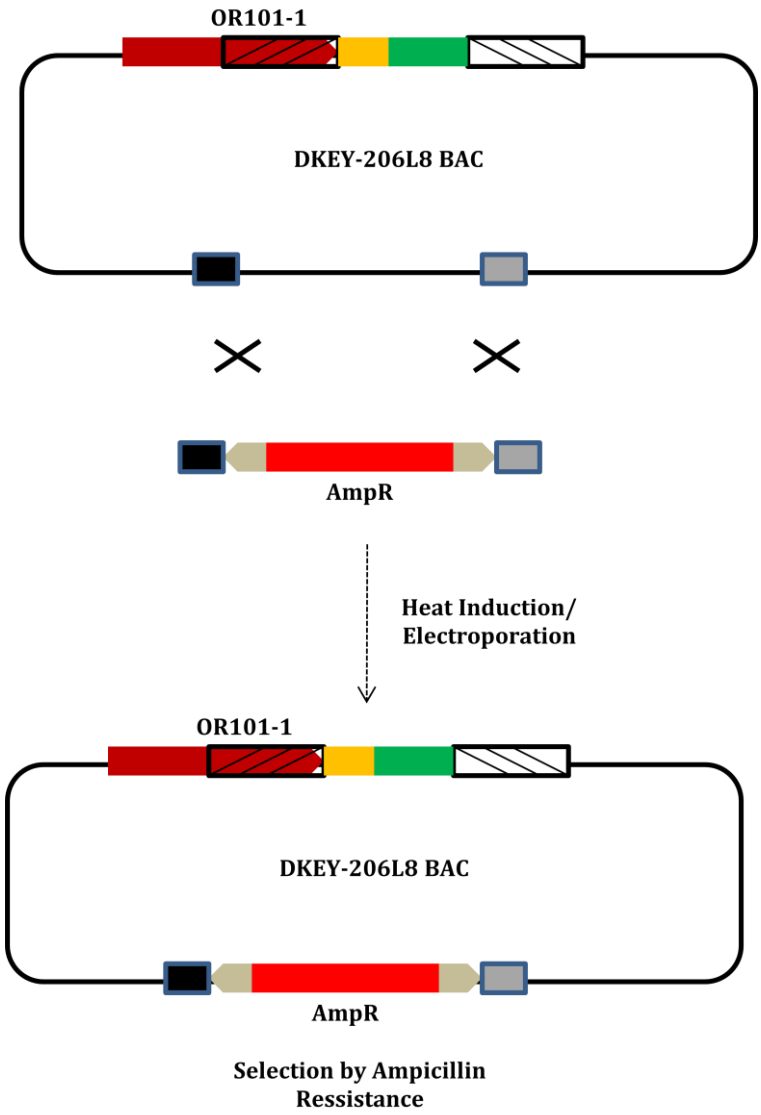


Figure 3.8.ITol2 Recombination.The amplified PCR product was treated with DpnI to exclude the template DNA from the reaction mixture. The iTol2 positive colonies were selected using ampicillin resistency.

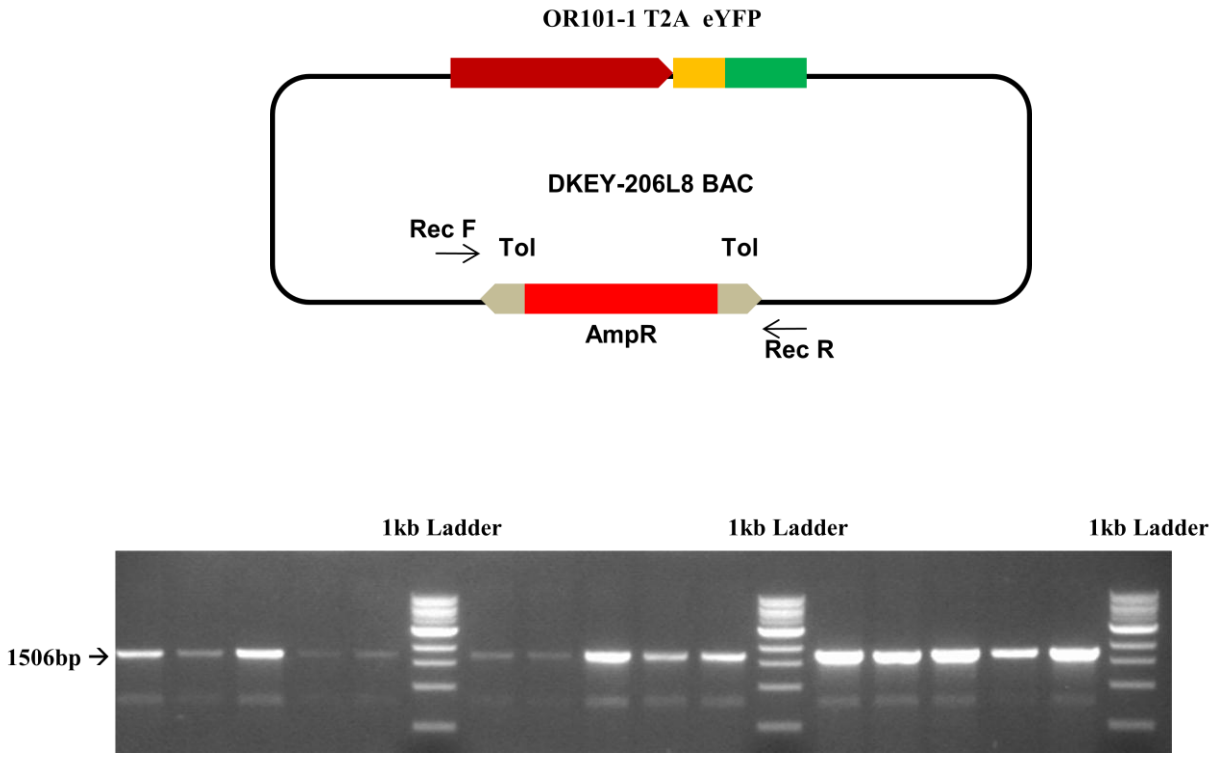


Figure 3.9.Confirmation of iTol2 Recombination with a Colony PCR Reaction. (Above)Verification of pIndigo BAC iTol2 constructs with the primers Indigo_Rec_F and Indigo_Rec_R. (Below) The iTol2 recombination efficiency was 100%.The expected band size was 1506bp.

Table 3.4: PCR Conditions for Recombination.

95^{o}C	5 minutes
6 cycles of	
95^{o}C	40 seconds
55^{o}C	40 seconds
72^{o}C	1 min./kb
19 cycles of	
95^{o}C	40 seconds
62^{o}C	40 seconds
72^{o}C	1 min./kb
72^{o}C	10 minutes
4^{o}C	Forever

The PCR product was treated with DpnI restriction endonuclease overnight to digest any remaining DNA template, and purified using regular gel purification protocols as explained above. Recombinaton reactions were performed using the protocol described above. The target PCR fragment with flanking homology arms was electroporated at 1.80 kV for 6.1ms. Transformed cells were grown at 32^{o}C for 4 hours and spread onto agar plates containing chloramphenicol and ampicillin, respectively. The plates were incubated at 32^{o}C.

3.2.21. Gap Repair Protocol

391bp of sequence located 3.5kb upstream of OR101-1 and 948bp of sequence coding for the OR101-1 gene were used as homology arms for gap repair cloning using recombination in *E.coli*. Specific primers containingSpeI and SacI restriction sites were designed for amplification of the upstream homology arm fromthe DKEY 206L8 BAC clone using regular PCR method and cloning into the target vector. Additionally, the OR101-1 coding sequence wasused as the downstream homology region, and cloned into the same target vector containing sequences coding for the self-cleavable peptide T2A and the yellow fluorescent reporter protein EYFP. Electrocompetent cell preparation and electroporation of the gap repair construct were conducted using an SW102 clone transformed with the DKEY206L8 BAC. Transformed cells were plated onLB plates containing ampicilin and incubated at 32^{o}C. For the success of the gap repair cloning was performed using regular colony PCRs, analytical digests and direct DNA sequencing through recombination junctions(Figure 3.10., and Figure 3.11.).

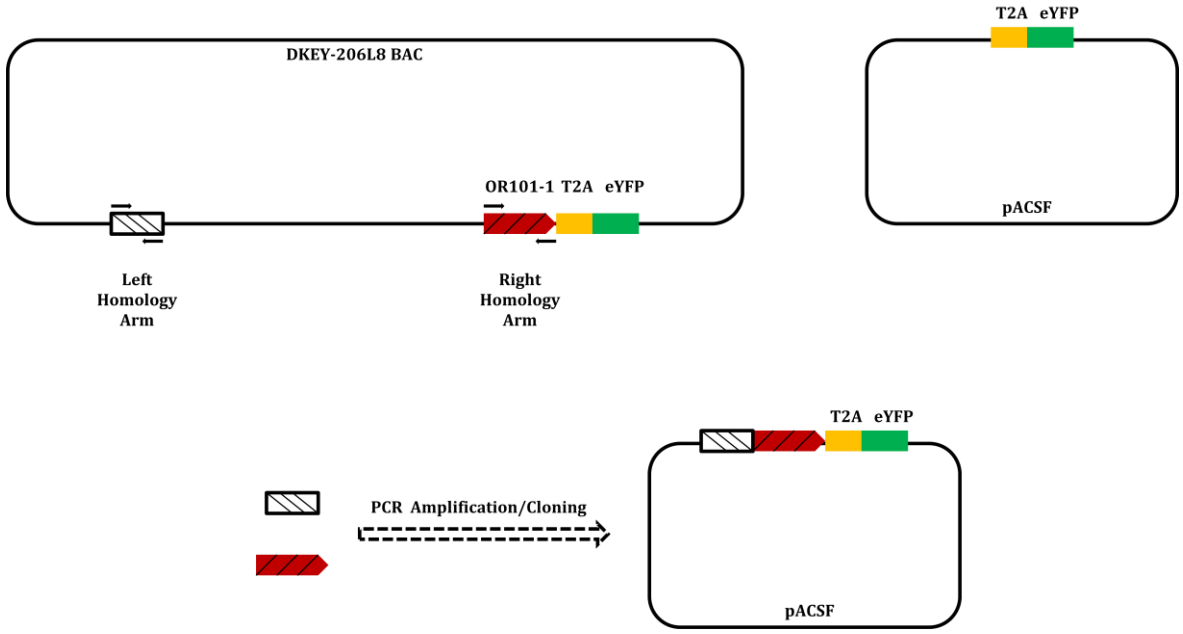


Figure 3.10.Cloning homology arms to pACSF Vector for Gap Repair Protocol. Left homology arm from 3.5kb upstream sequence of OR101-1 gene and right homology arm of OR101-1 gene was PCR amplified and cloned into pACSF.

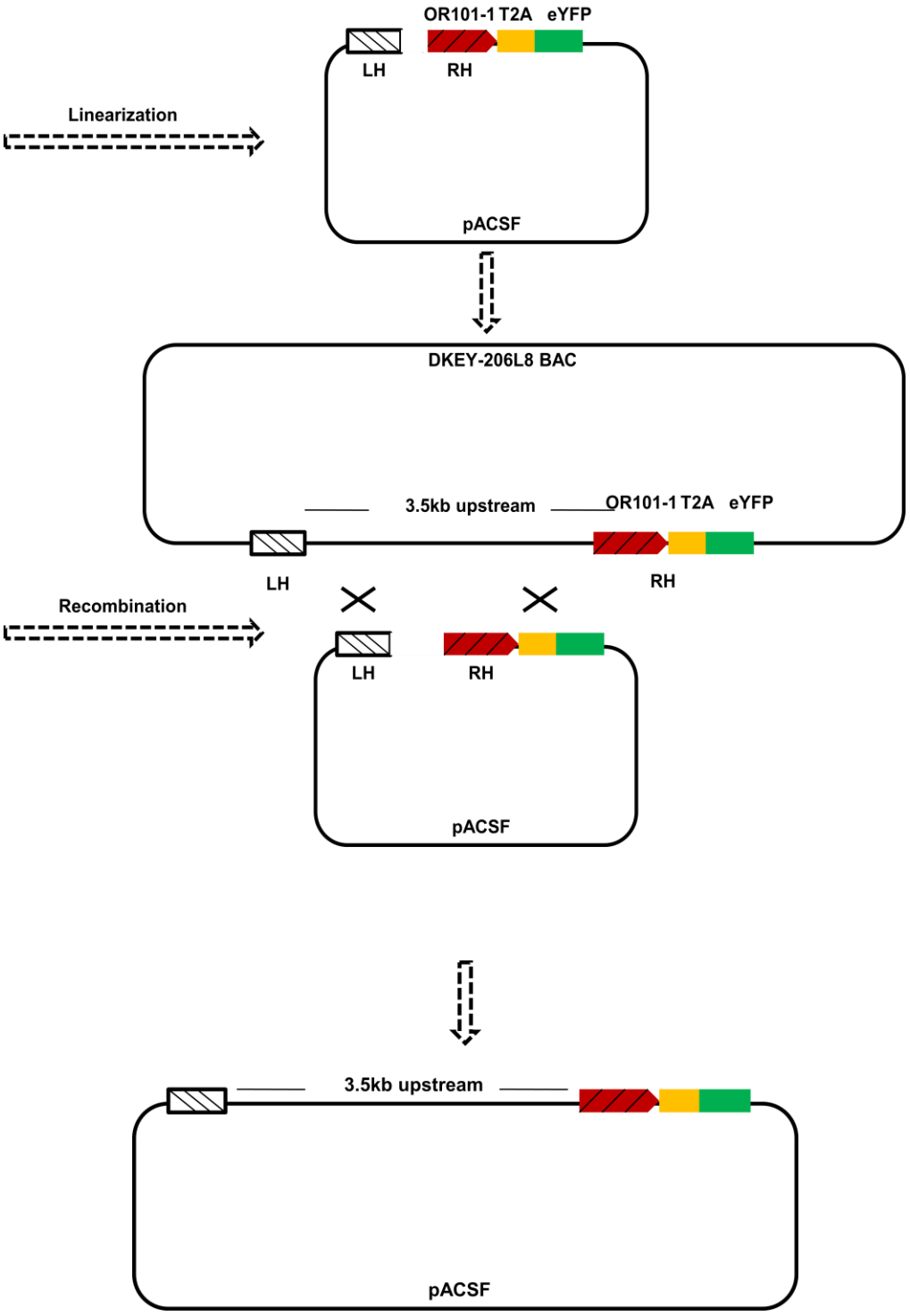


Figure 3.11. Capturing 3,5kb Upstream Genomic Sequence from DKEY-206L8 to pACSF Plasmid Using Gap Repair Protocol.The vector was linearized and the recombination procedure took place as mentioned previously. The 3,5kb genomic sequence in between of two homology arms was captured and transferredinto the pACSF.

OR101-1 deletion constructs with different upstream lengths were designed using sequential deletions and regular cloning methods.

First the 3.5kb upstream genomic sequence was fused with EYFP. To make this construct,(3.5-OR101-1upstr::OR101-1::YFP-pA), the gap repair construct was usedanda 2070bp fragment of this construct was deleted using NsiI and NotI restriction sites, and replaced with 1071bp fragment of(OR101-1upstr::EYFP-pA), which was constructed by Xalid Bayramli, by using the same NsiI and NotI restriction sites(Figure 3.12.).

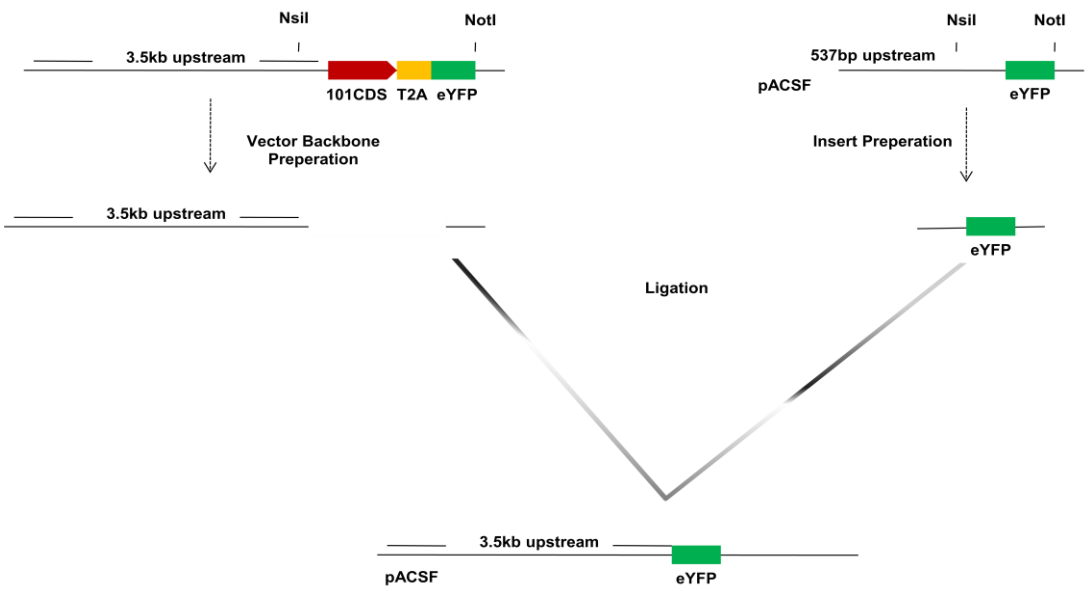


Figure 3.12. Preperation of -3.5-OR101-1upstr::EYFP-pA Transgenic Construct.-3.5-OR101-1upstr::EYFP-pA transgenic construct was prepared by using NsiI and NotI restriction sites on the -3.5-OR101-1upstr::OR101-1::EYFP-pA and -537bp-OR101-1upstr::EYFP-pA constructs.

3.2.22. Injection of DNA into Zebrafish Embryos

Freshly fertilized zebrafish eggs were collected in the morning and sorted on the side of a petri dish by pipetting out all the water. Glass capillary microinjection needles were filled with 2 µl of plasmid solution which consisted of 100 ng / µl target plasmid, 50ng/µl pOMP::mCherry-pA, 0.1% KCl and 0.01% Phenol-Red in distilled water. Approximately, 4 nanoliters of plasmid solution was injected into the cytoplasm of single-cell embryos using FemtoJet® Express (Eppendorf). The embryos were raised in E3 medium at 28^{o}C in an incubator for expression analyses.

3.2.23. Analysis of Transgene Expression

3 dpf injected zebrafish embryos were screened under a fluorescent stereomicroscope using a 2x objective lens. Transgene-expressing embryos were selected and further analyzed under a LEICA SP5-AOBS confocal microscope for verification of expression in the OE. The un-hatched embryos were manually dechorionated and positive fish were fixed overnight in 4% paraformaldehyde at 4^{o}C.

3.2.24. Removal of Pigments by Hydrogen Peroxide Treatment

Prefixed embryos were washed once with 1x PBS to remove the paraformaldehyde fixative. The embryos were incubated at room temperature in a freshly prepared 3% H_2O_2/0.5% KOH solution until all pigment disappeared. The embryos were washed for 4 minutes in 1X PBS to stop the bleaching reaction and progressively they were dehydrated for 5 minutes in 25% (vol/vol), 50% (vol/vol) 75% (vol/vol) methanol in 1x PBS. The dehydration reaction was stopped by placing the embryos in 100% Methanol. Dehydrated embryos were kept at -20^{o}C freezer until further processed for immunohistochemistry.

3.2.25. Immunohistochemistry

Dehydrated embryos were transferred into microfuge tubes or 24-well plates and rehydrated by successive incubations in 75% MeOH/25% PBS for 5 minutes, 50% MeOH/50% PBS for 5 minutes, 25% MeOH/75% PBS for 5 minutes with no agitation. Then they were transferred into 100% PBX solution (1xPBS containing 0,1-1,0% Triton X-100) for 4 times 5 minutes with rocking agitation. Blocking solution was prepared by adding 10% normal goat serum and 0,5% BSA into PBX, and embryos were transferred into blocking buffer and incubation in the blocking solution overnight at 4^{o}C with rocking agitation. The following day, primary antibodies were diluted to desired concentration in blocking bufferand embryos were incubated in antibody solution at 4^{o}C overnight with rocking agitation. The next day, the antibody was washed with PBX briefly, then 5x 5 minutes intervals and then 3x 20 minutes intervals.The embryos were blocked again as mentioned above, and on the next day secondary antibody was applied in blocking solution. Following overnight incubation at 4^{o}C with rocking agitation, unbound secondary antibody was washed away with PBX by 5x 5 minutes intervals and then 3x 20 minutes intervals. Finally, embryos were transferred to PBS for further documentation or mounting for microscopic imaging purposes.

4. RESULTS

The olfactory subgenome can comprise up to 3% of all genes of an organism in some cases. Typically, ORgenes are arranged in clusters in distinct loci that are located on most chromosomes. OR gene expression was previously shown to be monoallellic and monogenic in the mouse but this appears to hold true in the zebrafish olfactory system as well (Chess *et al.*, 1994; Malnic *et al.*, 1999; Serizawa *et al.*, 2000, 2003). Several hypothesis were proposed to explain how OR genes are selected for expression, most of which have in common that during the development of OSNs, OR gene loci may become activated by either a stochastic process or a limiting factor, allowing only a small number of OR gene loci in the genome to be capable of expressing OR mRNA in a given OSN. In addition, long-range interacting genomic elements, such as the H- (Fuss *et al.*, 2007, Nishizumi *et al.*, 2007) and P- elements in mouse (Vassalli *et al.*, 2002; Rothman *et al.*,2005), and the E15 elements in zebrafish (Nishizumi *et al.*, 2007), have been experimentally shown to affect OR gene expression in some cases. As soon as the OR protein starts to accumulate in the cell membrane, a negative feedback mechanism is initiated to stabilize the gene choice process and silence any other OR genes to be expressed (Serizawa *et al.*, 2003; Lewcock and Reed, 2004; Nguyen *et al.*, 2009). The details of this negative feedback mechanism, however, are mostly unknown.

In this thesis a set of comprehensive transgenic studies is presented where the regulation of a model zebrafish OR gene, OR101 is examined in zebrafish. OR101-1 locates on the minus strand of zebrafish chromosome 21. It comprises a single coding exon, adjacent, but in reverse orientation to the Class I OR115 gene family. According to the possible involvement of long- and short- range transcriptional regulators, four different experimental approaches were used to understand the mechanism of OR101-1 expression, and to identify candidated factors that affect the expression of the OR101-1 gene in zebrafish: a BAC transgenesis approach, 5-RLM RACE, a short promoter approach, and bioinformatic analysis.

For the BAC transgenesis approach, a BAC clone that contains the OR101-1 gene was modified bykeeping upstream- and downstream-flanking sequences intact, and tagging the OR101-1 gene with enhanced yellow fluorescent protein (eYFP) using the *galK*-based method of recombination in bacteria. The modified zebrafish BAC clone DKEY-206L8 was injected into single cell zebrafish oocytes and developing embryos were scored for transgene expression. Scoring of OR101-1 transgene expression was performed at 3dpf by examining the presence of the YFP signals in the OE using fluorescence microscopy.

BAC transgenesis is now widely used in mouse models, but recently also became popular in zebrafish. This method allows the transfer of large DNA sequences (up to 300kb) into a host genome. The integration efficiency of BAC DNA is about 1-3%, and facilitated by the use of specific transposition sites (iTol2) (Yang *et al.*,2006). By using this technique, the target gene or modified experimental locus can be successfully integrated into the host genome including long distance *cis-acting* elements. Another advantage of the approach is that iTol2 BACs might integrate as low copy number integrations, reducing copy number variation side effects. It was previously shown in mouse studies, thatBAC transgenesis can have disadvantegous side effects, when multiple BACs form concatamers and integrate into the genome as multiple copies with various orientations (Chandler *et al.*, 2007). This unwanted situation may create drastic over-expression of a target gene, may lead to silencing of the transgene, or to instability, and genetic lesions both inside and around the transgenes (Garrick *et al.*, 1998), which can sometimes limit the experimental applications.

To avoid these problems, transposon-mediated BAC transgenesis was employed, and BACs were modified to contain iTol2 sites for efficient introduction of single copies of the modified BAC into the genome through cut and paste mechanism (Suster *et al.* 2009a). In this method, modified BAC DNA containing iTol2 sites was injected with an appropriate transposase mRNA, into zebrafish oocytes, and it was expected, that with the help of the transposase, a single copy of the BAC integrated into the genome by recombination. Thus, reliable results could be obtained in the F1 generation of zebrafish embryos (Figure 4.1., and Figure 4.2.).

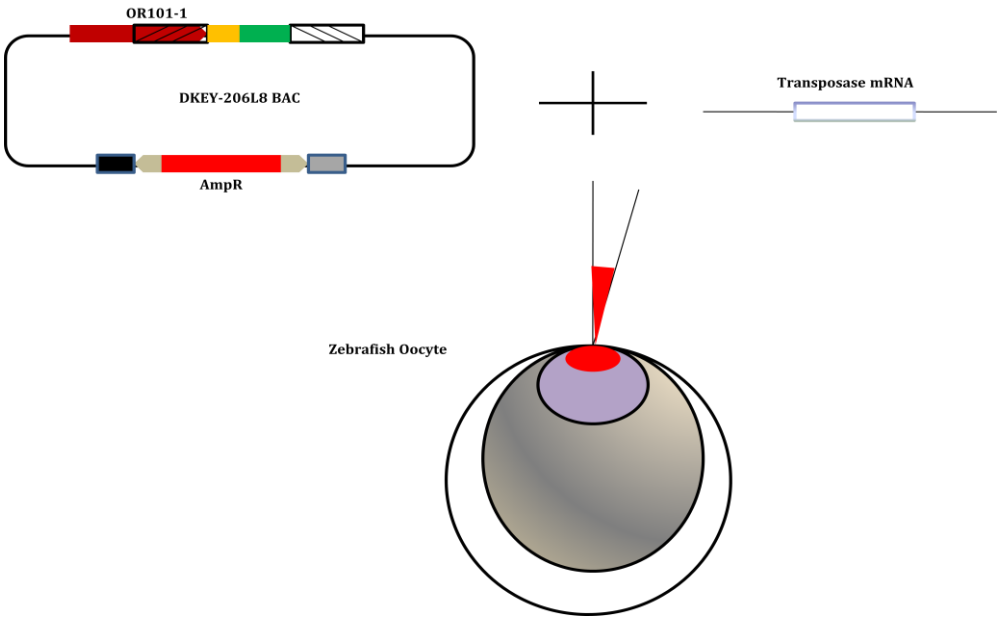


Figure 4.1. OR101-1 T2AEYFP in DKEY 206L8-iTol2 Injection Using Transposase mRNA.Transposase mRNA, which was synthesized *in vitro*, was co-injected with BAC DNA into the single-cell staged zebrafish eggs.

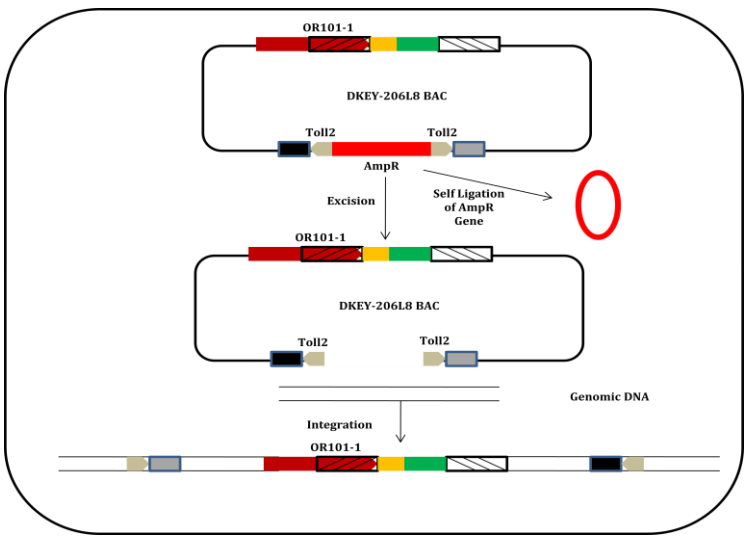


Figure 4.2.Transposase-mediated BAC Transgenesis.ITol2 sites maintain BAC DNA integration into the zebrafish genome by a simple copy and paste action.

Although transposasemediated BAC transgenesis is a widely used technique to obtain stable transgenic animals, and BAC sequences in many cases contain most of the relevant gene regulatory sequences, they have a number of limitations. For instance, their

large sizes are inversely proportional with their integration efficiency. Additionally, these large sequences might contain additional genes, whichcould drasticallyaffect transgenic expression of the target gene or even development of the organism. Therefore, minigenes, comprising only relatively short genomic upstream sequences and fluorescent reporter genes were also employed as experimental tools.

For the proximal promoter approach,minigenes comprising various lengths of genomic sequence upstream of the OR101-1 gene were used and cloned in front of a sequence coding for fluorescent reporter proteins, thereby replacing the OR101-1 coding sequence. Previous experiments showed that in some cases regulatory sequences are very compact and can locate in close distance to the TSSs of OR genes in mouse (Vassalli *et al.*, 2002; Rothman *et al.*,2005) and zebrafish (Mori *et al.*, 2000). Therefore, it was hypothesized that specific sequences in the region immediately upstream of the zebrafish OR101-1 could play a significant role in expression of this gene.

The proximal promoter approach not only resulted in a much higher efficiency of transgene expression when compared to BACs, but also gave much faster results, sincetransgene expression could be observed within 1 day after injection. However, the expression of the constructs was transient, and expression of the fluorescent reporter genes could only be observed up to one week after injection. But because the stability of reporter gene expression was not the primary concern in this study, and transient expression was sufficient for analyses,the proximal promoter approach was preferentially used for identification of candidate regulatory elements.

4.1. Transgenic BAC Injections

First, it was intended to analyze OR101-1-T2A-EYFP transgene expression in order to observe the OR101-1 expression pattern in the zebrafish OE. For this purpose, 300ng of transgenic BAC DNA was injected into single-cell stage fertilized zebrafish oocytes. However, among 462 injected and screened embryos, no transgene expression could

beobserved in the OE. However, 12 of the injected embryos showed ectopic expression mostly inthe striated muscles of the tail myotome and in cardiac muscle cells.

By using the modified iTol2-containing OR101-1-T2A-EYFP BAC and co-injection with Tol2-transposase mRNA, 2 out of 100 injected embryos could be identified that weakly expressed the OR101-1-T2A-EYFP transgene(Figure 4.3.).

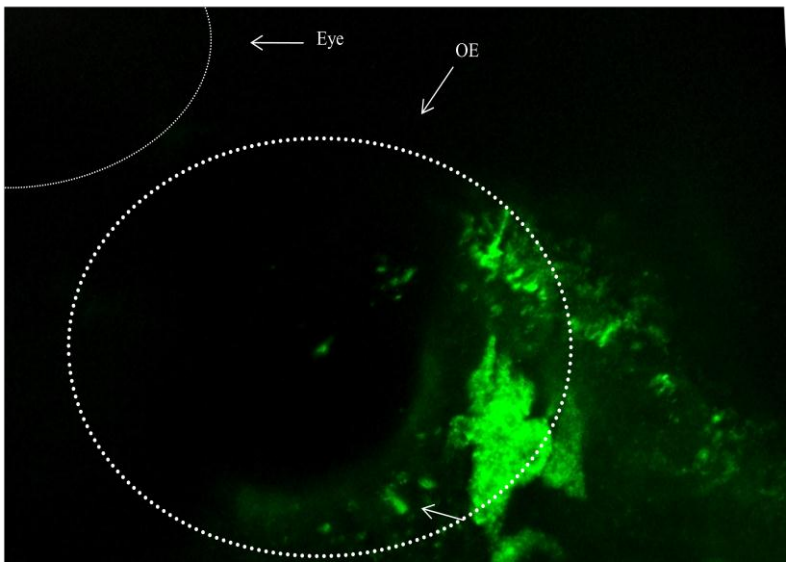


Figure 4.3. OR101-1 T2AEYFP in DKEY 206L8 BAC Transposase Mediated Expression in the OEof 3dpf Zebrafish Embryo.Modified OR101-1 T2AEYFP in DKEY-206L8 BAC with iTol2 sites was co-injected with transposase mRNA into the 1-cell staged zebrafish eggs. After 3 days, 2/100 injected embryos showed weak EYFPexpression.

The unexpectedly inefficient transgene expression of BAC constructs made alternative approaches necessary. Based on the previous findings, it was suspected that the OR101-1 coding sequence, which is included in the transgenic construct, may inhibit expression of the transgene by a negative feedback mechanism (Nguyen *et al.*, 2007, 2010). Similar effects have been reported in mice, and thus the presence of the OR101-1 coding sequence might drastically reduce the efficiency of expression. Therefore, BAC construct previously prepared by Xalid Bayramlı (unpublished), in which the OR101-1 coding sequence has been replaced with the yellow fluorescent protein Venus was used. When this construct was co-injected into fertilized zebrafish eggs with iTol2-transposase mRNA, the frequency of transgene expression slightly increased and 3 out of 60 injected

embryos that survived until 3dpf were positive for Venus expression in the OE (Figure 4.4.).This result, to some degree, supports the idea, that the presence of the OR101-1 coding sequence might affect efficiency of transgene expression by a negative feedback mechanism.

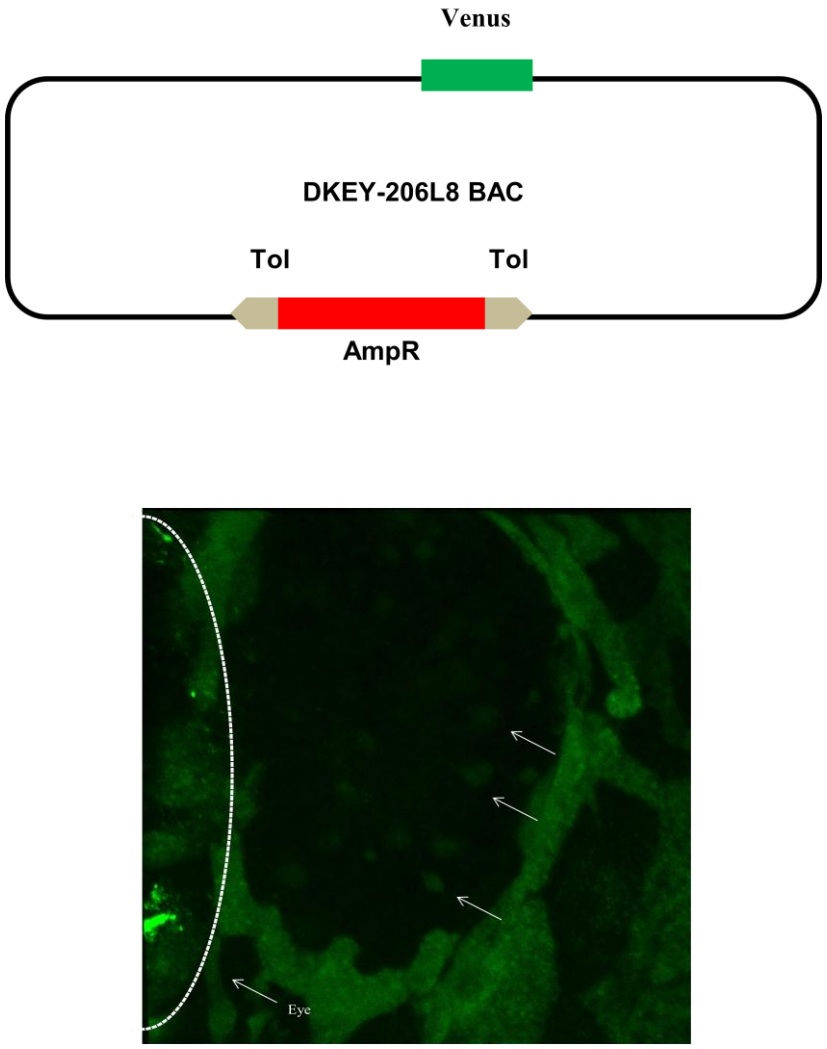


Figure 4.4. OR101-1 Deleted DKEY-206L8 Construct Flanking iTol2 Sites and Transgenic Expression.OR101-1 deletion construct having tol sites in DKEY 206L8 BAC. The reporter protein used for this construct is Venus (Green) (Above). Transgenic expression efficiency was 5% of the OR101-1 deletion in 3dpf zebrafish embryos (Below).

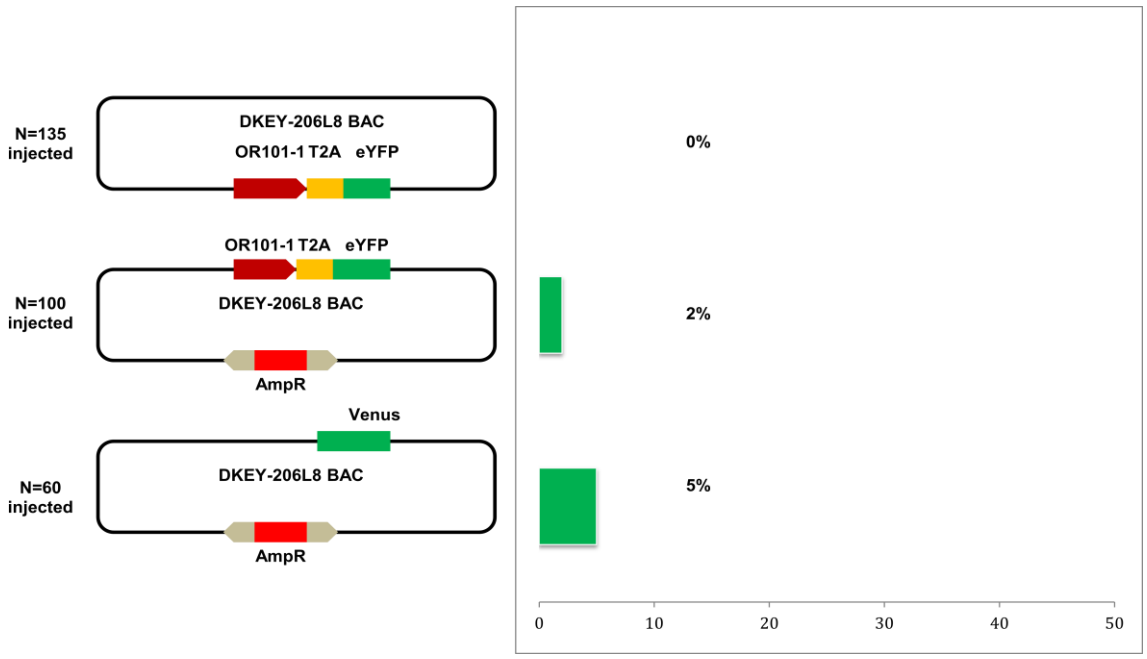


Figure 4.5.Summary of Expression Efficiencies of BAC Transgenes.Expression efficiencies of BAC constructs were shown above. Without transposase co-injection,transgenic expression was not observed. OR101-1 transgenic expression with transposase co-injection was 2%, and OR deletion co-injection efficiency was 5%.

Regardless, the level of reporter protein expression and the number of reporter gene-positive cells was not sufficient for detailed studies on regulatory elements of OR expression. Only weakly expressing cells were obtained at low frequency (Figure 4.5.). This might be because, even though the BAC sequence contains most of 5'- and 3'-flanking sequencesimportant for OR101-1 expression, they did not integrate sufficiently into the genome at early stages. Additionally, the presence of other genes on the BAC might reduce the expression probability of the OR101-1 gene. In order to overcome these limitations, it was decided to construct minigenes, containing a candidate promoter region of OR101-1 along with OR101-1 coding sequence.

4.2. Structure of OR101-1 Gene

Before engaging in a detailed promoter analysis of the OR101-1 gene, it was necessary to examine the cDNA structure of OR101-1 and to map the ocation of the TSS within the sequence upstream of the OR101-1 coding sequence. This analysis would be informative in a sense that would allow for the selection of sequences upstream of the TSS, which could have promoter ctivity. For this reason, the 5'-RNA Ligase-Mediated Rapid Amplification of cDNA Ends (5'-RLM RACE) was performed.

For the 5'-RLM RACE experiment, 3 different OR101-1 gene-specific RACE primers *(See Materials and Methods),* which locate 85bp,316bp, and353bp inside the OR101-1 codon sequence were designed. Outer and inner universal RACE primers were supplied by the kit andwere complementary to the adapter sequence thatwas ligated during the 5'-RLM RACE cDNA preparation step(Figure 4.6., and Figure 4.7.).

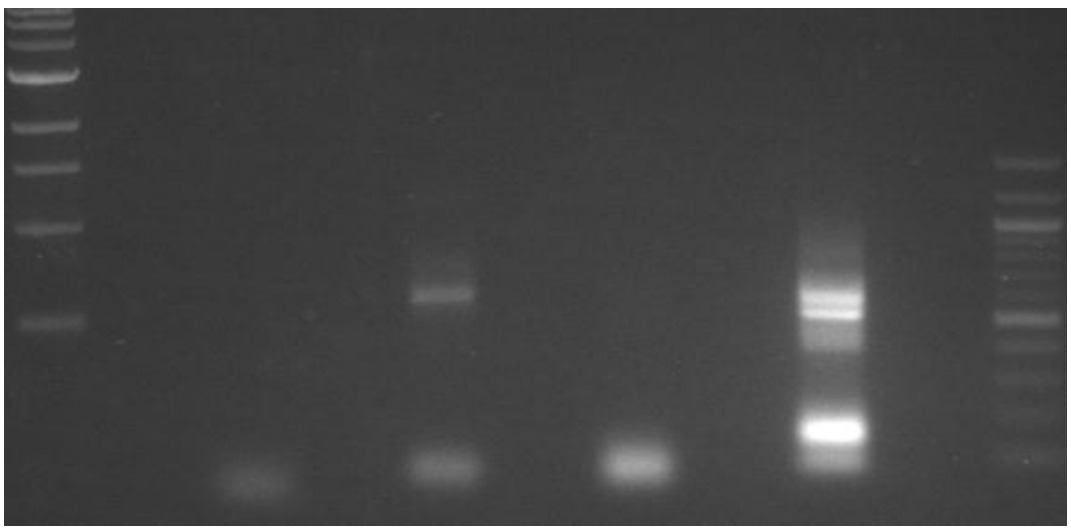


Figure 4.6.5'RLM RACE Analysis with 3 Sets of Gene Specific Primer Trials.(Lanes2-3-4) The PCRs were conducted using OuterRACE_F/101_5'RACE_R1, OuterRACE_F 101_5'RACE_R2, and OuterRACE_F/101_5'RACE_R3.(Lane 5) Nested PCR with the primers InnerRACE_F/ 101_5'RACE_Reverse3.

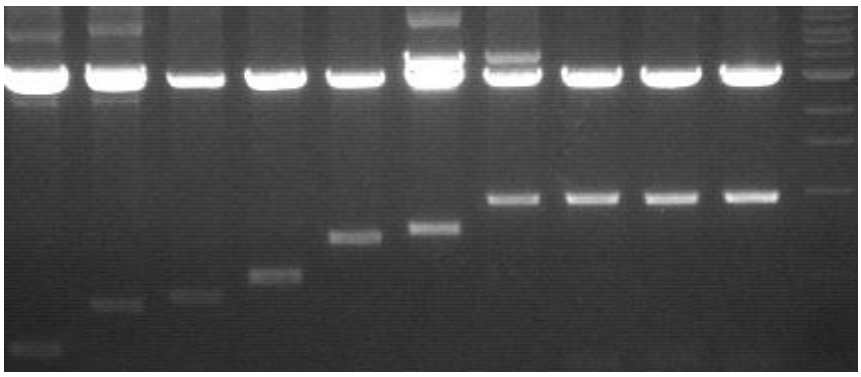


Figure 4.7.EcoR1 Analytical Digestion of pGEM®-T Vector, Containing 5' RACE Products.The first 7 plasmid sent for sequencing and the 7th plasmid matched to the OR101-1 gene locus on zebrafish chromosome 21.

RACE results revealed that the TSS of OR101-1 gene locates 584bp upstream of the OR101-1 coding sequence and consist of an185bp5'-non-coding exon followed by a 350bp intronic region and 49bp non-coding sequence of the second exon(Figure 4.8.).

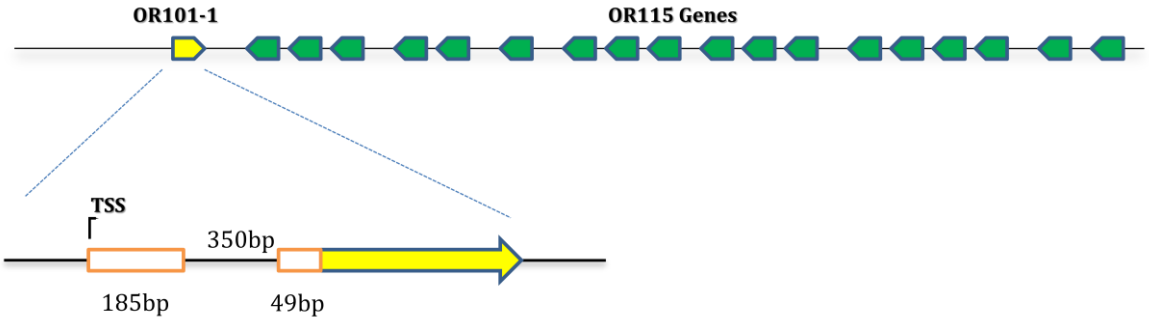


Figure 4.8.The Location and Structure of the OR101-1 Gene.OR101-1 gene locates downstream of the OR 115 gene family. The 5'RLM RACE results indicate that the TSS of OR101-1 is 584bp upstream of the coding sequence, containing 185bp of non-coding 1st exon, 350bp intron, and 49bp non-coding 2nd exon.

4.3. Gap Repair cloning of 3.5kb of Genomic Upstr. of the OR101-1 CDS T2A EYFP Construct

Short promoter constructs used in the promoter bashing approach were derived from a construct which was cloned by capturing 3.5kb of genomic sequence upstream of OR101-1 into a vector that contained the OR101-1 coding sequence followed by a sequence coding for T2A-EYFP. The T2A-EYFP-modified BAC DKEY-206L8 was used for cloning by sequence retrieval through gap repair (Warming, 2005). Promoter bashing contructs with different lengths or parts of the upstream sequence were prepared after deleting the OR101-1 coding sequence from the retrieved fragment, putting the candidate promoter regions directly in front of a sequence coding for yellow fluorescent protein.

4.4. OR101-1 Transgene Expression Using Proximal Promoter Approach

The longest minigene (-3.5-OR101-1upstr::OR101-1::YFP-pA) containing the full 3.5 kb sequence upstream of OR101-1, was co-injected with pOMP::mCherry-pAas a control for injection efficiency.Injection of the pOMP::mCherry-pA constructs results in reliable transgene expression in up to 90% of injected embryos. Thus, YFP expression from the minigenes could be normalized according to the efficiency of mCherry expression driven by the pOMP promoter.

Among the 248 embryos that were injected with the -3.5-OR101-1upstr::OR101-1::YFP-pA construct and survived until 3 dpf but none of them expressed the transgene in the OE, while 107 of them expressed the pOMP::mCherry-pA control construct. Unexpectedly, strong YFP signals were observed in striated muscle cells of the tail and in cardiac muscle cells. To further analyze the source of this ectopic expression, a plasmid containing only OR101-1::YFP-pA, but lacking the putative promoter sequence was constructed. This plasmid was subsequently co-injected with pOMP::mCherry-pA and 45 out of 51 surviving embryos expressed EYFP ectopically in striated muscle cells. This result indicates that either the OR101-1 coding sequence or the T2A sequence promotes ectopic transgene expression in muscle cells by interacting with muscle specific transcriptional regulators.

Thus, similar to the results obtained for BAC transgenesis, short promoter constructs that contain the OR101-1 coding sequence show insufficient reporter gene expression in olfactory tissue. It was previously suspected, that the OR101-1 coding sequence itself negatively affects expression of the constructs. To test for this possibility, the OR101-1 coding sequence was removed and replaced with a sequence coding for the yellow fluorescent protein.

4.5. OR101-1 Promoter Bashing Approach

Promoter bashing experiments were performed to identify sequences that drive reporter gene expression with high efficiency and selectivity in olfactory sensory neurons. Therefore, series of expression constructs with different lengths of genomic upstream sequence were engineered. An initial set of constructscontained 3.5kb (-3.5-OR101-1upstr::YFP-pA), 2.5kb (-2.5-OR101-1upstr::YFP-pA), 2.0kb (-2.0-OR101-1upstr::YFP-pA), and 1.2kb(-1.2-OR101-1upstr::YFP-pA) of upstream genomic sequence fused toEYFP and followed by an SV40 polyadenylation signal.

4.5.1. 3.5kb Upstream Genomic Sequence Analysis

Firstlythe -3.5-OR101-1upstr::YFP-pAconsruct was tested by co-injecting it into one cell zebrafish oocytes with pOMP::mCherry-pA as a control for injection efficiency(Figure 4.9.).

Embryos were scored for transgenic expression by fluorescent microscopy for the EYFP marker in live fish at 2 dpf and after immunohistochemistry for EYFP and mCherry at 3 dpf. From 458 injected embryos that survived until 3dpf, 205 expressed the pOMP::mCherry-pA construct. Of those, 25 embryos also expressed the -3.5-OR101-1upstr::YFP-pA construct. Thus, the normalized transgenic expression efficiency for this construct was 12.1%. The results are consistent with the notion that removal of the OR101-1 coding sequence drastically improves the efficiency of transgenic constructs.

The positive, transgene expressing cells were true sensory neurons in the OE, showing the typical round cell bodies and long dendrites that extended to the nasal cavity of OE. On occasion an axon emerging from the cell body could be distinguished.

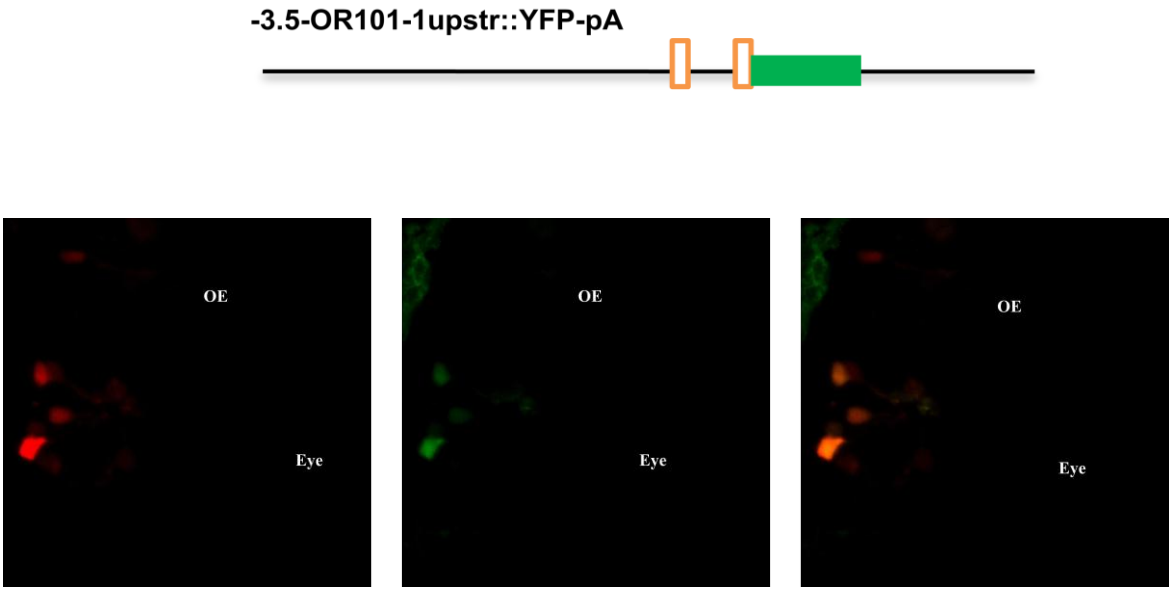


Figure 4.9.-3.5-OR101-1upstr::YFP-pA/pOMP::mCherry-pACo-expression.pOMP::mCherry-pA in pTol (left) and -3.5-OR101-1upstr::YFP-pAin pACSF (middle)were co-injected into single cell stage zebrafish eggs. The overlaid imageof two expression patterns(right). *In vivo* and/or IHC imaging was done on the 3 dpf zebrafish embryos.

4.5.2. 1.2kb Upstream Genomic Sequence Analysis

It has been reported that shorter transgenic constructs can have a higher transgenic efficiency compared to constructs containing longer sequences (Mori *et al.*, 2000, Argo *et al*., 2003). To test this possibility a shorter construct containing only 1.2kb of upstream sequence was engineered. The -1.2-OR101-1upstr::YFP-pAwas prepared by using direct PCR amplification and tested by injection into fertilized zebrafish oocytes (Figure 4.10.). The results of the co-injection of the shorter -1.2-OR101-1upstr::YFP-pA and pOMP-mCherry constructs was surprising. Among 881 fish that survived for 3 days after injection, 642 embryos were positive for pOMP::mCherry-pA. Out of those 394 fish also expressedyellow fluorescent protein from the -1.2-OR101-1upstr::YFP-pA. Thus, the

efficiency of expression, standardized to the number of pOMP::mCherry-pA-expressing embryos,increased to 61.3% when compared to the 12.1% observed for the -3.5-OR101-1upstr::YFP-pA construct. This result suggests that positive regulatory sequences affecting OR101-1 expression are located within the first 1.2kb directly upstream of OR101 coding sequence. It also suggests that negative regulators of expression might be located between -1.2 and -3.5 upstream of the OR101-1 coding sequence. In subsequent experiments the exact location of these candidate negative regulatory sites was further investigated.

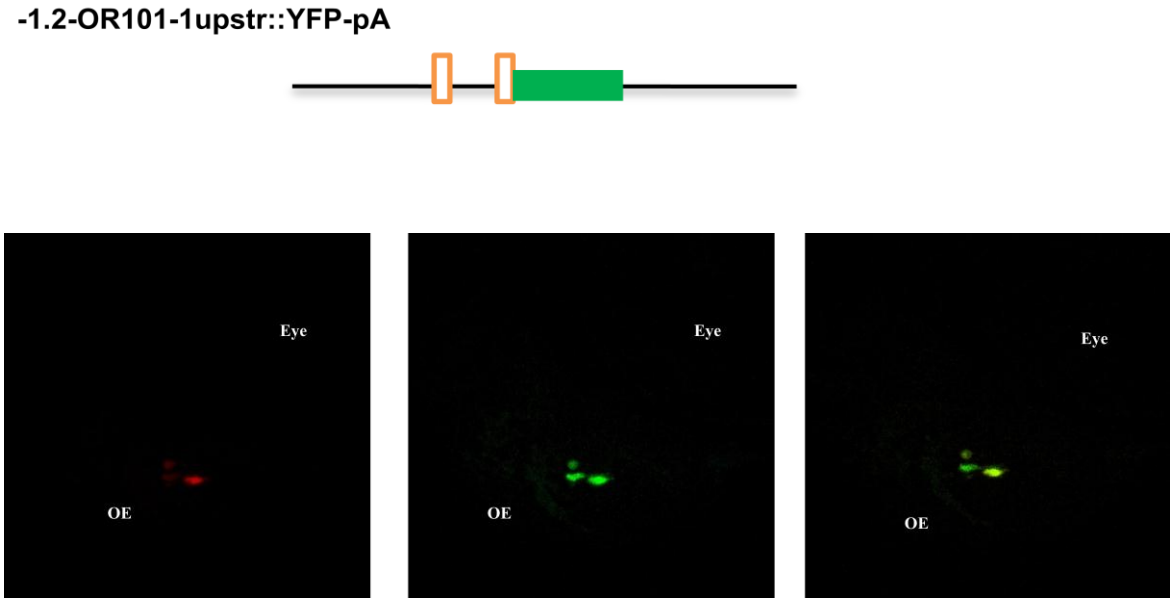


Figure 4.10.-1.2-OR101-1upstr::YFP-pA/pOMP::mCherry-pA Co-expression.-pOMP::mCherry-pA in pTol (left) and -1.2-OR101-1upstr::YFP-pAin pACSF (middle) were co-injected into single cell stage zebrafish eggs. The overlaid image of two expression patterns(right). *In vivo* and/or IHC imaging was done on the 3 dpf zebrafish embryos.

4.5.3. 2.5kb Upstream Genomic Sequence Analysis

A strong decrease of efficiency of transgene expression was observed when the length of the promoter constructs was increased from 1.2kb to 3.5kb. To identify the breakpoint of efficiency within the candidate OR101-1 promoter a construct, containing 2.5kb of genomic upstream sequence was engineered (Figure 4.11.).The expression efficiency of this construct, when compared to the efficiency of the -1.2-OR101-1upstr::YFP-pA and -3.5-OR101-1upstr::YFP-pA constructs would show where a potential

inhibitory sequence might be located. If, for instance, the inhibitory sequence would be more upstream than 2.5kb the efficiency of expression would be similar to the efficiency observed for the -1.2-OR101-1upstr::YFP-pA construct, whereas if the potential inhibitory sequence were located within 2.5kb upstream of the OR101-1 gene, the efficiency would be expected to be more similar to the results obtained for the -3.5-OR101-1upstr::YFP-pA construct.

The analysis revealed that among 568 fish that survived up to the day of analysis at 3dpf, 290 fish expressed the positive control pOMP::mCherry-pA. Out of those, 37 fish were positive for EYFP from the -2.5-OR101-1upstr::YFP-pA construct. The normalized expression efficiency was thus 12.7% for the-2.5-OR101-1upstr::YFP-pAconstruct. This efficiency of 12.7% is very similar to the efficiency of the -3.5-OR101-1upstr::YFP-pAconstructs but significantly lower than the efficiency that was observed for -1.2-OR101-1upstr::YFP-pA construct. These data suggested that a potential inhibitory sequence might be located between 2.5kb and 1.2kb upstream of the OR101-1 gene.

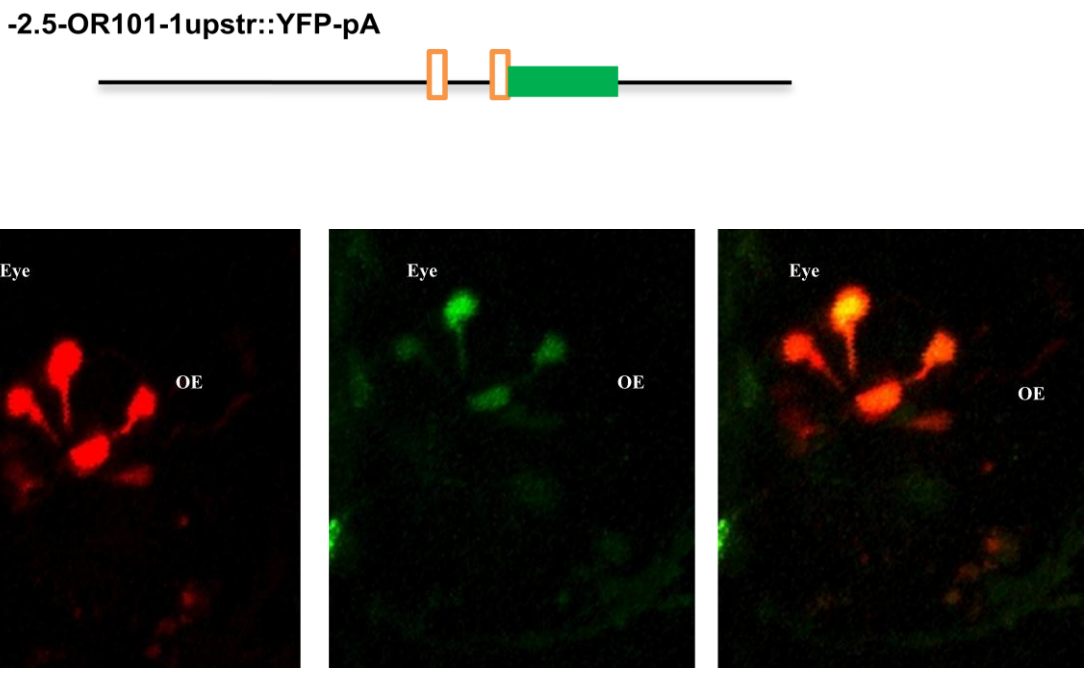


Figure 4.11.-2.5-OR101-1upstr::YFP-pA /pOMP::mCherry-pA Co-expression.pOMP::mCherry-pA in pTol (left) and -2.5-OR101-1upstr::YFP-pAin pACSF (middle) were co-injected into single cell stage zebrafish eggs. The overlaid image of two expression patterns (right). *In vivo* and/or IHC imaging was done on the 3 dpf zebrafish embryos.

4.5.4. 2.0kb Upstream Genomic Sequence Analysis

To further narrow down the location of a potential inhibitory regulatory site a construct containing 2kb of upstream sequence (-2.0-OR101-1upstr::YFP-pA) was engineered and analyzed as before (Figure 4.12.).Out of 512 injected fish thatsurvived up to 3dpf, 310 of themexpressedpOMP::mCherry-pA. Among those, 200 were positive for -2.0-OR101-1upstr::YFP-pA. Thus, the normalized expression efficiency for this construct was 64.5 %, which is very similar to the transgenic expression efficiency of the -1.2-OR101-1upstr::YFP-pA construct but significantly higher than the results observed for the -3.5-OR101-1upstr::YFP-pA and -2,5-OR101-1upstr::YFP-pA constructs. This indicates that a potential inhibitory sequence might be located between 2.0kb and 2.5kb upstream of the OR101-1gene (Figure 4.13.).

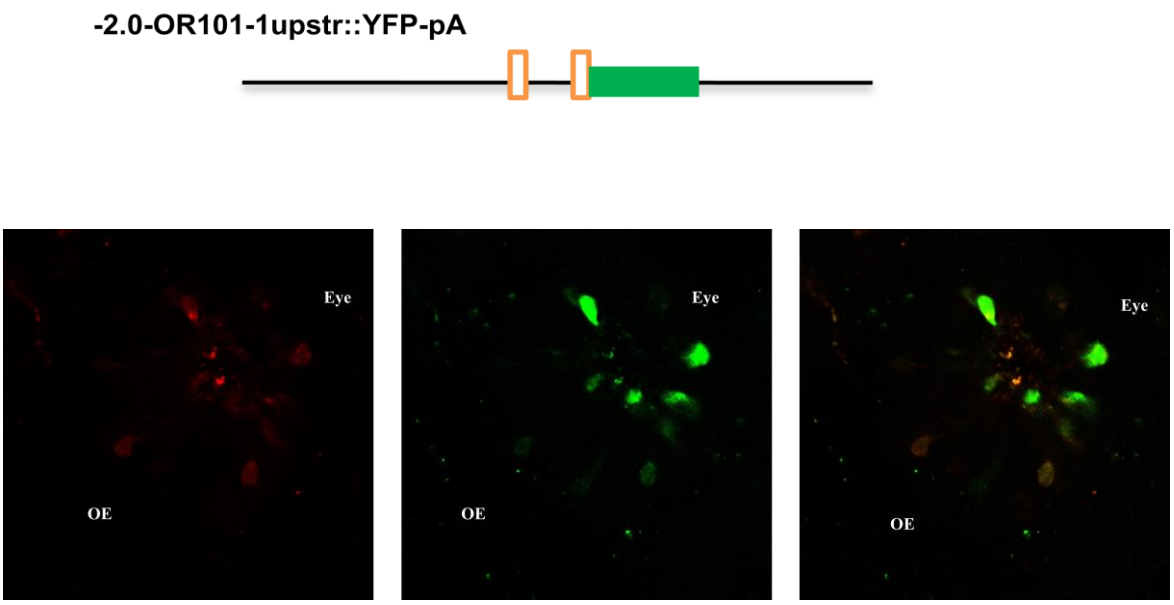


Figure 4.12.-2.0-OR101-1upstr::EYFP-pA/pOMP::mCherry-pA Co-expression.pOMP::mCherry-pA in pTol (left) and -2.0-OR101-1upstr::YFP-pA in pACSF (middle) were co-injected into single cell stage zebrafish eggs. The overlaid image of two expression patterns (right). *In vivo* and/or IHC imaging was done on the 3 dpf zebrafish embryos.

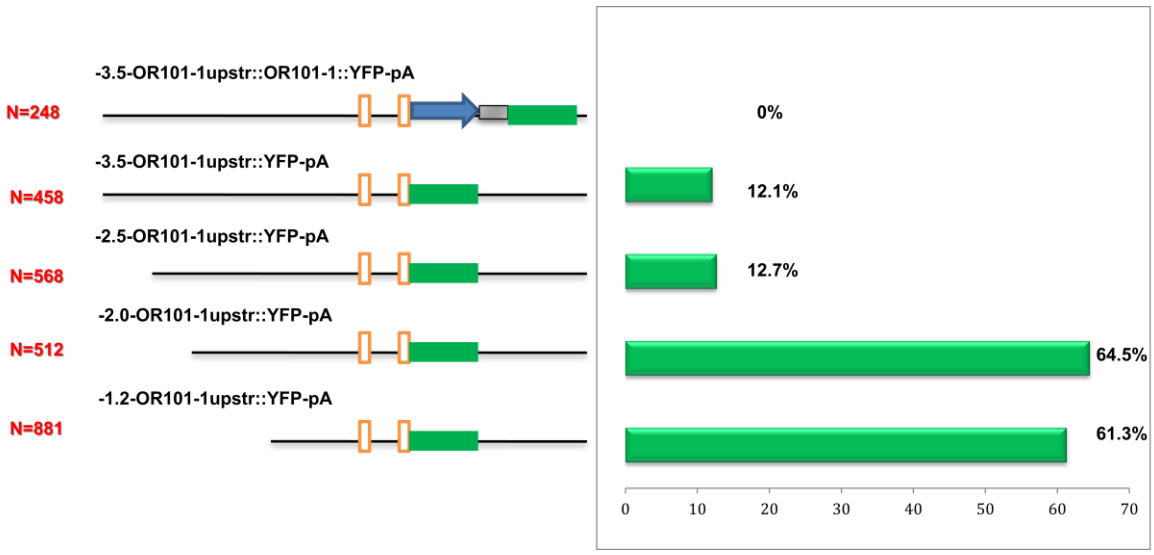


Figure 4.13.Summary of Transgenic Constructs and Their Expression Efficiencies.Initially five transgenic constructs with different upstream lengths were prepared. (Left) Longer promoters gave around 12% expression efficiency, while shorter promoters gave around 64% expression efficiency. OR101-1 CDS containing construct did not expressed in zebrafish OE.

4.6. Analysis of Potential Inhibitory Sequence of OR101-1 Gene Promoter

The reduction of the expression ratio from 64.5% to 12.7% between the 2.0kb upstream genomic sequence and 2.5kb upstream sequence indicates that the 500bp sequence, which is different between those two constructs, might contain binding sites for a negative regulatory factor of OR gene expression. In order to further characterize this potentially inhibitory sequence, 5 different approaches were used: 1) deletion of the 500bp potential inhibitory sequence from the full 3.5kb sequence of the -3.5-OR101-1upstr::YFP-pA construct, 2) cloning of the 500bp inhibitory sequence in front of the -1.2-OR101-1upstr::YFP-pAsequence, 3) Deletion of the first half of 500bp potential inhibitory sequence from the -3.5-OR101-1upstr::YFP-pAsequence, 4) deletion of the second half of the 500bp potential inhibitory sequence from -3.5-OR101-1upstr::YFP-pAsequence, and 5) cloning of the H-enhancer in frontof the -2.5-OR101-1upstr::YFP-pAsequence.

4.6.1. 500bp Potential Inhibitor + 1.2kb Upstream Genomic Sequence

The verification of an inhibitory effect of the 500bp sequence on transgene expression was first performed by cloning this sequence in front of a promoter construct with high transgene efficiency. The shortest working promoter constructs, the -1.2-OR101-1upstr::YFP-pAsequence, was selected for this purpose. Ifthe 500bp sequence contains any binding sites for negative regulators, it should confer this effect to the highly efficient -1.2kb construct(Figure 4.14). The co-injection results of the -500i-1.2OR101-1upstr::YFP-pA and pOMP::mCherry-pA plasmids showed,that among 71 injected and surviving fish, only 8 of them expressed EYFP, while 57 embryos expressed pOMP::mCherry-pA. Thus, the resulting efficiency of the -1.2kb construct decreased from 61.3% to 14 % in the presence of the 500bp sequence, supporting the hypothesis that the 500bp sequence contains binding site(s) for negative regulators of transcription.

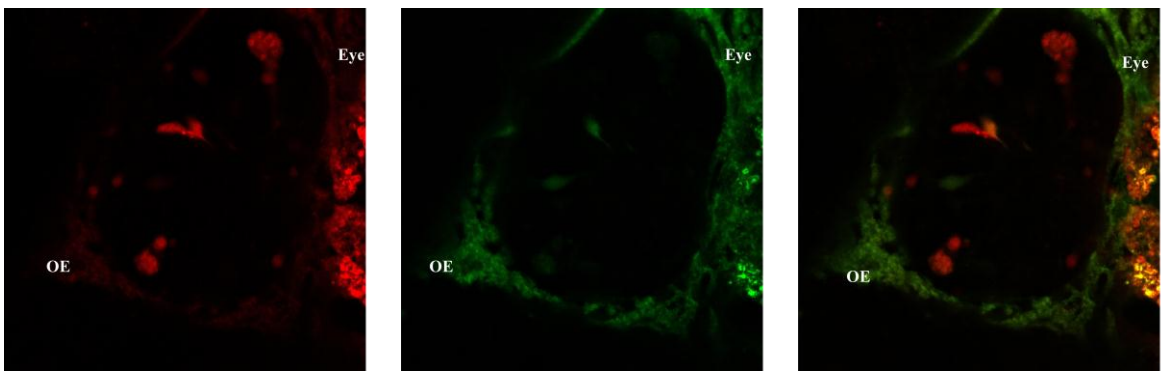


Figure 4.14.-500i+1.2-OR101-1upstr::YFP-pA/pOMP::mCherry-pA Co-expression.pOMP::mCherry-pA in pTol (left) and -500i+1.2-OR101-1upstr::YFP-pA in pACSF (middle) were co-injected into single cell stage zebrafish eggs. The overlaid image of two expression patterns (right). *In vivo* and/or IHC imaging was done on the 3 dpf zebrafish embryos.

4.6.2. 3.5kb Upstr. Genomic Sequence Analysis, Potential Inhibitor Deletion

Similarly, the deletion of the potential inhibitory sequence from the -3.5-OR101-1upstr::YFP-pA construct should remove the inhibitory effect of 500bp and significantly increase its efficiency. The expression efficiency of -3.5-OR101-1upstr::YFP-pA construct containing the inhibitory sequence was only 12.1%. When the -3.5-OR101-1upstr-Δi::YFP-pA construct was co-injected with pOMP::mCherry-pA, 154 out of 256 injected and surviving embryos expressed the positive control mCherry. Out of those, 62 embryos were positive for -3.5-OR101-1upstr-Δi::YFP-pA. Thus, the removal of the potential inhibitory sequence increased the transgene efficiency from 12.1% to 40.2%, supporting the idea that negative regulatory sites are located within that sequence.

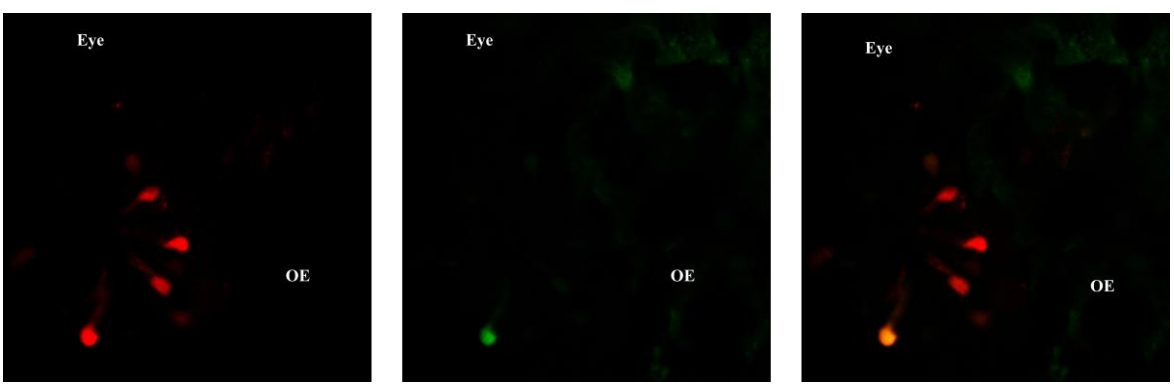


Figure 4.15.-3.5-OR101-1upstr-Δi::YFP-pA/pOMP mCherry Co-expression.pOMP::mCherry-pA in pTol (left) and -3.5-OR101-1upstr-Δi::YFP-pA in pACSF (middle) were co-injected into single cell stage zebrafish eggs. The overlaid image of two expression patterns (right). *In vivo* and/or IHC imaging was done on the 3 dpf zebrafish embryos.

4.6.3. 3.5 Upstream Genomic Sequence SpeI-StuI (ΔSS) Deletion Analysis

Transcription factor (TF) binding sites are on average 6-8bp in length; however, longer sequences for TF binding can also be found. Therefore, further analyses for identification of the TF binding sites were started by examining two-halves the 500bp sequence. These findings would be helpful to continue with a bioinformatics analyses to identify factors binding within in the inhibitory sequence. For this purposefirst the 5'-half of the 500bp sequence was deleted by digesting the construct with SpeI, bluntingthe sequence and removing the fragment using sequential StuI digestion followed with a CIP reaction and selfligating the vector backbone, resulting in the removal of the 5'-most 321bp of the 500bp candidate inhibitor region (Figure 4.16).

Then, the -3.5-OR101-1upstr-ΔSS::YFP-pA construct was co-injected with pOMP::mCherry-pA and fish were screened for fluorescent reporter gene expression. Among 73 injected and surviving fish, 15 of them expressed the -3.5-OR101-1upstr-ΔSS::YFP-pAtransgene. A total of 57 embryos expressed pOMP::mCherry-pA. Thus, the normalized transgene expression efficiency was 26.3%, which was higher than that observed for constructs containing the entire 500 bp of candidate sequence but lower than that observed for the smaller constructs that did not include the sequence. Thus, removal of the initial 321 bp of the 500bp sequence did not fully restore the high efficiency of transgenic expression that was observed when the full 500bp sequence was removed. Therefore, the 3'-half of the sequence was removed and tested for transgene expression.

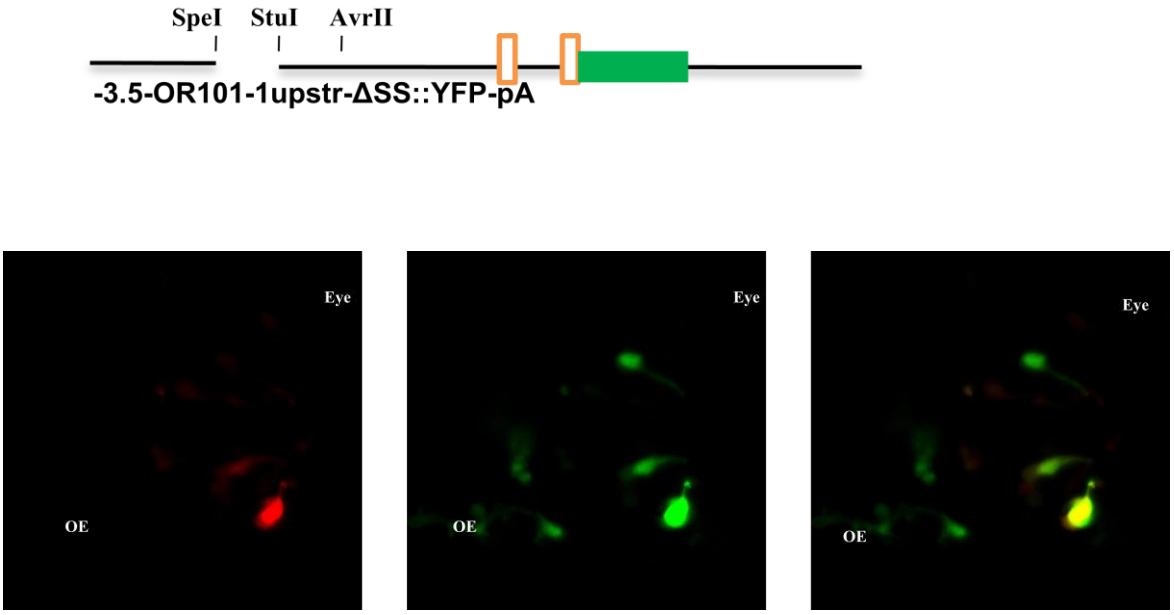


Figure 4.16.-3.5-OR101-1upstr-ΔSS::YFP-pA/pOMP::mCherry Co-expression.pOMP::mCherry-pA in pTol (left) and -3.5-OR101-1upstr-**Δ**SS::YFP-pA in pACSF (middle) were co-injected into single cell stage zebrafish eggs. The overlaid image of two expression patterns (right). *In vivo* and/or IHC imaging was done on the 3 dpf zebrafish embryos.

4.6.4. 3.5 Upstream Genomic Sequence Stu-AvrII (ΔSA) Deletion Analysis

The 3'-half of the 500bp sequence was deleted first digesting the parent construct with AvrII restriction enzyme, followed by a bluntingreaction, digestion with StuI and self-ligation. This procedure removed the 3'-fragment of 235bp from the 500bp (Figure 4.17.).

When the resulting -3.5-OR101-1upstr-ΔSA::YFP-pAconstruct was co-injected with the pOMP::mCherry-pA plasmid, 156 fish survived to day 3, and18 double positive embryos were obtained among 52 pOMP::mCherry-pA expressingfish. Thus, the observed efficiency of this transgenic construct reached 34.6% when normalized to pOMP::mCherry-pA. This result suggested that the 3'-half of the 500bp sequence has a stronger inhibitory effect than the 5'-half. However,none of the partial removals of the inhibitory sequence reached the full efficiency that was observed when the entire 500bp

were deleted from transgenic constructs, suggesting that more than one regulatory site might be located within the sequence.

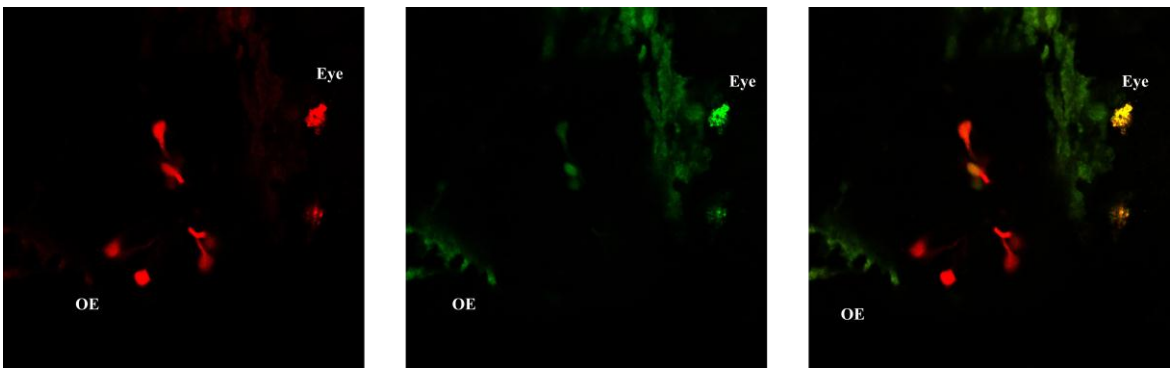


Figure 4.17.-3.5-OR101-1upstr-ΔSA::YFP-pA/pOMP::mCherry-pA Co-expression.pOMP::mCherry-pA in pTol (left) and -3.5-OR101-1upstr-ΔSA::YFP-pA in pACSF (middle) were co-injected into single cell stage zebrafish eggs. The overlaid image of two expression patterns (right). *In vivo* and/or IHC imaging was done on the 3 dpf zebrafish embryos.

4.6.5. H-Enhancer Cloning to the 2.5kb Upstream Genomic Sequence

In a next step, the mouse OR gene enhancer (H-Element) was cloned in front of the --2.5-OR101-1upstr::YFP-pA sequence(Figure 4.18.).In the absence of the H-enhancer, the expression efficiency of -2.5-OR101-1upstr::YFP-pA construct was only 12.7%. It was observed in previous experiments (Bayramli, unpublished) that the presence of the H-enhancer increase expression of the -1.2-OR101-1upstr::YFP-pA construct, but also promotes wide spread expression in the zebrafish OE. In the presence of the H-enhancer, cell types that are not strictly OR-expressing cells but rather VR-expressing cells could be observed as judged by their bottle shaped profiles and by their projection pattern to the

lateral OB. However, when the H-2.5-OR101-1upstr::YFP-pA was co-injected with pOMP::mCherry-pA, the following expression results were observed: Among 56 injected fish, only 7 of them expressed the trans gene and 37 of them expressed pOMP::mCherry-pA. The normalized transgene efficiency, which was 18.9%, was not as high as the other transgenic constructs containing H-element. Thus, the presence of the 500bp sequence might have an inhibitory influence even in the presence of strong enhancer, such as the H-element.

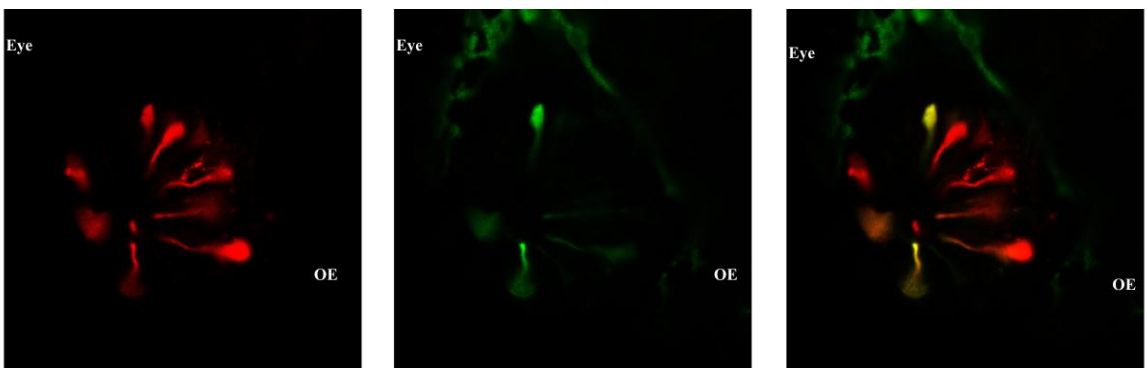


Figure 4.18. H + 2.5-OR101-1upstr::YFP-pA /pOMP::mCherry-pA Co-expression.pOMP::mCherry-pA in pTol (left) and H+2.5-OR101-1upstr::YFP-pA in pACSF (middle) were co-injected into single cell stage zebrafish eggs. The overlaid image of two expression patterns (right). *In vivo* and/or IHC imaging was done on the 3 dpf zebrafish embryos.

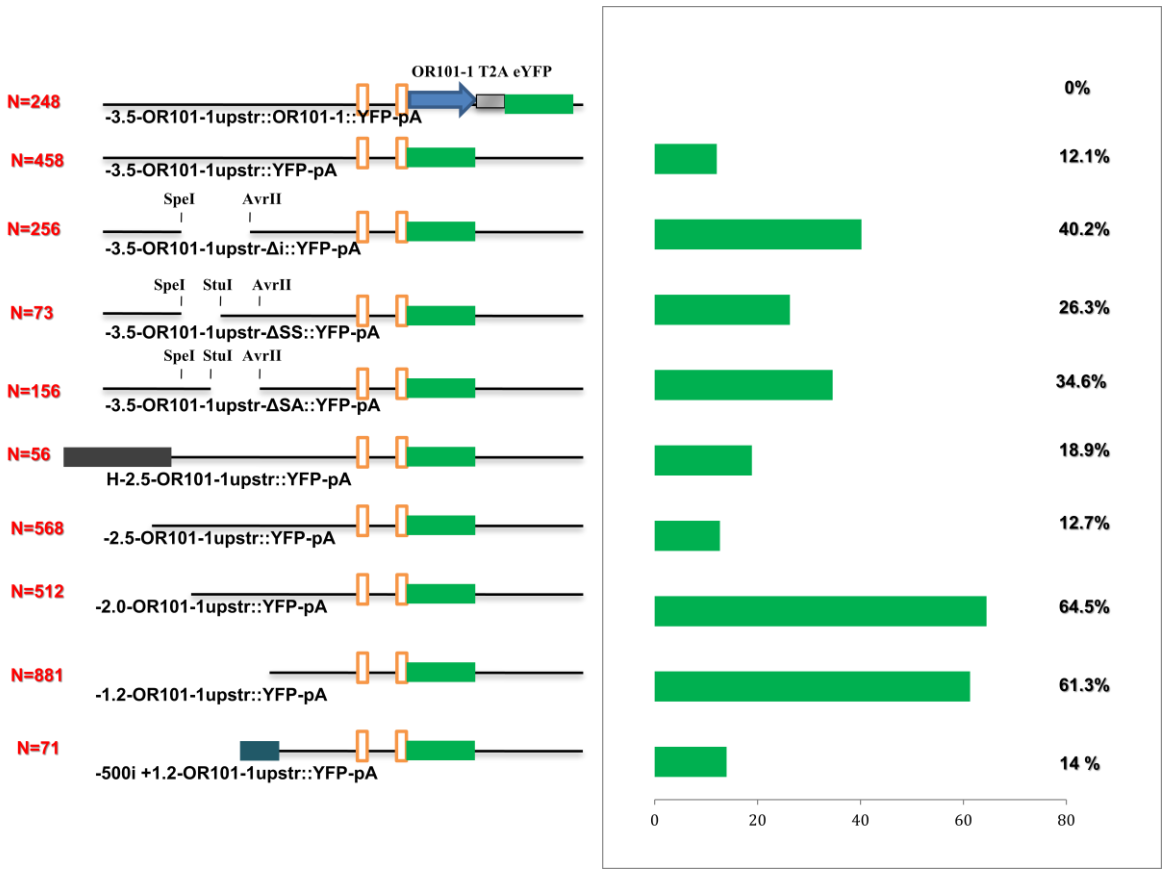


Figure 4.19.Summary of Expression Efficiencies of Ten Different Minigenes.All 10 transgenic constructs were injected at least three times for each construct. (Left) Red numbers indicate the number of surviving fishafter injections. Transgenic constructs were named and depicted according to their lengths of upstream sequences and corresponding deleted sites.

The overall results indicate, that a 500bp sequence located between 2,0kb and 2,5kb upstream of the OR101-1 gene excerts an inhibitory effect on transgene expression. When it was present in the transgenic construct, the expression efficiencies were consistently below 20%. However, when it was removed from the constructs, the expression efficiencies went up to 65%. Interestingly, intermediate results with 26,3 % and 34.6 % were obtained for 5'- and 3'- partial deletions of the 500bp inhibitory sequence, respectively. (Figure 4.19.).

Possible explanations for these results could be that the regulatory motif is located either directly at the junction of the two halfs that were experimentally tested, or that the

inhibition is caused by the presence of the whole 500bp sequence.In addition, multiple binding sites for negatively regulating transcription factors could be present within the 500 bp sequence and work in a combinatorial or complementary fashion.

Additionally, the results obtained for the H-2.5-OR101-1upstr::YFP-pA transgenic expression were surprising,becausethe number of cells expressing H-2.5-OR101-1upstr::YFP-pA was relatively low as compared to H-1.2-OR101-1upstr::YFP-pA, which was prepared by Xalid Bayramlı. The H element has a strong enhancing ability for OR transgene expression in the shorter construct, however, it also induces expression in cells that would normally not express the OR101-1 gene. Therefore, it was decided to analyze the morphology of the cells expressing the various promoter constructs.

4.7. Analysis of Transgene Expressing Cells

The results presented so far indicate that the presence of a 500bp region located between 2kb and 2.5kbupstream of the OR101-1 gene negatively influences transgene expression. This finding raises the question about a possible function of this inhibitory sequence. One plausible hypothesis isthat it might be necessary for expression of the OR101-1 gene in specific OSN subsets. This hypothesis was tested in two ways. First, the total numbers of YFP-positive cells from 10 noses were counted. For this data set, the results of the H-1.2-OR101-1upstr::EYFP-pA was included, which were performed by X. Bayramlı(Figure 4.20).In previous experiments, when the H element was used as an enhancer,a robust increase in the number and shift in the cell type of transgene expressing cells was observed.Therefore it was suspected, thatthe inhibitory effect of the 500bp sequence might not only present itself on the number of transgene expressing embryos, but also on the number and/or the morphology of cells.

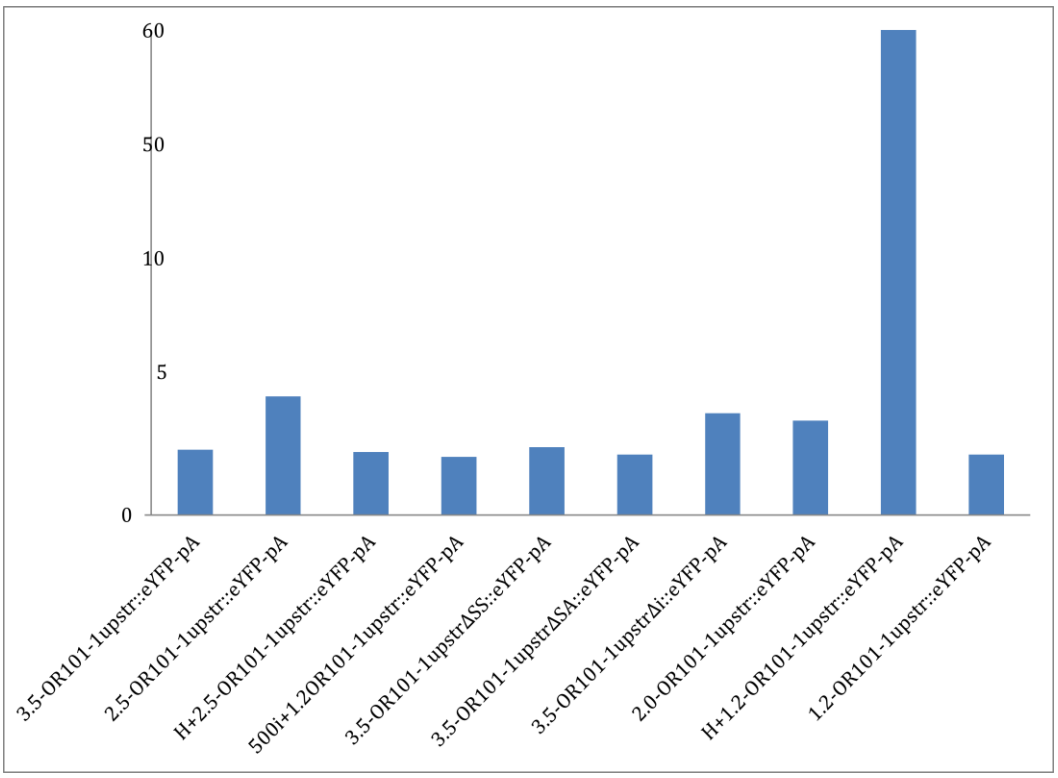


Figure 4.20. Total Number of YFP Positive Cells per Nose.By dividing the total cell numbers from 10 positive nosesto the total nose number for each construct, the average cell number/nose was calculated. There was no significant data verifying that 500bp inhibitory sequence reduces the transgene expressing cell numbers in the OE.

Surprisingly, the bumber of cells expressing the various constructs were not significantly affected by the presence of inhibitory sequence. Indeed, apart from H-1.2-OR101-1upstr::EYFP-pA, the highest number of cells was observed in fish that were injected with the -2.5-OR101-1upstr::EYFP-pA construct, which contains the inhibitory sequence.

Secondly, it was suspected that the inhibitory region might have an effect on the particular type of chemosensory cell that expresses the constructs. Three types of cells express chemosoensory receptors in the zebrafish OE, the microvillous, ciliated, and crypt, which express ORs, V2Rs, and V1Rs, respectively. These cell types can easily be distinguished from each other based on their morphological differences. Therefore, different cell types expressing the various transgenes were counted based on their

morphology differences. Although the sample size was small, a trend of the inhibitory sequence to induce expression mostly to the ciliated cells could be onserved (Figure 4.21.).

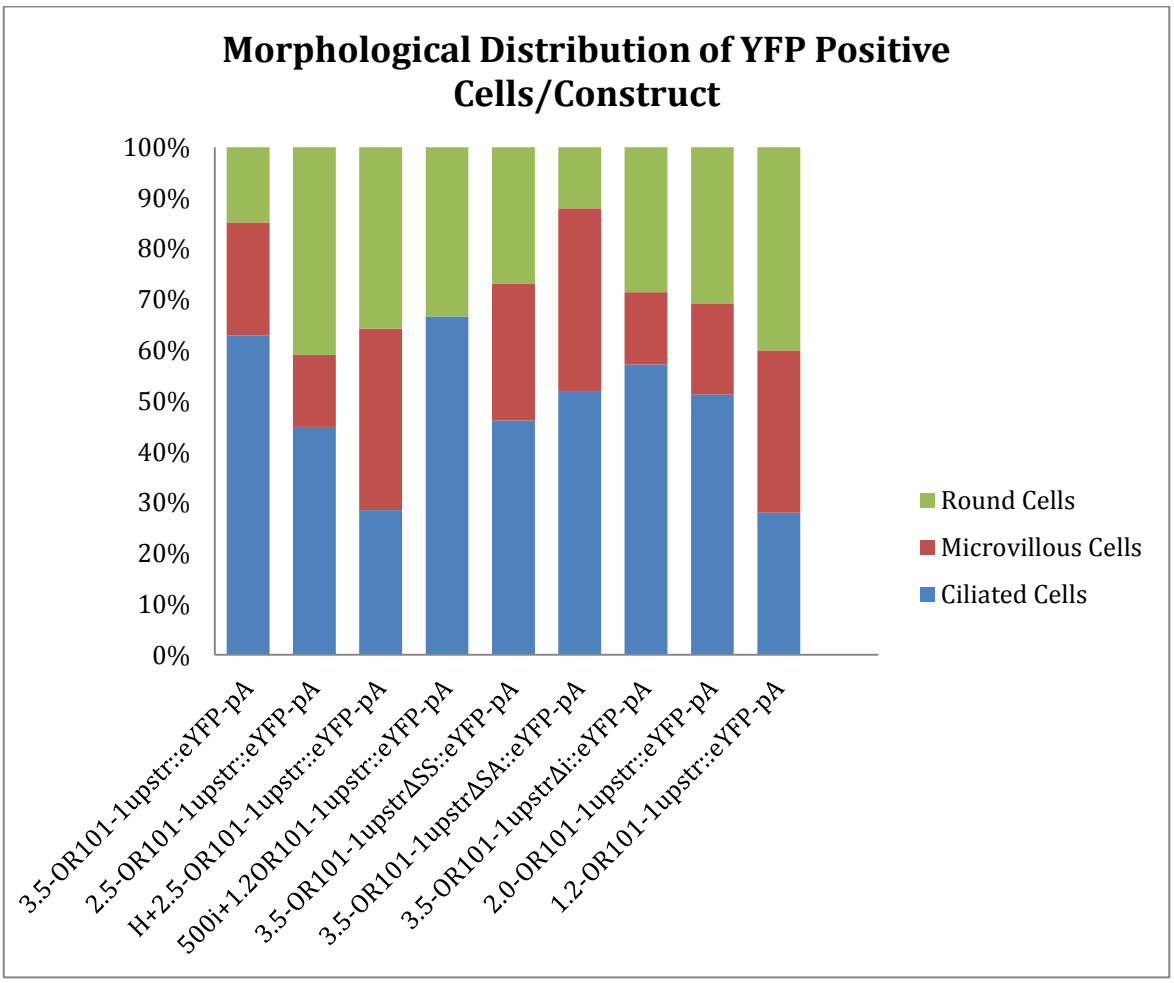


Figure 4.21.Morphological Analysis of the YFP Expressing Cells in the Zebrafish OE. Transgene expressing cell types in zebrafish OE. It can be clearly seen that 500bp inhibitor sequence containing constructs generally lead the expression of EYFP to the ciliated cells.

4.8. Bioinformatics Analysis of 3.5kb upstream Sequence

In parallel with the experimental results, a bioinformatic analysis was conducted to characterize the sequence upstream of the OR101-1 gene. BLAST analysis of the original 3.5kb sequence was performed to identify potentially conserved sequences. It was observed a region between 1836bp and 2509bp is extremely conserved in zebrafish and that related sequences are present on almost all chromosomes(Figure 4.22.).This stretch of high conservation includes the 500bp inhibitory sequence, thus raising the possibility thatthe same sequence might also be a potential inhibitor for other genes.

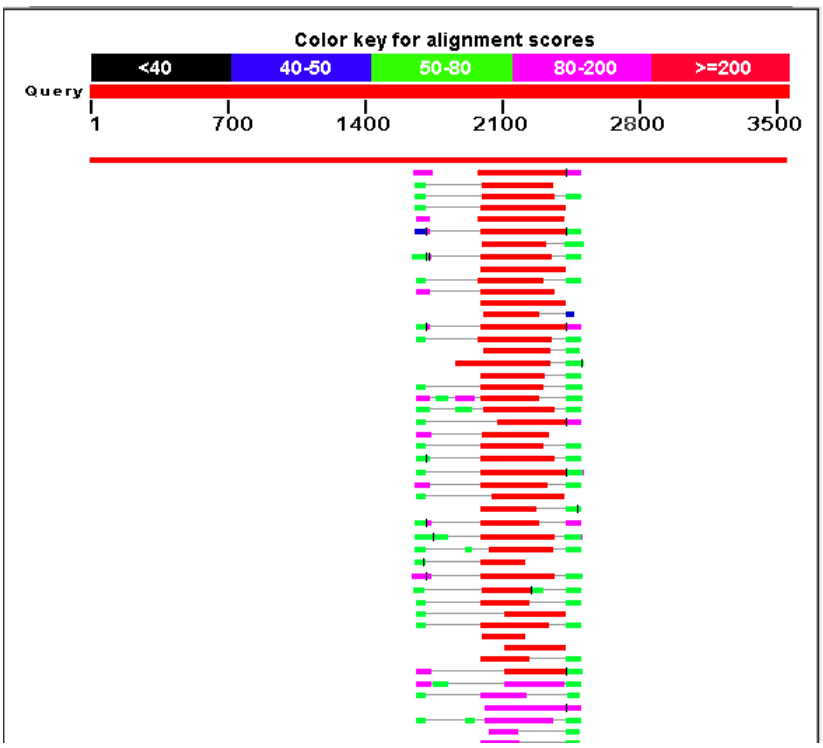


Figure 4.22.Bioinformatics Analysis of 3.5kb Genomic Upstream Sequence ofOR101-1 Gene.The blast alignment of 3.5kb genomic upstream of OR101-1 gene was done using NCBI Blast tool. The sequence between 1836bp and 2509bp was highly conserved in all zebrafish chromosomes (red).

Thus, genes locating close to this conserved sequence on other chrmosomes were identified. Surprisingly, the conserved sequence could not be found close to any OR gene cluster. However, many of the genes that were located close to occurrences of this conserved sequence were axonal path finding molecules.

The conserved and highly repetitive sequence was further analyzed by using NCBI BLAST and Repeatmasker programmes. It turned out, that the sequence between 1836 and 2509 is composed of 3 different DNA transposons, namely, ENSPM-2N_DR from DNA/CMC-EnSpm family, HE1_DR1from SINE/V family, and DNA-8-9 DR from DNA/hAT-Ac family, all of which are specific for zebrafish and thus might explain the highconservation across the genome.

Additionally, the presence of potential transcription factor binding sites within the 3.5kb upstream genomic region was analyzed. Recently, it was shown in the mouse that a potentially negative regulatory sequence motif is conserved within the first 500bp upstream of OR genes (Michaloski *et al*., 2012). These conserved sequences might be binding sites for members of the Zbtb gene family of transcriptional regulators. Alignment analyses revealed a motif in the mouse with the consensus sequence CNTCTGG, which resembles the motif for CH+BTB-POZ_RP58_M00532 in the TRANSFAC v11.3 database (Michaloski *et al.*, 2012). RP58 protein belongs to the BTB-ZF or POK (POZ and Krüppel) families of transcription factors, containing both Krüppel-like ZF DNA binding domain and BTB domain (Kelly and Daniel, 2006; Costoya, 2007).

The presence of this motif was examined in the 3.5kb upstream genomic region of OR 101-1. Only two occurrences of the CNTCTGG sequence could be found. Surprisingly, these two sites locate to the 500bp inhibitory region. Additionally, one site is located in the SpeI- StuI region, and the other one is located in the StuI- AvrII region (Figure 4.23.). These finding might thus explains why the expression efficiencies of the partial deletions of the inhibitory region are in between of expression efficiencies of-3.5-OR101-1upstr::EYFP-pA and -3.5-OR101-1upstr-Δi::EYFP-pA. RP58 binding domains

may have a combinatorial effect on suppressing OR101-1 gene expression, and maybe one of these sites is not sufficient to the complete regulation of OR101-1 expression.

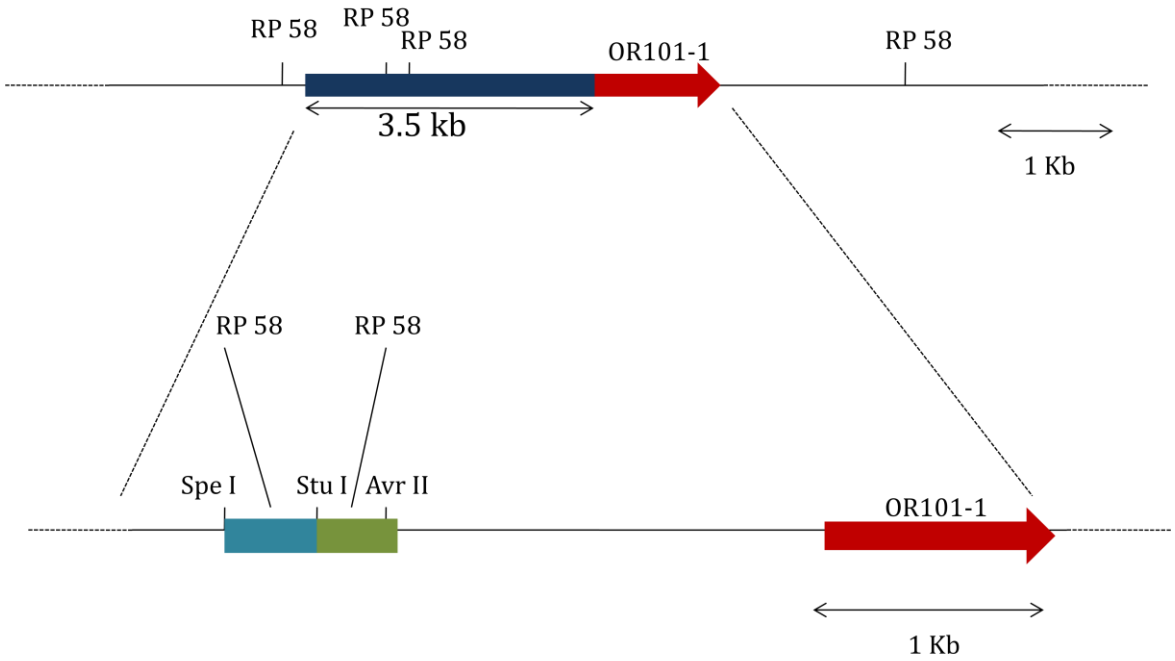


Figure 4.23.The Location of RP58 Binding Sites on the OR101-1 Locus.(Top) The PR58 recognizes CNTCTGG sequence. In the OR101-1 locus, there are only four RP58 binding sites are present in the 5kb upstream and 5kb downstream sequence of the OR101-1 locus, two of which are located in the 500bp inhibitory region.

5. DISCUSSION

It is widely believed that OR gene regulation is a combinatorial mechanism that occurs both at the transcriptional and translational level of the expression process. An important step during the regulation process is the initial selection of an OR gene for expression. It has long been reported that OR genes are expressed in monogenic and monoallelic fashion (Chess *et al.,* 1994; Mombaerts *et al.,* 1996; Serizawa *et al.,* 2000; Ishii *et al.,* 2004).However, there is limited knowledge about key factors that play a role in the transcriptional control of OR gene expression despite the fact that several *cis-acting* regulatory elements have been identified. Among those, long-range as well as short-range control elements have been described. Interestingly, expression of some OR genes depends strictly on long-range control elements (Serizawa *et al.,* 2003), while others are regulated by short proximal promoter elements. For instance, the mouse H-and (Serizawa *et al.,* 2003; Fuss *et al.,* 2007; Nishizumi *et al.,* 2007) P-elements (Bozza *et al.,* 2009) have been shown to act as long range regulators at the level of the MOR28 and P2 OR gene clusters, while the transcription factors LhX2, Emx2, and Olf1/EbF1 (Hirota *et al.,*2004) have been demonstrated to interact with specific binding sites that are located within proximal promoters of OR genes (Vassalli *et al.,* 2002; Rothman *et al.,*2005; Hirota and Mombaerts, 2004; McIntyre *et al.,* 2008).

On the other hand, systematic bioinformatics analysis predicted the functional implication of additional transcription factors, maybe in an OR subset-specific fashion (Hoppe *et al.,* 2006; Michaloski *et al.,* 2012). Of interest are recent findings, which revealed, thatin addition to positive regulators, inhibitory regulatory motifs might also be present in OR gene promoters (Michaloski *et al.,* 2012). This assumption is based on the finding that specific sequence motifs are enriched in sequences upstream of the TSS of a high number of OR genes, and which resemble binding sites of known inhibitory tranbscription factors. One such examples is a CNTCTGG motif, which resembles a binding site for theRP58 transcriptional repressor protein. This motif is present in around49–74% in the mouse V1R promoters, and around 40% in mouse OR promoters (Michaloski *et al.,* 2012).Regulation of OR expression might therefore include a combination of long-range activation and short range promoter function, which involves both, inhibitory and positive regulation by specific transcription factors. The functional

relevance of inhibitory transcriptional regulation on OR gene expression remains unknown. It could, however, be involved in specifying particular OR subsets in a hirarchical fashion, similar to OR gene regulation in the fruit fly, *Drosophila melanogaster*(Ray *et al.,* 2007), or it might be important to restrict expression of certain OR genes to particular cellular subtypes that are competent to express a limited range of ORs (Hoppe *et al.,* 2006; Bozza *et al.,* 2009).

The project presented in this thesis focuses on the regulation of OR101-1 gene expression in zebrafish. The OR101-1 gene appears to be different from most of the other 135 zebrafish ORs. It might be the only representative of a class II OR genes in the zebrafish genome because it shares a high level of protein similarity with human class II ORs and no other OR genes with high sequence similarity to OR101-1 is present in the zebrafish genome.

Despite the fact that OR101-1 appears to be the only member of its class and subfamily, it is part of a larger cluster of 16 OR genes on chromosome 21 and does not constitute an isolated gene in the genome. However, OR101-1 is located at the far end of the cluster and in reverse orientation of the OR115 genes located next to it. This genomic arrangement might facilitate the analysis of regulatory sequences involved in the expression control of OR101-1, as it might not be under the control of long-range cluster-specific regulatory sequences. However, this possibility cannot be fully excluded and it has been demonstrated that OR genes of different OR families can be controlled by the same cluster control region or could otherwise be coordinatedly regulated (Fuss *et al.,* 2007; Nishizumi *et al.,* 2007; Sato *et al.,* 2007). In the work presented in this thesis, short-range regulatory elements have been analyzed in detail, pointing to sequences immediately upstream of the OR101-1 TSS that have the potential to drive transgene expression in a high number of OSNs. However, this still is not a strict prove of the independence of OR101-1 expression from additional long-range elements that might be located within or around the gene cluster on chromosome 21.

However, evidence for an independent regulation of OR101-1 comes from a systematic analysis of the phenotype of a replacement of the OR101-1 coding with fluorescent proteins (Bayramli, unpublished). In these deletion experiments, ONSs that express the mutated OR gene locus that do not produce a functional OR protein often undergo a second OR gene choice (Serizawa *et al.*, 2003; Lewcock and Reed., 2004; Shykind *et al.*, 2004; Sato *et al.*, 2007; Bozza *et al.*, 2009).Second OR gene choice is not random, but typically a large number of OR genes, including unlinked ORs, can be expressed. This phenomenon manifests itself by the rather broad projection of transgenic OSN axons to multiple glomeruli in the OB. Different from what is observed for related experimensts in which class I OR genes have been deleted in zebrafish (Sato *et al.*, 2007), OR101-1 deletion transgene-expressing ORNs target a very limited number of glomeruli in the OB. This observation can be interpreted such that OR101-1-expressing ORNs are restricted to the exclusive expression of OR101-1 and that the projection pattern to the OB represents the fact that deletion transgene expressing cells select the endogenous OR101-1 alleles with high prevalence. Thus, expression of the OR101-1 and OR115 genes might not be controlled by common regulators, as the cells that usually express OR101-1 do not have the competence to express members of the OR115 family as second choice.

This is interesting in the light of recent studies in the mouse, which demonstrated that distinct ORN cell types are committed to the exclusive expression of class I and class II OR genes, respectively (Bozza *et al*., 2009). Interestingly, the difference in class I / class II regulation is defined by the individual OR gene promoter rather than by common cluster-specific regulatory elements.In the mouse genome, the approximately 120 class I OR genes are located within a single large cluster on chromosome 7 and are expressed exclusively in the dorsalmost OE. Similar to the OR101-1 deletion phenotype, OR deletion expressing ORNs in the mouse make a second choice exclusively among ORs of the same class as the OR that has been deleted. However, two atypical class I ORs (Tsuboi *et al.*, 2006) escape the class I OR expression phenotype and regulation, despite being embedded in the middle of the cluster. Interestingly, those two genes differ in their promoter struture and appear to be more similar to class II ORs (Hoppe *et al.*, 2006; Hirota *et al.*, 2007; Bozza *et al.*, 2009). For these reasons, OR101-1 mighty be an ideal candidate to study gene regulation and proximal promoter elements in zebrafish.

Because of the remaining uncertainty, whether OR101-1 is regulated by short-range, long-range, or a combination of short- and long-range regulators, initial transgenic experiments were based on the injection of modified BAC DNA sequences. The advantage of the BAC transgenic approach is that rather large BAC sequences might include even unknown long-range regulatory elements. In the case of OR101-1 BAC transgenes approximately 50kb to 60kb upstream and downstream of the OR101-1 coding sequence were included. This, in theory, should be sufficient sequence to include long-range regulators. By comparison the H-element is located 70kb upstream of the MOR28 cluster in mouse (Serizawa *et al.,* 2003), while the P-element is embedded within the P2 cluster (Khan *et al.,* 2012). The zebrafish genome is much more compact that the mammalian genome and intergenic distances are on average about 10-fold shorter. Two candidate cluster regulators E15.1 and E15.2 have been located within a single large OR cluster on chromosome 15 (Nishizumi *et al.,* 2009).

Two different modified BAC constructs were prepared for injection into fertilized zebrafish oocytes. In the first BAC modification the enhanced yellow fluorescent reporter protein linked to the OR101-1 coding sequence by a T2A peptide were used, to observe expression and axonal targeting of OR101-1-expressing neurons to the OB.

In the secondBAC modification, the OR101-1 coding sequence was replaced by a sequence coding for the yellow fluorescent protein Venus to observe the effect of the OR101-1 deletion on second OR gene choice. In addition, the OR101-1 deletion approach would offer further experimental advantages. It has been shown that the OR coding sequence is directly involved in OR gene regulation (Serizawa *et al.,* 2003; Nguyen *et al.,* 2009). Thus, the OR101-1 deletion experiments would eliminate any additional influence of the presence of the OR101-1 coding sequence on transgene expression.

Unfortunately, in our hands, expression frequency for both the BAC transgenic constructs was rather low and undetectable despite a high number of injected oocytes. This could be due to technical reasons because of low efficiency of transient transgene expression from the rather large BAC DNA that was injected. It could be that the amount of injected DNA was not sufficient to reach levels that are compatible with expression, or that concatamerization of multiple copies of the BAC DNA were subject to silencing by epigenetic events (Serizawa *et al.*, 2003; Lewcock and Reed, 2004; Nguyen *et al.*, 2009). Alternatively, it could also be due to reasons related to OR gene expression control. To eliminate the first possibility, the BAC sequences were further modified to include the Tol2 transposable emelements for efficient integration of BAC DNA into the genome.

However, the twoiTol2 transposase-containing constructs did not give the expected high successof expression either. When co-injected with transposase mRNA, the number of transgene positive fish slightly increased but was still very low. Only 2% of injected embryos expressed the BAC in which the OR101-1 sequence was tagged with yellow fluorescentprotein and 5% expressed the OR101-1 deletion BAC. The decreased efficiency of the BAC in which the OR101-1 coding sequence was tagged with a fluorescence reporter over the deletion BAC might be because of the presence of the OR coding sequence in the transgenic constructs. Similar effects have been observed in the mouse for some transgenic constructs, even when the OR coding sequence was expressed from non-OR promoters (Nguyen *et al.*, 2009).

Thus, overall, the rate of transgene expression that was obtained from the transient BAC transgenic approach was not sufficient for systematic experiments targeting elements that are involved in OR101-1 gene regulation and their statistical analysis.However, the limitations might be overcome bystable integration into the germline of zebrafish, rather than using transient expression. Strong expression of BAC transgenic constructs in a high number of ORNs has been reported for a BAC in which two OR genes have been deleted and replaced with fluorescent reporters when integrated into the germline (Sato *et al.*, 2007). Yet, the establishment of stable transgenic BAC lines is time consuming due to the rather long generation time of the zebrafish of three months (Westerfield, 2007) and not ideal for the identification of small binding sites within the OR101-1 promoter because

many different transgenic modifications have to be tested. Therfore, a short promoter transgenic approach and systematic promoter bashing were performed.

To prepare for the short promoter analysis the 5'-UTR region and cDNA structure of the OR101-1 gene were determined. For this purpose 5'-RLM RACE was performed on cDNA transcribed from 20 zebrafish olfactory epithelia. It was found that the OR101-1 5'-UTR locates 584bp upstream of the coding sequence and consist of a 185bp5'-non-coding exon followed by a 350bp intronic region and 49bpof non-coding sequence preceding the OR101-1 coding sequence in the second axon. The OR101-1 coda structure is consistent with the structure of most reported OR genes, which typically contain one or two non-coding 5'-exons and longer introns. The intron span might be rather compact in zebrafish, when compared to the mouse, as observed from limited trasncriptome and RNAseq data that is available for zebrafish OR genes in public databases. However, this transcriptome sequence information is rather limited for zebrafish OR genes and is slowly becoming available. RACE experiments are inherently prone to error and one can never be sure whether longer or alternative transcripts exist, unless a large number of clones are sequenced. The RT-PCR results for OR101-1 obtained several bands, which were cloned, checked for different insert sizes and sequenced. However, the number of sequenced samples was limited. While conducting the described studies, RNAseq results for the the OR101-1 transcript became available on the EBSEMBL database (www.ensembl.org). Despite the sample being sequenced from bone, the cDNA structure was identical to the result that was obtained by our RLM-RACE experiments, thereby confirming our results independently. Yet, both approaches do not rule out the possibility that longer transcripts with alternative promoters or alternatively transcribed mRNAs might exist. Alternative splicing has been reported for a variety of OR genes (Sosinky *et al.*, 2000; Kan *et al.*, 2001; Young *et al.*, 2003; Volz *et al.*, 2003). These alternative transcripts typically affect the inclusion and length of the various 5'-non coding exons but not the single continuous coding sequence. Yet, this might be relevant for expression and translation control of the OR101-1 gene.

Due to the more compact nature of the zebrafish genome and OR gene clusters, with the length of intergenic sequences between different OR genes being on average 1kb to 5kb, it is reasonable to speculate that regulatory elements in zebrafish OR genes are quite compact as well. Therefore, a screen for regulatory sequences in sequences upstream of the OR101-1 gene was conducted using a promoter bashing approach.

For this purpose, transgenic reporter constructs, containing up to 3.5kb upstream genomic sequence of the OR101-1 fused to a sequence coding for yellow fluorescent protein, were prepared. Initially, the coding sequence was kept intact in the transgenic construct in order to analyze OR101-1 expression and targeting of axons of OR101-1-expressing cells to OB glomeruli. Strikingly, the -3.5-OR101-1upstr::OR101-1::YFP-pA construct did not result in expression in the zebrafish OE, and only ectopic expression in striated muscle cells was observed. The observation that no expression was observed in the OE might be explained by the fact, that endogenous OR genes can block or suppress the expression of their transgenic counterparts (Tsuboi *et al.,* 2006; Nguyen *et al.,* 2007, 2010). On the other hand, ectopic expression in striated muscles of the tail myotome was quite consistent and always observed. To further analyze this effect, a promoterless construct, only comprising the OR101-1 coding sequence followed by T2A-YFP was injected, which also resulted in the same ectopic expression in muscle cells. Thus, somehow the combinations of sequences from the OR101-1 gene and / or the T2A linker promote expression in this cell type. Interestingly, this pattern of expression is never observed for the endogenous OR101-1 gene as judged by in situ hybridization using riboprobes specific for OR101-1.

To overcome the problems with constructs that contain the OR101-1 coding sequence, the sequence was omitted from subsequent transgenic constructs to eliminate the probability of any regulatory influence coming from the OR101-1 coding sequence. First, 4 different constructs with different promoter lengths were engineered. The initial set of constructs contained 3.5kb (-3.5-OR101-1upstr::YFP-pA), 2.5kb (-2.5-OR101-1upstr::YFP-pA), 2.0kb (-2.0-OR101-1upstr::YFP-pA), 1.2kb (-1.2-OR101-1upstr::YFP-pA), upstream genomic sequence fused with EYFP and followed by SV40 polyadenylation signal.

Interestingly, some of these constructs were expressed in a high number of injected embryos and in a high number of ORNs, further supporting the idea that the presence of the OR coding sequence, somehow is implicated in the control of OR gene expression (Nguyen *et al.,* 2009). Yet, the different constructs containing different lengths of upstream genomic sequence were not identical in their efficiency to promote transgene expression. Generally, shorter constructs containing up to 2kb of upstream sequence were expressed at high frequency, while constructs containing longer fragments of upstream sequences had a tendency to be less efficient.

The most striking difference in the efficiency of expression was observed between the -2.5-OR101-1upstr::YFP-pA and -2.0-OR101-1upstr::YFP-pA constructs, which marked the breakpoint in promoter efficiency. This result suggested the presence of an inhibitory sequence located withthe 500bp sequence between 2kb and 2.5kb upstream of the OR101-1 coding sequence. Further analysis of this 500bp sequence revealed that its presence indeed decreases the efficiency of transgene expression.

Removal of the 500bp sequence from the inefficient 3.5kb construct dramatically increased its rate of transgene expression by 4-fold. Similarly, when the 500bp sequence was fused to the highly efficient 1.2kb transgenic construct, the presence of this 500bp region dramatically suppressed the activity of the construct.

However, when the 500bpsequence was divided into two parts to identify the exact location of the contained inhibitory motif, none of the partial deletions reached transgene efficiencies that were as high as observed for a full removal of the sequence. These results were very surprising at first, and suggested the presence of multiple inhibitory motifs in the 500bp inhibitory sequence.

Using web-based sequence analysis tools for the3,5kb genomic upstream sequence it was found that the sequence between 1836bp and 2509bpis highly conserved in the zebrafish genome. This is due to the fact, that the entire 500bp region is derived from transposable elements, the ENSPM-2N_DR transposon, from DNA/CMC-EnSpm family, the HE1_DR1 transposon from the SINE/V family, and the DNA-8-9 DR transposon from the DNA/hAT-Ac family.

Among these three different transposon families, some members of ENSPM transposons have been demonstrated to regulate gene expression. The En-SpmN6_DR transposon contains sequences that can act as p53 response elements which are functionally responsive to Drp53, which can activate a p53-dependent apoptosis pathway (Loviglio *et al.,* 2011). It was shown that certain genes close to the insertion of the En-SpmN6 transposon were rendered responsive to p53 signaling. There is, however, no functional data available about the regulatory activity of the ENSPM-2N_DR transposon, insertion of which upstream of OR 101-1 may have caused the observed transcriptional inhibition in its gene expression.

When the integration sites of the En-SpmN6_DR transposon were analyzed at the genomic level no further integrations close to other OR gene s could be identified. Thus the OR101-1 gene might be the only gene that is under the influence of this particular transposable element. In this light it is interesting, however, that the closest conservation of the transposon sequence found upstream of OR101-1 was observed close to genes coding for axon path finding molecules, such as cadherins and Kirrel2/Kirrel3.

Olfactory sensory neuron projections are highly specific at the level of the OB. Cells expressing the same receptor come together in processing modules called glomeruli. The Kirrel and cadherin molecules were previously demonstrated to guide olfactory sensory axons to OB glomeruli and to be, at least partially, involved in the convergence of axons from OSNs that express the same OR gene onto a specific glomerulus (Serizawa *et al.,* 2006).Thus, if the insertion of the same transposon into other genomic loci close to genes coding for axon path finding molecules has similar effects as on the OR101-1 gene,

OR101-1 expressing ORNs could be endowed with a specific set of factors regulating axon convergence to glomeruli in the OB. Thus, the observed inhibition of transgene expression in the presence of the 5000bp sequence could be beneficial for OR101-1 OSNs in terms of specification of OR expression in particular ORN cell types and it might contribute to the formation of a new OR101-1 cell type if other genes are coordinately controlled as an effect of the insertion of the transposon. However, at this point it remains speculative whether the presence of the transposon close to the OR101-1 transcription start site has true functional significance.

A first indication, however, might come from a preliminary analysis of cell types that express transgenic constructs that did or did not include the 500bp sequence.To better understand the specific function of the repressive sequence, a statistical analysis for each construct except H-2.5-OR101-1upstr::YFP-pAwas performed on a dataset from 10 noses of double positive animals. The hypothesis was that the inhibitory element could reduce the number of transgene-expressing cells in the OE. However, no diference in the number of transgene-expressing OSNs was observed for the different constructs, regardless of the presence of the 500bp sequence. However, there was a bias for transgenic constructs that contained the 500bp sequence to be expressed predominantly in ciliated OSNs, while constructs that did not contain the sequence were also expressed in VR-expressing microvilluos cell types and in cells with round profiles that resembled crypt cells. However, the number of analyzed fish was small and a complication arises in embryonic zebrafish that undergo rapid development where cell types might not unequivocally be identified by morphological criteria alone. Therefore, the analysis should be repeated using defined molecular markers for the different chemosensory cell types in the zebrafish nose (Braubach *et al.*, 2012).

Despite the fact that the inhibitory sequence was of transposon origin, it could contain specific transcription factor binding sites that might be responsible for the observed effects, similar to the effect of a related transposon on p53 responsiveness (Loviglio *et al.*, 2011).

Recently, candidate negative regulatory sites, which are binding sites for RP58, have been described in a bumber of OR and VR promoters (Michaloski *et al.*, 2012). When looking for potential TF binding sites within the 500bp sequence, two RP58 binding sites could be observed. RP58 proteins are belong to the BTB-ZF family (Aoki *et al.*, 1998), and act as a transcriptional repressors by inducing epigenetic modifications. Moreover, specific members of the Zbtb gene family were found in mouse studies, to be strongly expressed in the OE. Their binding domains are conserved in the promoter sequences of both V1R and OR genes (Michaloski *et al.*, 2012) and zbtb7 was shown to physically interact with the conserved motif and to be expressed in the OE.

Interestingly, the two RP58 binding motives that can be identified within the 500bp sequence locate within each of the two halves of the inhibitory sequence. This might explain why only intermediary results were observed when only the 5'- or 3'- half of the sequence was removed from transgenic constructs. If these two motives have a combinatorial action, then both motifs should be removed to obtain the full effect. The effect of these two motives should be analyzed more, by designing a site-directed mutagenesis experiments on these CNTCTGG sequences while leaving the remaining sequence intact.

In the zebrafish genome 25 different zbtb genes can be identified, raising the question, which one or which combination might be responsible for the observed suppression of gene expression. However, at least one of them, Zbtb10, is strongly expressed in the OEand in the OB(Gompel *et al.*, 2001). Thus, the two RP58 binding sites, in combination with Zbtb10, may repress the OR101-1 gene expression. To confirm this speculation, the zbtb binding motifs should be mutated in long transgenic constructs. Alternatively, the function of zbtb genes expressed in the zebrafish OE could be tested by morpholino-mediated knock-down of gene function. If the two observed RP58 binding sites have functional significance for OR101-1 expression DNA-protein interaction could also be analyzed by yeast-1-hybrid screening.

It is intersting to note that no similar motifs have been observed for other zebrafish or mouse OR genes. Thus, the results presented in this thesis might be the first description of a negative regulation of OR gene expression in vertebrates. In Drosophila specific factors have been demonstrated that restrict expression to one of the two olfactory organs, the maxillary palp or thge antenna (Fuss and Ray, 2009). Yet, in vertebrates, the limited number of transcription factors that have been tested functionally promote OR expression. Removal of the homeodomain or O/E binding sites from the M71 promoter resulted in reduced expression of the transgene (Rothmann *et al*., 2005). Similarly, removal of the factors that are supposed to bind to these sites, such as Lhx2 (Hirota *et al.,* 2007), Emx2 (McIntyre *et al.,* 2008), or Olf-2/3 (Wang *et al.,* 2004) result in distubances of OR gene expression, typicaly the specific loss of expression of OR subsets.

An intersting hypothesis is that zbtb-mediated repression might have functional significance for class II OR expression. As outlined earlier, OR101-1 is the only member of the class II ORs in zebrafish and the only gene for which RP58 sites were identified so far. It should be analyzed, whether the observed zbtb-binding motifs in the mouse are also exclusively found in the promoters of class II OR genes or if they are also present around class I ORs. Thus, the zbtb-mediated repression might function as a class II OR specific regulation.

Identification of regulatory motifs responsible for OR gene expression sheds some light on the underlying mechanisms of OR gene choice, which might be a combination of activation and inhibition to restrict OR gene expression to appropriate ORN subsets. Currently, our understandding of the mechanisms that activate OR gene expression in specific spatial patterns in the OE is very limited. Thus, the identification of specific sequences that are presented in this thesis might provide new insight into the complex mechanism of OR gene regulation. The analysis presented here focuses exlcusively on sequences located 5’- and upstream of the transcription start site. It is possible, however, that further elelments are located in 3’-downstream regions of the OR genes or within introns. Experiments, similar to the ones presented here, could be employed to further analyze the role of these sequences for OR gene expression.

APPENDIX A: EQUIPMENTS

Table 6.1.Equipments.

4 ^{o}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:	Bio-Rad Labs, USA (ReadySub-Cell GT Cells)
Electroporation system:	Bio-Rad Labs, USA (Gene Pulser Xcell)
Fluorescence Microscope:	Leica Microsystems, USA (MZ16FA)
Freezer:	-20^{o}C Arçelik, Turkey -80^{o}C Thermo Electron Corp., USA (Thermo Forma723)
Gel documentation:	Bio-Rad Labs, USA (GelDoc XR)
Incubator:	Weiss Gallenkamp, UK Nuve, Turkey
Incubating shaker:	Thermo Electron Corp, USA (Forma Orbital Shaker)
Laboratory glass bottles:	Isolab, Germany
Luminometer:	Fluoroskan Ascent FL luminometer (Thermo Scientific)
Micropipetters:	Gilson, USA (Pipetman) Eppendorf, Germany (Research)
Microwave oven:	Vestel, Turkey
Microinjector:	Eppendorf, Germany (FemtoJet)
Refrigerator:	Arcelik, Turkey
Softwares:	Invitrogen, USA (Vector NTI)
Thermal cyclers (PCR):	Bio-Rad Labs, USA (C1000)
Vortex:	Scientific Industries, USA

APPENDIX B: SUPPLIES

Table 6.2.Supplies.

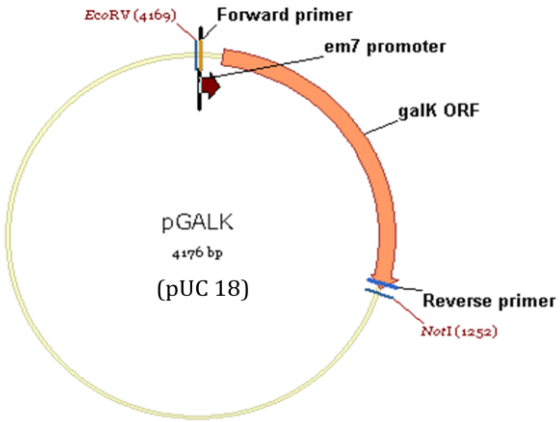
Disposable Labware	
14 ml Culture Tubes:	Greiner Bio-One,Belgium (187261)
CELLSTAR® Centrifuge Tubes,15 ml:	Greiner Bio-One,Belgium (186161)
CELLSTAR® Centrifuge Tubes,50 ml:	Greiner Bio-One,Belgium (227261)
Filtered Pipette Tips:	Greiner Bio-One,Belgium (771288, 772288, 740288)
Micro-centrifuge tubes:	Citotest, China (34730015)
PCR Tubes:	Bio-Rad, U.S.A. (TBS0201)
Chemical Supplies	
1kb DNA Ladder:	New England Biolabs, U.S.A. (N3232)
100bp DNA Ladder:	New England Biolabs, U.S.A. (N3231L)
Advantage® 2 Polymerase Mix	Clontech, U.S.A. (639201)
5X GoTaq® Flexi Buffer:	Promega, U.S.A (M890A).
Ampicillin sodium salt:	Sigma-Aldrich, U.S.A. (A9518-25G)
AseI:	New England Biolabs, U.S.A (R0526M).
AvrII:	New England Biolabs, U.S.A (R0174L)
BamHI:	Promega, U.S.A (R602A).
Bovine Serum Albumin:	Sigma-Aldrich, U.S.A. (A-9647)
Chloramphenicol:	Roche (85369821)
2-Deoxi-D-Galactose:	Sigma-Aldrich, U.S.A. (D4407-5G-A)
Diethylpyrocarbonate	Sigma-Aldrich, U.S.A. (D5758-100ML)
EcoRV:	Promega, U.S.A (R635A).
EcoRI:	New England Biolabs, U.S.A (R0101 S).
Ethanol Absolute:	Sigma-Aldrich, U.S.A. (34870)
Ethidium Bromide 10 mg/ml:	Sigma Life Sciences, U.S.A. (E1510-1M L).
Ethylenediaminetetraacetic acid (EDTA) disodium salt:	Sigma-Aldrich, U.S.A. (E5134 - 1KG)
Ethyl 3-aminobenzoate methanesulfonate salt:	Sigma-Aldrich, U.S.A. (A5040-25G)
Glycerol, for molecular biology:	Sigma-Aldrich, U.S.A. (G5516-500M L).
GoTaq® Flexi DNA Polymerase:	Promega, U.S.A (M830B).
LB Agar:	Sigma Life Sciences, U.S.A. (SL08394).
LB Broth:	Sigma-Aldrich, U.S.A. (L7658-1 KG).
Low Gel Temperature Agarose:	Mallinckrodt Chemicals (B08643)
Magnesium Chloride,25mM:	Promega, U.S.A (A3511

Table 6.2.Supplies. (Continued)

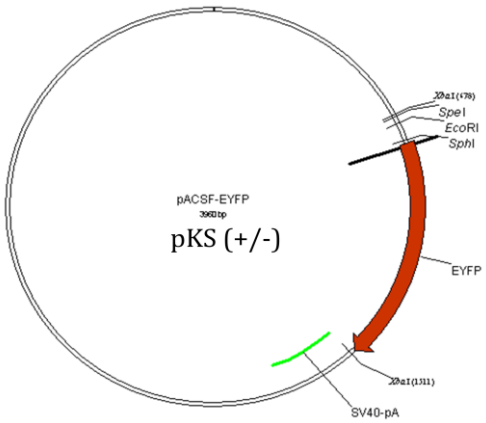
Magnesium Sulfate:	Sigma-Aldrich, U.S.A. (M7506)
McConkey Agar Base:	Becton, Dickinson and Company, U.S.A (0228424)
Nco I:	Promega, U.S.A (R6515).
Not I:	Promega, U.S.A (R6435).
Nsi I:	Promega, U.S.A (R653A)
pGEM®-T Easy Vector System:	Promega, U.S.A (A1360).
Phenol Red:	Sigma-Aldrich, U.S.A. (P3532-5G)
Phenol : Chloroform : Isoamyl alcohol:	Sigma-Aldrich, U.S.A. (P2069)
Potassium Chloride:	Sigma-Aldrich, U.S.A. (P9541)
Potassium Hydroxide:	Atabay Kimya Sanayii LTD. ŞTI
Pst I:	New England Biolabs, U.S.A (R140 L)
Sal I:	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 - 1KG)
Sodium hydroxide:	Sigma-Aldrich, U.S.A. (S8045 - 1KG)
Spe I:	New England Biolabs, U.S.A (R0133)
Sph I:	New England Biolabs, U.S.A (R0182)
Stu I:	New England Biolabs, U.S.A (R0187S)
T4 DNA Ligase:	New England Biolabs, U.S.A (M0202L) Promega, U.S.A (M1804)
Trizma® Base:	Sigma-Aldrich, U.S.A. (T6066)
TRIzol™	Invitrogen, U.S.A. (15596 - 026)
Triton X-100	Roche, (12754421)
Xho I:	New England Biolabs, U.S.A (R0146).
Commercial Kits	
FirstChoice® RLM-RACE Kit	Ambion, U.S.A (AM1700)
QIAprep® Spin Miniprep Kit (250):	Qiagen®, U.S.A. (27106).
High Pure PCR Purification Kit	Roche, Germany (11732676001)

APPENDIX C: VECTOR MAPS

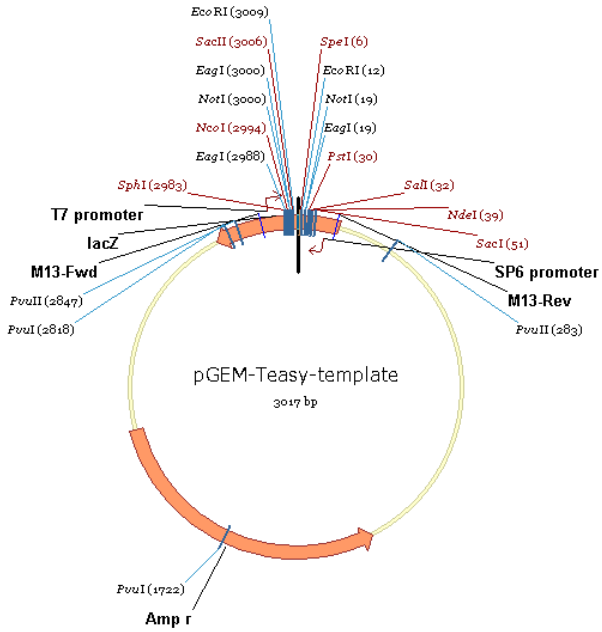
(a)



(b)



(c)



(d)

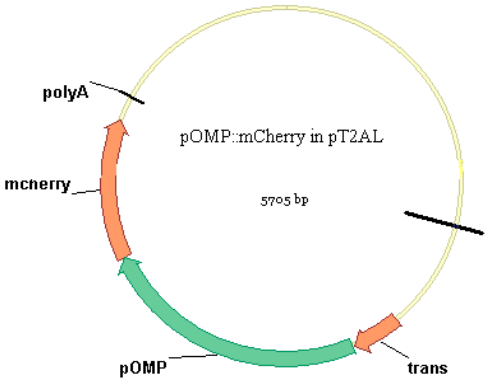


Figure 6.1.Vector maps.

(a) Vector map of pgalK. (b) Vector map of pACSF-EYFP. (c) Vector map of pGEM®-T Easy. (d) Vector map of pOMP::mCherry in pT2AL.

REFERENCES

Alioto, T. S. and J. Ngai, 2005. "The Odorant Receptor Repertoire of Teleost Fish." *BioMed Central Genomics,* Vol.6,No.173.

Andres, G. and E. Roessler, 1970. "Transfer of the Species-Specific Behavior Componets by Xenoplastic Transplantation of the Brain Rudiments between Xenopus Laevis And Hymenochirus Boettgeri (Amphibia, Anura)." *Revue Suisse de Zoologie,* Vol.77, No.4, pp. 959-62.

Aoki, K., G. Meng, K. Suzuki, T. Takashi, Y. Kameoka, K. Nakahara, R. Ishida, and M. Kasai,1998. "RP58 Associates with Condensed Chromatin and Mediates a Sequence-Specific Transcriptional Repression." *Journal of Biological Chemistry,* Vol.273, No.41, pp.26698-704.

Argo, S., F. Weth, and S.I. Korsching,2003. "Analysis of Penetrance and Expressivity during Ontogenesis Supports a Stochastic Choice of Zebrafish Odorant Receptors from Predetermined Groups of Receptor Genes." *Europian Journal of Neuroscience,* Vol.17, No.4, pp.833-43.

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 Reserch Communications,* Vol. 221, No.2, pp.240-7.

Baier, H. and S. Korsching, 1994. "Olfactory Glomeruli in the Zebrafish Form an Invariant Pattern and are Identifiable across Animals." *Jornal of Neuroscience,* Vol. 14, No.1, pp. 219-30.

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. 220-33.

Braubach, O. R., A. Fine, and R.P. Croll, 2012. "Distribution and Functional Organization of Glomeruli in the Olfactory Bulbs of Zebrafish (Danio Rerio)." *The Journal of Comperative Neurology,* Vol.520, No.11, pp. 2317-39.

Breer, H., J. Fleischer, and J. Stromann,2006. "The Sense of Smell: Multiple Olfactory Subsystems." *Cellular and Molecular Life Sciences,* Vol.63, No.13, pp.1465-75.

Brennan, P. A. and E. B. Keverne,1997. "Neural Mechanisms of Mammalian Olfactory Learning." *Progress in Neurobiology,* Vol.51, No.4, pp.457-81.

Brennan, P. A. and F. Zufall,2006. "Pheromonal Communication in Vertebrates." *Nature,* Vol.444, No.7117, pp.308-15.

Britten, R. J. and E. H Davidson, 1969. "Gene Regulation for Higher Cells: a Theory." *Science,*Vol.165, No.3891, pp.349-57.

Buck, L. and R. Axel,1991. "A Novel Multigene Family may Encode Odorant Receptors: A Molecular Basis for Odor Recognition." *The Cell,* Vol.65, No.1, pp.175-87.

Buck, L. B., 2000. "The Molecular Architecture of Odor and Pheromone Sensing in Mammals." *The Cell,* Vol.100, No.6, pp.611-8.

Bulger, M., M. A. Bender,J.H van Doorninck, B. Wertman, C.M. Farrell, G. Felsenfeld, M. Groudine, and R. Hardison,2000. "Comparative Structural and Functional Analysis of the Olfactory Receptor Genes Flanking the Human and Mouse Beta-Globin Gene Clusters." *Proceedings of the National Academy of Sciences,U S A,* Vol.97, No.26, pp.14560-5.

Chandler, K. J., R. L. Chandler, E.M. Broeckelmann, Y. Hou, E.M Southard-Smith, and D.P. Mortlock, 2007."Relevance of BAC Transgene Copy Number in Mice: Transgene Copy Number Variation across Multiple Transgenic Lines and Correlations with Transgene Integrity and Expression." *Mammalian Genome,* Vol.18, No.10, pp.693-708.

Chess, A., L. Buck, M.M Dowling, R. Axel, and J. Ngai, 1992. "Molecular Biology of Smell: Expression of the Multigene Family Encoding Putative Odorant Receptors." *Cold Spring Harbor Symposia on Quantitative Biology*, No.57, pp.505-16.

Chess, A., I. Simon, H. Cedar, and R. Axel, 1994. "Allelic Inactivation Regulates Olfactory Receptor Gene Expression."*The Cell*, Vol.78, No.5, pp.823-34.

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.1249-59.

Costoya, J. A.,2007. "Functional Analysis of the Role of POK Transcriptional Repressors." *Briefinngs in Functional Genomics and Proteomics,* Vol.6, No.1, pp.8-18.

Czerny, T., G. Halder, U. Kloter, A. Souabni, W.J. Gehring, and M. Busslinger, 1999. "Twin of Eyeless, a Second Pax-6 Gene of Drosophila, Acts Upstream of Eyeless in the Control of Eye Development." *Molecular Cell,*Vol.3, No.3, pp.297-307.

de Felipe, P. and M. D. Ryan, 2004. "Targeting of Proteins Derived from Self-Processing Polyproteins Containing Multiple Signal Sequences." *Traffic,* Vol.5, No.8, pp.616-26.

Demir, R., L. B. Haberly, and M.B. Jackson,1998. "Voltage Imaging of Epileptiform Activity in Slices from Rat Piriform Cortex: Onset and Propagation." *Journal of Neurophysiology,* Vol.80, No.5, pp.2727-42.

DiBello, P. R., D. A. Withers,C.A. Bayer, J.W. Fristrom, and G.M. Guild,1991. "The Drosophila Broad-Complex Encodes a Family of Related Proteins Containing Zinc Fingers." *Genetics,* Vol.129,No.2, pp.385-97.

Donnelly, M. L., L. E. Hughes, G. Luke, H. Mendoza, E. ten Dam, D. Gani, and M.D. Ryan,2001. "The 'Cleavage' Activities of Foot-And-Mouth Disease Virus 2A Site-Directed Mutants and Naturally Occurring '2A-Like' Sequences." *Journal of General Virology,*Vol.82, No. 5, pp.1027-41.

Duboule, D,1998. "Vertebrate Hox Gene Regulation: Clustering and/or Colinearity?" *Current Opinions of Genetics and Development,* Vol.8, No.5, pp.514-8.

Dugas, J. C. and J. Ngai,2001. "Analysis and Characterization of an Odorant Receptor Gene Cluster in the Zebrafish Genome." *Genomics,*Vol.71, No.1, pp.53-65.

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

Fuchs, T., G. Glusman, S. Horn-Saban, D. Lancet, and Y. Pilpel, "The Human Olfactory Subgenome: From Sequence to Structure and Evolution." *Human Genetics*Vol.108, No.1, pp.1-13.

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." *The Cell,* Vol.130, No.2, pp.373-84.

Fuss, S. H. and A. Ray,2009. "Mechanisms of Odorant Receptor Gene Choice in Drosophila and Vertebrates." *Molecular and Cellular Neuroscience,* Vol.41, No.2, pp.101-12.

Garrick, D., S. Fiering, D.I.Martin, and E. Whitelaw,1998. "Repeat-induced Gene Silencing in Mammals." *Nature Genetics,* Vol.18, No.1, pp.56-9.

Gehring, W. J.,2002. "The Genetic Control of Eye Development and Its Implications for the Evolution of the Various Eye-Types." *International Journal of Developmental Biology,*Vol.46, No.1, pp.65-73.

Godfrey, P. A., B. Malnic, and L.B. Buck,2004. "The Mouse Olfactory Receptor Gene Family." *Proceedings of the National Academy of Sciences, U S A,* Vol.101, No.7, pp.2156-61.

Godt, D., J. L. Couderc, S.E. Cramton, and F.A. Laski,1993. "Pattern Formation in the Limbs of Drosophila: Bric a Brac is Expressed in Both a Gradient and a Wave-Like Pattern and is Required for Specification and Proper Segmentation of the Tarsus." *Development,* Vol.119, No.3, pp.799-812.

Gompel, N., N. Cubedo, C. Thisse, B. Thisse, C. Dambly-Chaudière, and A. Ghysen,2001. "Pattern Formation in the Lateral Line of Zebrafish." *Mechanisms of Development,* Vol.105, No.1-2,pp.69-77.

Gossen, M., S. Freundlieb, G. Bender, G. Müller, W. Hillen, and H. Bujard,1995. "Transcriptional Activation by Tetracyclines in Mammalian Cells." *Science,*Vol.268, No.5218, pp.1766-9.

Hamdani el, H. and K. B. Doving,2007. "The Functional Organization of the Fish Olfactory System." *Progress in Neurobiology,* Vol.82, No.2, pp.80-6.

Hansen, A. and T. E. Finger,2000. "Phyletic Distribution of Crypt-Type Olfactory Receptor Neurons in Fishes." *Brain, Behaviour and Evolution,* Vol.55, No.2, pp.100-10.

Hansen, A., S. H. Rolen, K. Anderson, Y. Morita, J. Caprio, and T.E Finger,2003. "Correlation between Olfactory Receptor Cell Type and Function in the Channel Catfish." *Journal of Neuroscience, 2003.*Vol.23, No.28, pp.9328-39.

Hansen, A. and E. Zeiske,1998. "The Peripheral Olfactory Organ of the Zebrafish, Danio Rerio: An Ultrastructural Study." *Chemical Senses,*Vo.23, No.1, pp.39-48.

Harrison, S. D. and A. A. Travers,1990. "The Tramtrack Gene Encodes aDrosophila Finger Protein that Interacts with the Ftz Transcriptional Regulatory Region and Shows a Novel Embryonic Expression Pattern." *EMBO Journal,*Vol.9, No.1, pp.207-16.

Hauck, B., W. J. Gehring, and U. Walldorf,1999. "Functional Analysis of an Eye Specific Enhancer of the Eyeless Gene in Drosophila." *Proceedings of the National Academy of Sciences U S A,* Vol.96, No.2, pp.564-9.

Heras, S. R., M. C. Thomas,M. García-Canadas, P. de Felipe, J.L. García-Pérez, M.D. Ryan, and M.C. López,2006. "L1Tc Non-LTR Retrotransposons from Trypanosoma Cruzi Contain a Functional Viral-Like Self-Cleaving 2A Sequence in Frame with the Active Proteins They Encode." *Cellular and Molecular Life Sciences,*Vol.63, No.12, pp.1449-60.

Herrada, G. and C. Dulac, 1997. "A Novel Family of Putative Pheromone Receptors in Mammals With a Topographically Organized and Sexually Dimorphic Distribution." *The Cell,* Vol.90, No.4, pp.763-73.

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 U S A,*Vol.101, No.23, pp.8751-5.

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.679-88.

Hoppe, R., H. Breer, and J. Stromann,2006. "Promoter Motifs of Olfactory Receptor Genes Expressed in Distinct Topographic Patterns." *Genomics,*Vol.87, No.6, pp.711-23.

Hoppe, R., H. Frank,H. Breer, and J. Stromann,2003. "The Clustered Olfactory Receptor Gene Family 262: Genomic Organization, Promoter Elements, and Interacting Transcription Factors." *Genome Research,* Vol.13, No.12, pp.2674-85.

Hoppe, R., M. Weimer, A. Beck, H. Breer, and J. Stromann,2000. "Sequence Analyses of the Olfactory Receptor Gene Cluster Mor37 on Mouse Chromosome 4." *Genomics,*Vol.66, No.3, pp.284-95.

Ishii, T., M. Omura, and P. Mombaerts,2004. "Protocols for Two- and Three-Color Fluorescent RNA İn Situ Hybridization of the Main and Accessory Olfactory Epithelia in Mouse." *Journal of Neurocytology,* Vol.33, No.6, pp.657-69.

Kan, Z., E.C. Rouchka, W.R. Gish, and D.J. States, 2001. "Gene Structure Prediction and Alternative Splicing Analysis Using Genomically Aligned Ests." *Genome Research,* Vol.11, pp.889–900.

Kawakami, K. and A. Shima,1999. "Identification of the Tol2 Transposase of the Medaka Fish Oryzias Latipes that Catalyzes Excision of a Nonautonomous Tol2 Element in Zebrafish Danio Rerio." *Gene,*Vol.240, No.1, pp.239-44.

Kawakami, K., H. Takeda, N. Kawakami, M. Kobayashi, N. Matsuda, and M. Mishina, 2004. "A Transposon-Mediated Gene Trap Approach Identifies Developmentally Regulated Genes in Zebrafish." *Development Cell,*Vol.7, No.1, pp.133-44.

Kelly, K. F. and J. M. Daniel,2006. "POZ for Effect--POZ-ZF Transcription Factors in Cancer and Development." *Trends in Cell Biology,* Vol.16, No.11, pp.578-87.

Khan, M., E. Vaes, and P. Mombaerts,2011. "Regulation of the Probability of Mouse Odorant Receptor Gene Choice."*The Cell,*Vol.147, No.4, pp.907-21.

Kolterud, A., M. Alenius, L. Carlsson, and S. Bohm,2004. "The Lim Homeobox Gene Lhx2 is Required for Olfactory Sensory Neuron Identity." *Development,* Vol.131, No.21, pp.5319-26.

Krieger, J. and H. Breer,1999. "Olfactory Reception in Invertebrates." *Science,*Vol.286, No.5440, pp.720-3.

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.465-75.

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.3002-14.

Kwan, K. M., E. Fujimoto,C. Grabher, B.D. Mangum, M.E. Hardy, D.S. Campbell, J.M. Parant, H.J. Yost, J.P. Kanki, and C.B. Chien,2007. "The Tol2kit: A Multisite Gateway-Based Construction Kit for Tol2 Transposon Transgenesis Constructs." *Developmental Dynamics,* Vol.236, No.11, pp.3088-99.

Lane, R. P., T. Cutforth, J. Young, M. Athanasiou, C. Friedman, L. Rowen, G. Evans, R. Axel, L. Hood, and B.J. Trask,2001. "Genomic Analysis of Orthologous Mouse and Human Olfactory Receptor Loci." *Proceedings of the National Academy of Sciences U S A,* Vol.98, No.13, pp.7390-5.

Lewcock, J. W. and R. R. Reed,2004. "A Feedback Mechanism Regulates Monoallelic Odorant Receptor Expression." *Proceedings of the National Academy of Sciences U S A,*Vol.101, No.4, pp.1069-74.

Li, J., T. Ishii, P. Feinstein, and P. Mombaerts,2004. "Odorant Receptor Gene Choice is Reset By Nuclear Transfer from Mouse Olfactory Sensory Neurons." *Nature,* Vol.428, No.6981, pp.393-9.

Lotto, R. B. and D. J. Price,1994. "Evidence that Molecules Influencing Axonal Growth And Termination in the Developing Geniculocortical Pathway are Conserved between Divergent Mammalian Species." *Brain Research Developmental Brain Research,*Vol.81, No.1, pp.17-25.

Loviglio, M., L. Micale, C. Fusco, A. Calcagnì, B. Augello, G. Cotugno, E. V. D'Addetta, and G. Merla,"Enspm-N6_DR DNA Transposons Shape the Repertoire of P53 Target Genes in Zebrafish." American Society of Human Genetics/ICHG 2011 Meeting.

Ma, M.,2007. "Encoding Olfactory Signals via Multiple Chemosensory Systems."*Critical Reviews in Biochemistry and Molecular Biology,* Vol.42, No.6, pp.463-80.

Magklara, A., A. Yen, B.M. Colquitt, E.J. Clowney, W. Allen, E. Markenscoff-Papadimitriou, and Z.A. Evans, 2011. "An Epigenetic Signature for Monoallelic Olfactory Receptor Expression." *The Cell,* Vol.145, No.4, pp.555-70.

Malnic, B., J. Hirono,T. Sato, and L.B. Buck,1999. "Combinatorial Receptor Codes for Odors." *The Cell,* Vol.96, No.5, pp.713-23.

Matsunami, H. and L. B. Buck,1997. "A Multigene Family Encoding a Diverse Array of Putative Pheromone Receptors in Mammals." *The Cell,* Vol.90, No.4, pp.775-84.

McIntyre, M. L. and C.D Jones,2008. " Evolution of Gene Expression in the Drosophila Olfactory System. " *Molecular Biology and Evolution,* Vol.25, No.6, pp.1081–1092.

McClintock, T. S.,2010. "Achieving Singularity in Mammalian Odorant Receptor Gene Choice." *Chemical Senseses,* Vol.35, No.6, pp.447-57.

Meredith, M.,1991. "Sensory Processing in the Main and Accessory Olfactory Systems: Comparisons and Contrasts." *The Journal of Steroid Biochemistry and Molecular Biology,* Vol.39, No.4B, pp.601-14.

Michaloski, J. S., P. A. 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.1091-8.

Michaloski, J. S., P. A. Galante, M.H. Nagai, L. Armelin-Correa, M.S. Chien, H. Matsunami, and B. Malnic,2012. "Common Promoter Elements in Odorant and Vomeronasal Receptor Genes." *PLoS One,*Vol.6, No.12, pp.e29065.

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." *Journal of Neuroscience,* Vol.25, No.14, pp.3586-92.

Mombaerts, P.,1999. "Molecular Biology of Odorant Receptors in Vertebrates." *Annual Review of Neuroscience,* Vol.22, pp.487-509.

Mombaerts, P.,1999. "Seven-Transmembrane Proteins as Odorant and Chemosensory Receptors." *Science,*Vol.286, No.5440, pp.707-11.

Mombaerts, P., F. Wang, C. Dulac, S.K. Chao, A. Nemes, M. Mendelsohn, J. Edmondson, and R. Axel,1996. "Visualizing an Olfactory Sensory Map." *The Cell,*Vol.87, No.4, pp.675-86.

Mori, K., H. Nagao, and Y. Yoshihara,1999. "The Olfactory Bulb: Coding and Processing of Odor Molecule Information." *Science,*Vol.286, No.5440, pp.711-5.

Mori, K., H. von Campenhause,and Y. Yoshihara,2000. "Zonal Organization of the Mammalian Main and Accessory Olfactory Systems." *Philosophical Transactions of the Royal Society B: Biological Sciences,*Vol.355, No.1404, pp.1801-12.

Mori, T., K. Kurihara, I. Matsuoka, and K. Kurihara, 2000. " Analysis of Promoter Activity of 5'-Upstream Regions of Zebrafish Olfactory Receptor Genes." *Bio Pharmaceutical Bullettin,*Vol.23, No.2, pp.165-73.

Nei, M.,1969. "Gene Duplication and Nucleotide Substitution in Evolution." *Nature,*Vol.221, No.5175, pp.40-2.

Nei, M., X. Gu, and T. Sitnikova,1997. "Evolution by The Birth-and-Death Process in Multigene Families of the Vertebrate Immune System." *Proceedings of the National Academy of Sciences USA,* Vol.94, No.15, pp.7799-806.

Nemazee, D., A. Martensson, and L. Verkoczy, 2002. "Haplotype Exclusion and Receptor Editing: Irreconcilable Differences?"*Seminar Immunology,*Vol.14, No.3, pp.191-8; discussion 222-4.

Nguyen, M. Q., C. A. Marks,L. Belluscio, and N.J. Ryba, 2010. "Early Expression of Odorant Receptors Distorts the Olfactory Circuitry." *Journal of Neuroscience,* Vol.30, no.27, pp.9271-9.

Nguyen, M. Q., Z. Zhou, C.A. Marks , N.J. Ryba, and L. Belluscio, 2007. "Prominent Roles for Odorant Receptor Coding Sequences in Allelic Exclusion." *The Cell,*Vol.131, No.5, pp.1009-17.

Niimura, Y. and M. Nei,2005. "Evolutionary Dynamics of Olfactory Receptor Genes in Fishes and Tetrapods." *Proceedings of the National Academy of SciencesU S A,* Vol.102, No.17, pp..6039-44.

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 SciencesU S A,* Vol.104, No.50, pp.20067-72.

Oikawa, T., K. Suzuki, T.R. Saito, K.W. Takahashi, and K. Taniguchi,1998. "Fine Structure of Three Types of Olfactory Organs in Xenopus Laevis." *The Anatomical Record,*Vol.252, No.2, pp.301-10.

Oka, Y., K. Kobayakawa, H. Nishizumi, K. Miyamichi, S. Hirose, A. Tsuboi, and H. Sakano,2003. "O-MACS, A Novel Member of the Medium-Chain Acyl-Coa Synthetase Family, Specifically Expressed in the Olfactory Epithelium in a Zone-Specific Manner." *European Journal of Biochemistry,*Vol.270, No.9, pp.1995-2004.

Osborn, M. J., A. Panoskaltsis-Mortari, R.T. McElmurry, S.K. Bell, D.A. Vignali, M.D. Ryan, A.C. Wilber, R.S. McIvor, J. Tolar, and B.R. Blazar,2005. "A Picornaviral 2A-Like Sequence-Based Tricistronic Vector Allowing for High-Level Therapeutic Gene Expression Coupled to a Dual-Reporter System." *Molecular Therapy,*Vol.12, No.3, pp.569-74.

Papatsenko, D., A. Nazina, and C. Desplan,2001. "A Conserved Regulatory Element Present in All Drosophila Rhodopsin Genes Mediates Pax6 Functions and Participates in the Fine-Tuning of Cell-Specific Expression." *Mechanisms of Development,*Vol.101, No.1-2, pp.143-53.

Pifferi, S., M. Dibattista, C. Sagheddu, A. Boccaccio, A. Al Qteishat, F. Ghirardi, R. Tirindelli, and A. Menini, 2009. "Calcium-Activated Chloride Currents in Olfactory Sensory Neurons from Mice Lacking Bestrophin-2." *Journal of Physiology,*Vol.587, No.Pt 17, pp.4265-79.

Qasba, P. and R. R. Reed,1998. "Tissue and Zonal-Specific Expression of an Olfactory Receptor Transgene." *Journal of Neuroscience,*Vol.18, No.1, pp.227-36.

Rawson, N. E., J. Eberwine, R. Dotson, J. Jackson, P. Ulrich, and D. Restrepo, 2000. "Expression of mRNAs Encoding for Two Different Olfactory Receptors in a Subset of Olfactory Receptor Neurons." *Journal of Neurochemistry,* Vol.75, No.1, pp.185-95.

Ray, A., W.G. van Naters, T. Shiraiwa, and J.R. Carlson,2007. "Mechanisms of Odor Receptor Gene Choice in Drosophila. " *Neuron,*Vol.53, No.3, pp.353–369.

Reed, R. R.,2000. "Regulating Olfactory Receptor Expression: Controlling Globally, Acting Locally." *Nature Neuroscience,*Vol.3, No.7, pp.638-9.

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

Ressler, K. J., S. L. Sullivan, and L.B. Buck,1994. "A Molecular Dissection of Spatial Patterning in the Olfactory System." *Current Opinions in Neurobiology,*Vol.4,No.4,pp.588-96.

Riviere, S., L. Challet, D. Fluegge, M. Spehr, and I. Rodriguez,2009. "Formyl Peptide Receptor-Like Proteins are a Novel Family of Vomeronasal Chemosensors." *Nature,*Vol.459, No.7246, pp.574-7.

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.535-46.

Ryba, N. J. and R. Tirindelli, 1997. "A New Multigene Family of Putative Pheromone Receptors." *Neuron,*Vol.19,No.2, pp.371-9.

Sambrook, L. and D. Russel.,2001. Molecular Cloning: A Laboratory Manual. New York, USA, Cold Spring Harbor Laboratory.

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." *Journal of Neuroscience,*Vol.27, No.7, pp.1606-15.

Serizawa, S., T. Ishii, H. Nakatani, A. Tsuboi, F. Nagawa, M. Asano, K. Sudo, J. Sakagami, H. Sakano, T. Ijiri, Y. Matsuda, M. Suzuki, T. Yamamori, Y. Iwakura, and H. Sakano,2000. "Mutually Exclusive Expression of Odorant Receptor Transgenes." *Natural Neuroscience,*Vol.3, No.7, pp.687-93.

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,*Vol.302, No.5653, pp.2088-94.

Serizawa, S., K. Miyamichi, and H. Sakano,2004. "One Neuron-One Receptor Rule in the Mouse Olfactory System." *Trends in Genetics,* Vol.20, No.12, pp.648-53.

Serizawa, S., K. Miyamichi, H. Takeuchi, Y. Yamagishi, M. Suzuki, and H. Sakano,2006. "A Neuronal Identity Code for the Odorant Receptor-Specific and Activity-Dependent Axon Sorting." *The Cell,* Vol.127, No.5, pp.1057-69.

Sheng, G., E. Thouvenot, D. Schmucker, D.S. Wilson, and C. Desplan,1997. "Direct Regulation of Rhodopsin 1 By Pax-6/Eyeless in Drosophila: Evidence for a Conserved Function in Photoreceptors." *Genes and Development,*Vol.11, No.9, pp.1122-31.

Shykind, B. M.,2005. "Regulation of Odorant Receptors: One Allele at a Time." *Human Molecular Genetics,*14 Spec No.1, pp.R33-9.

Shykind, B. M., S. C. Rohani, S. O'Donnell, A. Nemes, M. Mendelsohn, Y. Sun, R. Axel, andG. Barnea,2004. "Gene Switching and the Stability of Odorant Receptor Gene Choice." *The Cell,*Vol.117, No.6, pp.801-15.

Smale, S. T. and D. Baltimore,1989. "The "Initiator" as a Transcription Control Element." *The Cell,*Vol.57, No.1, pp.103-13.

Sosinsky, A., G. Glusman, and D. Lancet, 2000. "The Genomic Structure of Human Olfactory Receptor Genes." *Genomics,*Vol.70, No.1, pp.49-61.

Strotmann, J., I. Wanner, T. Helfrich, A. Beck, C. Meinken, S. Kubick, and H. Breer,1994. "Olfactory Neurones Expressing Distinct Odorant Receptor Subtypes are Spatially Segregated in the Nasal Neuroepithelium."*Cell and Tissue Research,*Vol.276, No.3, pp.429-38.

Strotmann, J., I. Wanner, J. Krieger, K. Raming, and H. Breer,1992. "Expression of Odorant Receptors in Spatially Restricted Subsets of Chemosensory Neurones." *Neuroreport,*Vol.3, No.12, pp.1053-6.

Sullivan, S. L., K. J. Ressler, and L.B. Buck, 1995. "Spatial Patterning and Information Coding in the Olfactory System."*Current Opinions in Genetics and Development,* Vol.5, No.4, pp.516-23.

Suster, M. L., K. Sumiyama, and K. Kawakami,2009. "Transposon-Mediated BAC Transgenesis in Zebrafish and Mice." *BMC Genomics,*Vol.10, pp.477.

Thermes, V., C. Grabher,F. Ristoratore, F. Bourrat, A. Choulika, J. Wittbrodt, and J.S. Joly.2002. "I-SceI Meganuclease Mediates Highly Efficient Transgenesis in Fish." *Mechanisms of Development,* Vol.118, No.1-2, pp.91-8.

Tinaztepe, E., 2009."Class Restriction in Odorant Receptor Gene Regulation." Molecular Biology and Genetics. Istanbul, Bogazici University, Master: 106.

Touhara, K., S. Sengoku, K. Inaki, A. Tsuboi, J. Hirono, T. Sato, H. Sakano, and T. Haga, 1999. "Functional Identification and Reconstitution of an Odorant Receptor in Single Olfactory Neurons." *Proceedings of the National Academy of Sciences USA,*Vol.96, No.7, pp.4040-5.

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.1436-44.

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.381-96.

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.681-96.

Vassar, R., S. K. Chao, R. Sitcheran, J.M. Nuñez, L.B. Vosshall, and R. Axel, 1994. "Topographic Organization of Sensory Projections to the Olfactory Bulb." *The Cell,*Vol.79, No.6, pp.981-91.

Vassar, R., J. Ngai, and R. Axel,1993. "Spatial Segregation of Odorant Receptor Expression in the Mammalian Olfactory Epithelium." *The Cell,*Vol.74, No.2, pp.309-18.

Vincent, J. D. and P. M. Lledo, 2001. "Neuronal Connectivity and Chemical Mediators İnvolved in Olfactory Message Transmission." *Bulletin de l'Academie Nationale de Medecine,*.Vol.185, No.4, pp.689-703; discussion 703-5.

Wang, G., M. Tanifuji, and K. Tanaka, 1998. "Functional Architecture in Monkey Inferotemporal Cortex Revealed by in Vivo Optical Imaging." *Neuroscience Research,*Vol.32, No.1, pp.33-46.

Wang, M. M., R. Y. Tsai, K.A. Schrader, and R.R. Reed,1993. "Genes Encoding Components of the Olfactory Signal Transduction Cascade Contain aDNA Binding Site that may Direct Neuronal Expression." *Molecular and Cellular Biology,*Vol.13, No.9, pp.5805-13.

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." *Journal of Neuroscience,*Vol.17, No.11, pp.4149-58.

Warming, S., N. Costantino, D.L. Court, N.A. Jenkins, and N.G. Copeland, 2005. "Simple and Highly Efficient BAC Recombineering using GalK Selection." *Nucleic Acids Research,*Vol.33, No.4, pp.e36.

Westerfield, M., "THE ZEBRAFISH BOOK, 5th Edition; A Guide for the Laboratory use of Zebrafish (Danio Rerio), 2007."

Volz, A., A. Ziegler, A. Ehlers, R. Younger, S. Forbes, J. Trowsdale, D. Schnorr, and S. Beck,2003. "Complex Transcription and Splicing of Odorant Receptor Genes."*The Journal of Biological Chemistry*, Vol.278, pp.19691–19701.

Xiong, W. C. and C. Montell,1993. "Tramtrack is a Transcriptional Repressor Required for Cell Fate Determination in the Drosophila Eye." *Genes and Development*,Vol.7, No.6, pp.1085-96.

Xu, Y., M. Baldassare, P. Fisher, G. Rathbun, E.M. Oltz, G.D. Yancopoulos, T.M. Jessell, and F.W. Alt, 1993. "LH-2: A LIM/Homeodomain Gene Expressed in Developing Lymphocytes and Neural Cells." *Proceedings of the National Academy of Sciences USA*, Vol.90, No.1, pp.227-31.

Yang, Z., H. Jiang, T. Chaichanasakul, S. Gong, X.W. Yang, N. Heintz, and S. Lin, 2006. "Modified Bacterial Artificial Chromosomes for Zebrafish Transgenesis." *Methods*,Vol.39, No.3, pp.183-8.

Young, J. M., C. Friedman,E.M. Williams, J.A. Ross, L. Tonnes-Priddy, and B.J. Trask, 2002. "Different Evolutionary Processes Shaped the Mouse and Human Olfactory Receptor Gene Families." *Human Molecular Genetics*,Vol.11, No.5, pp.535-46.

Young, J.M., B.M. Shykind, B. J. Trask, R.P. Lane, L. Tonnes-Priddy,J.A. Ross, M. Walker, and E.M. Williams, 2003. "Odorant Receptor Expressed Sequence Tags Demonstrate Olfactory Expression of over 400 Genes, Extensive Alternate Splicing and Unequal Expression Levels" *Genome Biology*,Vol.4, No.11, pp.R71.

Zhang, X. and S. Firestein,2002. "The Olfactory Receptor Gene Superfamily of the Mouse." *Nature Neuroscience*,Vol.5, No.2, pp.124-33.

Zhang, X. and S. Firestein,2007. "Comparative Genomics of Odorant and Pheromone Receptor Genes in Rodents." *Genomics,*Vol.89, No.4, pp.441-50.

Zollman, S., D. Godt, G.G. Privé, J.L. Couderc, and F.A. Laski,1994. "The BTB Domain, Found Primarily in Zinc Finger Proteins, Defines an Evolutionarily Conserved Family that Includes Several Developmentally Regulated Genes in Drosophila." *Proceedings of the National Academy of Sciences USA*, Vol.91, No.22, pp.10717-21.

Zozulya, S., F. Echeverri, and T. Nguyen, "The Human Olfactory Receptor Repertoire." *Genome Biology*,*2001.*Vol.2, No.6, RESEARCH0018.