CO-EXPRESSION OF ODORANT RECEPTOR GENES IN ZEBRAFISH

by

Arif Murat Atasoy

B.S., Molecular Biology and Genetics, Boğaziçi University, 2009

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

2011

ACKNOWLEDGEMENTS

I would like to thank Assist. Prof. Stefan Fuss for his guidance through my research. I learned everything from scratch in his lab. He was always willing to share his in-depth knowledge, which made working in Fishlab such a valuable experience.

I would like to thank Xalid Bayramlı; working with him was a genuine pleasure. He's truly one of the greatest people I have had the chance to meet. I would like to thank İbrahim Taştekin with whom I shared so many good discussions, faced difficulties with experiments but never ran out of ideas to exchange, to Nuray Söğünmez who always cared for her experiments as if they were her children and to Emir Tınaztepe for countless fun evenings.

I would like to thank Assist. Prof. Arzu Çelik, Ece Terzioğlu-Kara, Sercan Sayın, Mustafa Talay, Güner Kaçmaz, Bahar Şahin and Burak Tepe. I would also like to thank Assist. Prof. İbrahim Yaman and Nehir Banaz for their help with cell culture experiments; Assist. Prof. Hikmet Budak and Melda Kantar for their help with RACE experiments.

I would also like to thank Renin Hazan, Çiğdem Köroğlu and Ayşe Güven, whose friendship and support kept me motivated in the last two years.

I would like to thank TÜBİTAK BİDEB for the masters scholarship they kindly provided during my research.

Finally, I would like to thank members of my family, Nihal Atasoy, Aydın Atasoy and Selim Atasoy. They have always helped me make the best decisions in life.

ABSTRACT

CO-EXPRESSION OF ODORANT RECEPTOR GENES IN ZEBRAFISH

Olfactory receptors constitute the largest G-Protein-Coupled Receptor family in humans and in other vertebrates. Olfactory receptor genes are expressed in the olfactory sensory neurons in the olfactory epithelia and are responsible in detection of odorant molecules. The mechanisms that regulate OR gene choice are still largely unknown, but it has been shown through various methods that the majority of olfactory sensory neurons only express a single OR gene. This phenomenon has been named the one receptor-one neuron rule and it has been shown, during development, the OSNs that express more than one are selectively reduced, possibly by apoptosis. In zebrafish, which has a repertoire of 136 intact OR genes, an exception to the one receptor-one neuron rule has been reported. Members of the OR103 family have been shown to be co-expressed and the nature of this co-expression was OR103-1 being co-expressed with at least one of OR103-5 and OR103-2 in every cell OR103-1 is expressed. In this thesis, I designed experiments to find the underlying mechanism of this co-expression. I used 5'-RACE, RT-PCR, and gene-specific RACE to obtain transcript information of several OR103 members, which have revealed a possible single transcript encompassing both OR103-1 and OR103-5 coding sequences, separated by the intergenic region. In order to understand if both OR103-1 and OR103-5 could be translated from this single mRNA, which would require IRES activity by the intergenic sequence, I used tissue specific promoters to express, in *vivo*, two different fluorescent proteins that have been separated by the intergenic region. Finally, I used Hk-2 cell cultures to quantitatively test for possible IRES activity of the intergenic sequence in vitro.

ÖZET

ZEBRABALIĞINDA KOKU RESEPTÖR GENLERİNİN BİRLİKTE EKSPRESYONU

Koku alma genleri insanlarda ve diğer omurgalılardaki en büyük G-protein-eşli reseptör familyasını oluşturur. Koku alma genleri koku duyu epitelindeki koku duyu nöronlarında anlatılılar ve koku moleküllerinin algılanmasında rol oynarlar. Koku alma gen seçimini kontrol eden mekanizmalar hala büyük ölçüde bilinmemektedir ama koku alma duyu nöronlarının büyük çoğunlukla tek bir koku alma geni anlattıkları çeşitli metodlarla gösterilmiştir. Bu fenomen tek reseptör-tek nöron kuralı olarak adlandırılmıştır ve gelişim sürecinde birden fazla koku alma geni anlatan koku duyu nöronlarının büyük ihtimalle kontrollü hücre ölümü yoluyla sayılarının azaldığı gözlemlenmiştir. 136 koku duyu geni olan zebrabalığında tek reseptör-tek nöron kuralına bir istisna daha önceden gösterilmiştir. OR103 familyasının üyelerinde tek hücrede ikili ekspresyon görülmüştür ve bu ikili ekspresyonun şekli her OR103-1 pozitif hücrede OR103-5 ve/ya OR103-2'nin birlikte anlatıldığıdır. Bu tezde, bu ikili anlatımın altında yatan sebepleri bulmaya çalıştım. RACE, RT-PCR ve gen-spesifik RACE kullanarak OR103 üyeleri için transkript bilgisini buldum ve OR103-1 ve OR103-5'in tek bir transkriptte aralarındaki intergenik bölgeyle ayrılmış şekilde birlikte anlatıldığını gözlemledim. Bu tek mRNA'dan iki genin proteinlerinin sentezlenebiliceğini, yani intergenik bölgenin IRES aktivitesine sahip olup olmadığını test etmek için zebrabalığında doku spesifik promotörlerle iki farklı floresan proteinin anlatıldığı in *vivo* çalışmalar yaptım. Son olarak, Hk-2 hücre kültüründe intergenik sekansın IRES aktivitesini kuantitatif şekilde test ettim.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS / ABBREVIATIONS

BAC	Bacterial Artificial Chromosome
bp	Base Pair
cAMP	Cyclic Adenosinemonophosphate
cDNA	Complementary Deoxyribonucleic Acid
DNA	Deoxyribonucleic Acid
EMCV	Encephalomyocarditis Virus
Gap43	Growth Associated Protein-43
GFP	Green Fluorescent Protein
GPCR	G-protein-coupled Receptor
IRES	Internal Ribosome Entry Site
kb	Kilobase Pair
mRNA	Messenger Ribonucleic Acid
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
RLM	RNA-ligase Mediated
RFP	Red Fluorescent Protein
RNA	Ribonucleic Acid
RT-PCR	Reverse-transcription Polymerase Chain Reaction
UTR	Untranslated Region

1. INTRODUCTION

1.1. The Olfactory System

1.1.1. Overview

Olfaction starts when odorant molecules from the environment bind to olfactory receptors (ORs), which are located on the cilia of olfactory sensory neurons (OSNs). OR gene constitute the largest gene family in every genome and individual OR genes code for seven-transmembrane G-protein-coupled receptors (GPCRs). Binding of the odor ligand to the OR activates an OSN-specific heterotrimeric G-protein that contains the olfactory speific a subunit Ga_{olf}, which in turn activates a specific adenyl cyclase type III. Increased levels of cAMP open an olfactory specific isoform of the non-selective cation channel OCNC1, which allows for the conductance of Ca^{2+} ions. Ca^{2+} ions then activate a Ca^{2+}-activated chloride channel, opening of which accounts for most of the depolarization current in OSNs.

Stem cells in the basal layer of the OE give rise to developing OSNs, which make a selective choice of OR gene expression from the large repertoire of OR genes in the genome (Malnic *et al*., 1999). After an initial choice has been made, the OSNs commit to expression of only the chosen OR for the rest of their lives (Serizawa *et al.*, 2003; Lewcock and Reed, 2004; Shykind *et al.*, 2004) however, switching of OR gene expression during an early phase of the lifespan of OSNs has been reported for some OR genes (Shykind *et al*., 2004). As they mature, OSNs project their axons to the olfactory bulb, a specilized region of the forebrain, where they form the glomeruli. Here OSN axons form synapses with mitral cell projection neurons and local interneurons (Ressler *et al.*, 1993; Vassar *et al.*, 1993; Mombaerts *et al.*, 1996). OSNs that express the same OR project their axons to one of two homogenous glomeruli in the OB. The array of glomeruli in the olfactory bulb

forms a layer of relay centers with extensive lateral connections between glomeruli before mitral cells take the information to the olfactory cortex in the brain (Mori *et al.,* 1999).

The targeting of axons to the right glomerulus strongly depends on the expressed OR (Wang *et al*., 1998) and OSN activity has been shown to play a role in OSN survival and fine mapping of anatomical connections in the olfactory system (Watt *et al.,* 2004; Zhao and Reed, 2001, Tian and Ma, 2008).

OR genes are phylogentically classified into two distinct classes: Class I and Class II OR genes. Class I genes are thought to have evolved earlier and respond primarily to water soluble molecules. Class II ORs are the predominant type in mammals, constituting about 90% of the OR repertoire, and are thought to detect primarily airborne odorants. A strong evidence for this comes from a study on the amphibian African crawfrog *Xenopus*, for which it has been shown that the two different classes of ORs are expressed in different cavities of the *Xenopus* nose, one that is accessible to water and another one that is accesible to air (Freitag *et al.*, 1995).

Studies have shown that the identity of an odor is encoded by a combinatorial code in the olfactory system (Friedrich and Korsching, 1997; Malnic *et al.,* 1999; Meister and Bonhoeffer, 2001). These study showed by calcium imaging of OSNs expressing single ORs (Malnic *et al*., 1999) or by activity recording from olfactory bulb glomeruli (Friedrich and Korsching, 1997; Meister and Bonhoeffer, 2001) that each OR can be activated by a number of different molecules with different affinities. Moreover, a single odorant molecule has been shown to be able to activate more than one OR. The identity of the odor is thus translated into the combinaton of ORs to which the odor molecules can bind and the glomeruli that are activated.

1.1.2. The Olfactory System in Zebrafish

The anatomy and molecular organization of the zebrafish olfactory system is similar to the rodent olfactory system. However, the OR repertoire is much smaller, with approximately 136 ORs, most of which are thought to be Class I receptors (Alioto and Ngai, 2005; Niimura and Nei, 2005). Also, the number of pseudogenes seems to be lower in zebrafish than in mammals, where about 10% of the OR genes are pseudogenes. Corresponding to the lower number of OR genes, there are also fewer glomeruli in the OB (Baier and Korsching, 1994). Different odorants have been shown to activate fixed patterns of glomeruli in the zebrafish OB (Friedrich and Korsching, 1997).

1.2. Monogenic and Monoallelic Expression of OR genes

It has been shown that OR genes are expressed in a monogenic fashion in OSNs (Malnic *et al*., 1999; Serizawa *et al.*, 2000, 2003). Thus, every single OSN expresses only a single member of the large OR gene repertoire. Interestingly, individual OSNs also only express one of the two parental alleles (Chess *et al*., 1994). When mice from two different backgrounds, carrying a polymorphic I7 allele were bred, single cell RT-PCR revealed that individual I7 cells did this in a monoallelic fashion (Chess *et al.*, 1994). The study used small pools of isolated OSNs, ensuring by statistics that each pool, most of the time, by probability would only get one cell expressing I7 gene. Only one of two polymorphic alleles could be amplified from most of the pools analyzed. The same study has shown that there was no preference for each of the different alleles and they were expressed with equal frequencies.

In other experiments, strict monoallelic expression of M71 has been demonstrated when the two alleles of the M71 OR gene were tagged with different fluorescent proteins (Li *et al.*, 2004). Distinct OSN cell populations were observed that were positive for either of the two reporter genes in equal numbers. Interestingly, even transgenic OR sequences are subject to mechanisms that ensure monoallelic and monogenic expression and further

proved each cell chooses only one allele from all available OR genes (Serizawa *et al.,* 2000).

The combination of monogenic and monoallelic expression of OR genes has become famous as the “one receptor-one neuron rule” in the olfactory system and is often also referred to as “singular expression”. The mechanisms that underlie the rule are largely unknown. However, its presence indicates that expression OR genes is strictly regulated (Fuss and Ray, 2009).

1.2.1. One Receptor-One Neuron Rule

The one receptor-one neuron rule is a widely accepted but equally doubtful phenomenon (Mombaerts, 2004). The rule implies that each OSN chooses and expresses only one receptor from the large genomic repertoire. The evidence to this rule comes from two different approaches. The first approached used single-cell RT PCR approach in which cDNA from a single cell is obtained and amplified by degenerate primers designed to amplify any OR gene that might be expressed by the cell (Malnic *et al.,* 1999; Touhara *et al.,* 1999). However, this approach is not very conclusive since most of the time it failed to amplify an OR from the OSN due to low amounts of RNA, thus, making it doubtful if two OR genes could be amplified by this method even if they were present (Mombaerts, 2004). In fact, one study could show an exception to the rule with a similar approach in which two OR genes from cDNA of a single mature rat OSN have been amplified (Rawson *et al.,* 2000). The second approach is a statistical argument based on the expression frequencies of OR genes in the OE using in situ hybridization. It has been shown in mice that OR genes can be found with approximate expression frequencies of 0.1% (Vassar *et al.,* 1993; Ressler *et al.,* 1993). Given that there are about 1000 to 1500 OR genes expressed in the mouse olfactory tissue, each OR would be expressed in a different cell population. In zebrafish, the same approach has shown expression frequencies of 0.5–2% with approximately 130 expressed OR genes (Chess *et al.,* 1992) which is in support of the idea that every OSN expresses only one OR.

Finally, somewhat indirect evidence comes from the number of glomeruli in the OB of the mouse. Since OSNs expressing a given OR mostly project their axons to two glomeruli per OB together with only the cells that are expressing the same OR, the number of glomeruli would be expected to be around 3000 for each bulb, which is indeed observed (Bozza *et al.,* 2009).

1.2.2. Silencing by a Functional OR and Second Choice

Expression of a functional OR protein has been shown to actively prevent expression of more than one OR gene per OSN. If the OR coding sequence is mutated by deletion or a frame-shift mutation, the same OSNs that express the mutated OR locus also express another OR gene (Serizawa *et al.,* 2003; Lewcock and Reed, 2004). This phenomenon has been named "second choice". How the cell detects the presence of a functional OR protein and the mechanisms that prevent second choice if one OR is already successfully expressed by the OSN are not understood. Often second choice occurs from the other allele of the deleted OR gene or from other ORs that are expressed in the same expression zone in the OE (Shykind *et al.,* 2004). This mechanism might have evolutionary relevance for the olfactory system in general because there are many pseudogenes in the mammalian olfactory gene repertoires (Young *et al.,* 2002) and second choice allows the cells that initially choose a non-functional OR to switch to a functional OR protein. Otherwise, more and more functionally inactive OSN cell populations expressing pseudogenes would accumulate in the OE over evolutionary time.

Second choice has also been reported for the zebrafish olfactory system. When a bacterial artificial chromosome (BAC) that contains 16 OR genes on it including the entire 111 and 103 subfamilies was modified so that the OR103-1 and OR111-7 coding sequences have been replaced with sequences coding for fluorescent proteins and injected into zebrafish oocytes, it has been shown that fluorescently labeled OSNs express other OR genes that were linked to the mutated locus. Thus, mostly genes from the corresponding subfamilies were seen, with the exception of OR106-2 (Sato *et al.,* 2007).

There's also evidence for switching of functional OR genes early during the development of OSNs (Shykind *et al.,* 2004). When cells that expressed a functional MOR28 allele at some point throughout their lives have been marked permanently with RFP through Cre-mediated recombination in lineage tracing experiments, it has been shown that 10% of the marked cells have switched to express other receptors.

1.3. Co-Expression

While the majority of OSNs only express a single receptor in a monogenic and monoallelic fashion in mice and the zebrafish, certain exceptions to this rule have been reported (Rawson *et al.,* 2000, Tian and Ma, 2008, Sato *et al.,* 2007). Co-expression of OR genes are an interesting exception to the above discussed one receptor-one neuron rule, for both developmental and physiological aspects. Understanding how and why certain neurons break the rule to which the majority of OSNs obeys might lead to a better understanding of how co-expression is prevented in most OSNs.

1.3.1. Co-Expression in Mouse Olfactory System

A study taking advantage of a combination of single-cell RT PCR using degenerate primers and 3'-RACE identified cells co-expressing the I9 and HGL-SL2* OR genes in the rat olfactory epithelium (Rawson *et al.,* 2000) The study used 3'-RACE to select for ORs in the single cell isolate and to eliminate the possibility of genomic amplification, while using degenerate primers, instead of OR-specific primers, to targ*et al*l OR genes expressed in the cells analyzed. Interestingly all HGL-SL2*-positive OSNs were positive for I9 while only a fraction of I9 cells was positive for the second OR. Thus, while the co-expression was systematic, it was not 100% complete.

It has further been shown, through in situ hybridizations on the mouse septal organ, that MOR256-3 is co-expressed with at least one of eight other receptors at a frequency of approximately 2% in newborn mice (Tian and Ma, 2008). The same study shows the same

pattern of co-expression in four weeks old mice, but with a much lower frequency that has dropped to 0.2% over this period. Interestingly, when one side of the nose was kept under sensory deprivation by naris closure, the co-expression rate remained high (1.5%) even at 4 weeks after birth. Those results suggest an activity dependent elimination of neurons expressing more than one receptor. To prove that there is a selective elimination of co-expressing cells by apoptosis, the authors used Bax null-mutant mice, which have impaired apoptotic capability. When Bax null-mutant mice were tested for co-expression of MOR256-3 and the other eight OR genes, the co-expression rate has been shown to be around 1.6% at 4 weeks after birth, suggesting active cell death as the mechanism for elimination of co-expressing OSNs.

Due to the large pool of odorant receptors in mice, it's difficult to analyze them all one by one or all with possible combinations for systematic co-expression. This is also true for the zebrafish even though it has a relatively smaller repertoire of odorant receptors where the 136 ORs allow for 9180 different combinations. Therefore, a comprehensive analysis of co-expression testing high portions of the available gene pool has not yet been attempted.

1.3.2. Co-expression in the Zebrafish Olfactory System

Among the two subfamilies (103 and 111) that have been examined for co-expression in the adult wild-type zebrafish, OR103-1 and OR103-5/2 have been found to be co-expressed (Sato *et al.,* 2007). The study used in situ hybridization under stringent conditions to show that members of 111 were never co-expressed with members of 103 family; members of 111 family were also never co-expressed with each other, but OR103-1 was always co-expressed with OR103-5 and/or OR103-2. Whether only one or both of OR103-5 and OR103-2 were co-expressed with OR103-1 could not be determined due to high sequence similarity between those two genes. When two zebrafish OR genes, OR103-1 and OR111-7 have been replaced with fluorescent proteins in the context of a transgenic BAC construct and integrated into the genome of the zebrafish, most of the second choice occurred for OR103-5/2. However, this may not be a classic "second choice", but rather

the cells that would normally express OR103-1 and OR103-5/2 continue to express OR103-5/2 even though OR103-1 was replaced with a fluorescent reporter gene.

The observation that there is systematic co-expression of OR genes among members of the OR103 family forms the starting point of this thesis, which aims at an investigation of the underlying mechanism of this co-expression.

1.4. IRES Sequences

Internal ribosome entry sites were initially discovered in RNA viruses (Pelletier and Sonenberg, 1988, Jang *et al.,* 1989). Normally, translation from a capped, monocistronic mRNA occurs after the ribosome assembles at the 5'-end of the mRNA and scans for the start codon (Kozak, 1989). However, viral RNAs do not have 5'-caps and use sequences with IRES activity to recruit the ribosomes for translation of such transcripts. IRES sequences have fast become popular as tools in biological research because of their ability to report gene expression from bicistronic transcripts without affecting the original gene. The IRES sequence most commonly used is from Encephalomyocarditis virus EMCV (Jang *et al.,* 1989).

There have also been reports of IRES sequences from eukaryotic systems (Macejak and Sarnow, 1991, Stoneley *et al.,* 1998). However, many claimed IRES sequences from eukaryotic systems actually have turned out to have cryptic promoter activity and the number of actual eukaryotic IRES sequences is much lower than proposed (Bert *et al.,* 2006).

1.4.1. IRES Sequences in the Zebrafish Genome

Although there have been reports that the common EMCV IRES sequence can be used to drive translation from bicistronic transcripts in tissues in the zebrafish (Fahrenkrug *et al.,* 1999), attempts to use it to report OR protein expressing cells in the zebrafish

olfactory system have failed (Sato *et al.,* 2007). This presents a need for a different sequence to label cells expressing OR proteins in the OE of the zebrafish.

There are two known IRES sequences in the zebrafish genome (Ul-Hussain *et al.,* 2008; Lekven *et al.,* 2001). The first one resides on the 5' UTR of connexin-55.5 mRNA (Ul-Hussain *et al.,* 2008). However, this IRES sequence has been shown to be highly regulated to show activity only in neurons where polyprimidine tract binding protein is present. The second IRES sequence known in the zebrafish genome resides in between the two open reading frames of *wnt8*. However, it has been shown to have low efficiency in vitro and its efficiency in *vivo* remains unknown (Lekven *et al.,* 2001).

Here we present studies to investigate the nature and the mechanism of the co-expression of the OR103-1 and OR103-5/OR103-2 genes that has been described in the zebrafish olfactory system. We first examined whether co-expression of two OR genes is an independently regulated process or if it results from a failure of transcriptional termination of the OR103-1 gene. We could show that a single long transcript might exist that could account for the observed co-labeling of a specific cell population by in situ hybridization for the previously co-expressed OR genes. This finding led me to the question, whether co-expression is only seen at the level of the mRNA or if the two OR genes could also be co-translated. We obtained preliminary evidence that the intergenic region between the OR103-1 and OR103-5 genes might have IRES function and could promote translation of two coding sequences from a bicistronic transcript. We tested the IRES activity of the intergnic region using expression analysis of transgenic constructs in zebrafish in *vivo* and in cell culture systems in vitro. A high level of co-translation was observed with both approaches, making the OR103-1/OR103-5 an interesting new tool for studies of gene expression and manipulation in zebrafish.

2. PURPOSE

In this thesis we aimed to find the underlying mechanism of the only reported co-expression of OR genes in the zebrafish. OR103-1 has previously been reported to be always co-expressed with at least one of OR103-5 and OR103-2. Since this constitutes an exception to the widely true one receptor-one neuron rule, it would be interesting to understand how a stable co-expression of OR genes can occur and what makes this case different so that it can escape the tight regulation that ensures expression of only a single gene. One of alternative scenarios to test were if the co-expressed OR genes were on a single transcript and expressed by the activation of a single promoter. The second possible scenario was that the co-expressed OR genes had their individual promoters activated simultaneously in the same OSN, giving rise to individual transcripts.

3. MATERIALS AND METHODS

3.1. Materials

3.1.1. Fish

Adult and embryonic specimens of 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.

3.1.2. Equipment and Supplies

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 according to the manufacturer's 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 at 28 ^{o}C and a 14 hour light/10 hour light/dark cycle. They were fed four times a day, three times with flake food and once with live brine shrimp (*Artemia sp.*).

Matings were set up in 1 L breeding tanks in the previous evening with separators separating male and female fish. The separators were removed in the morning and the eggs were collected immediately after fertilization.

3.2.2. Olfactory Tissue Extraction

For extraction of total RNA from olfactory epithelia, 10 fish were euthanized in ice water for 10 minutes and the olfactory epithelia were dissected out in 1X PBS under 20X magnification with a stereomicroscope. First the head was separated from the body and then the jaw was separated with surgical scissors. The tissues around the epithelia were loosened by scraping around the olfactory rosettae with foreceps until the epithelia could be removed intact. The epithelia were kept overnight in TRIzol™ (Invitrogen) in an 1.5 ml eppendorf in -80 ^{o}C.

3.2.3. Total RNA Isolation

Total RNA isolation of olfactory epithelia was performed following a standard TRIzol™ (Invitrogen) protocol according to the manufacturer's instructions. The extracted and precipitated total RNA was dissolved to a final volume of 20 µl with DEPC-treated H_2O.

3.2.4. Polymerase Chain Reaction

PCR reactions were carried out with GoTaq® Flexi DNA Polymerase (Promega) or Advantage® 2 Polymerase Mix (Clontech) according to manufacturer's protocols. Template DNA amounts were between 1 ng to 100 ng. A maximum of 24 cycles were used in standard PCR program.

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 Reaction Components

1 µl Template
5 µl 10X Buffer or 10 µl 5X buffer
3 µl $MgCl_2$ (25 mM) (only with Gotaq polymerase)
1 µl dNTP mix (10mM)
2.5 µl Forward Primer (10µM)
2.5 µl Reverse Primer (10µM)
0.5 µl Enzyme
up to 50 µl dH_2O

Table 3.2. Standart PCR Program

95 ^oC 2 minutes
24 cycles of:
95 ^oC 30 seconds
$T_{annealing}$ ^oC 30 seconds
72 ^oC 1 minute for each kb of expected product
72 ^oC 7 minutes
4 ^oC 20 minutes

3.2.5. First-Strand cDNA Synthesis

First-Strand cDNA synthesis was performed using the SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen) according to the manufacturer's manual. Shortly, 1 µg of total RNA of olfactory epithelia was mixed with 1 µl of 10 mM dNTP mix 1µl of 0.5 µg/ml oligo$(DT)_{12-18}$ primer and 6 µl of depc-treated H_2O. The mix was incubated at 65 ^oC for 5 minutes and placed on ice. 2 µl of 10X RT buffer 4 µl of 25mM $MgCl_2$, 2 µl of 0.1M DTT and 1µl RNaseOUT™ was added and the reaction was kept on 42 ^oC for 2 minutes. Then, 1 µl of SuperScript™ II RT was added and the mix was incubated at 42 ^oC for 50 minutes. The reaction was stopped by incubating at 70 ^oC for 15 minutes. 1µl of RNase H was added and the reaction was incubated 20 minutes at 37 ^oC. The sample was then stored at -20 ^oC.

3.2.6. RT-PCR on First-Strand cDNA

Advantage® 2 Polymerase Mix (Clontech) was used for RT-PCR on First-Strand cDNA. 1 µl of the first-strand cDNA was used as template. 1 µl of 10 mM dNTP mix, 2.5 µl of 10µM stocks of each primer was used with supplied 10X buffer. Thermal cycler (Bio-Rad) was programmed to 2 minutes at 95 ^oC and cycles of 30 seconds at 95 ^oC, 30 seconds at 4-5 ^oC degrees below the annealing melting temperature of the primers as determined by the Oligocalc online tool[1], 1 minute per expected kilobase of PCR product at 72 ^oC. Up to 35 cycles were performed and followed by 5 minute at 72 ^oC. The products were analyzed on 1% agarose gel with EtBr and stored at -20 ^oC.

3.2.7. RNA Ligase Mediated Rapid Amplification of cDNA Ends (RLM-RACE)

For RLM-RACE, the FirstChoice® RLM-RACE Kit (Ambion) was used. The steps were performed according to the manufaturer's manual. 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 degraded mRNAs, rRNAs and tRNAs. CIP was removed by

1- http://www.basic.northwestern.edu/biotools/oligocalc.html

phenol:choloform followed by chloroform extraction and the sample was treated with 1µl Tobacco Acid Pyrophosphatase 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 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 manufacturer's manual. 1 µl of the first-strand was used as template for PCR with outer primer and gene specific outer primer. 35 cycles of PCR was performed and 1 µl of the product was used as template for another PCR with inner primer and gene specific inner primer for 35 cycles. Gotaq® DNA Polymerase (Promega) was used in those PCR reactions with supplied 5X buffer and 1 µl of 10mM dNTP mix and 3 µl of 25mM $MgCl_2$. The product was cloned to pGEMT-easy vector (Promega) and sequenced.

3.2.8. Restriction Digests of DNA

For restriction digests of DNA, restriction enzymes from New England Biolabs or Promega were used. Digestion reactions were set up with 3 - 5 Units / µg of enzyme per microgram of DNA sample in final concentration of 1X of supplied buffer. The reactions were incubated at 37 ^{o}C for up to 8 hours.

3.2.9. Gel Electrophoresis and Gel Extraction of DNA

DNA samples were run on 1% agarose gels stained with EtBr. As a marker 1 kb DNA Ladder (NEB) has been used in all gels except for Figure 4.8a where 100 bp DNA Ladder (NEB) has been used.

Table 3.3. Primers used in this thesis.

Name of Primer	Sequence (5'-3')
RT-PCR-forward	GCTATGGTTGCGTTGCTTACTCCCC
RT-PCR-reverse	CAGAAGTGCCGATACAAGGCCATAATAATGAGGC
103-1RACE1	GCATCAGAGCAGCCACTGAGTCAAAGTATCCC
103-1RACE2	GGCCAGGGGAGTAAGCAACGCAACCATAGC
103-5RACE	CAGAAGTGCCGATACAAGGCCATAATAATGAGGC
Inner Primer	CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG
Outer Primer	GCTGATGGCGATGAATGAACACTG
103-5-2-R	CACGGATGACCAGTGCCAAAGG
103-5-5UTR-R	GCAGTAGTCTGTGAAACAACTTTGGC
Intergenic Reg-SphI-forward	TGCATGCTTTTAGAAGTCATATATTATAAATGC
Intergenic Reg-NcoI-reverse	ACCATGGTGAGTATTAGTCGGGGGGGGAAATCTC
Bactin-forward	CTGGGATGACATGGAGAAGATCTG
Bactin-reverse	CCTTGATGTCACGGACAATTTCTC
T7-highTM	TAATACGACTCACTATAGGGCGAATTGG
M13-R-highTM	GGAAACAGCTATGACCATGATTA
GFP-forward	GCGACGTAAACGGCCACAAGTT
GFP-reverse	TACTTGTACAGCTCGTCCATGCCG
103-1-5-Int-F	CGAGTACGCCCCATTATGGAAATG

The High Pure PCR Purification Kit (Roche) was used to extract DNA fragments from agarose gels after separation. The sample was run on agarose gel and the band of interest was cut with a scapel under UV light. The gel slice was melted in the supplied Binding Buffer and centrifuged through a spin column. The column was washed twice with Wash Buffer and DNA was eluted by elution buffer. The purity and the concentration of the obtained DNA was measured using a NanoDrop Spectrophotometer.

3.2.10. Purification of DNA

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

3.2.11. Ligation of DNA to Vectors

DNA fragment to be cloned (insert) and the required vector was prepared by suitable restriction digests and DNA purifications as described above. 3 µl of the purified insert and vector was run on agarose gel side by side to estimate relative intensity. Using the sizes and the estimated relative intensity of the samples, a 1:3 molar vector to insert ratio was estimated. Vector and insert DNAs were mixed together not exceeding a total of 100 ng DNA and H_2O was added to bring the reaction volume to 8 µl. 1 µl of T4 DNA Ligase (NEB) was added after addition of 1 µl of supplied 10X T4 Ligase Buffer. The reactions were kept at room temperature for 1 hour and transformed to competent cells.

For ligation of PCR products directly, the pGEMT-easy Vector System (Promega) was used. 3.5 µl of PCR products was mixed with 0.5 µl of pGEMT-easy vector and 5 µl of supplied 2X Rapid Ligase Buffer. 1 µl of T4 DNA Ligase (Promega) was added and the reaction was kept at room temperature for 30 minutes and transformed to competent cells.

3.2.12. Preparation of Competent Cells

A single colony of the E. Coli strain TOP10-MRF' was inoculated to 100 mL LB and grown at 37 ^{o}C in a shaking incubator a few hours until the culture reached an OD_{550} of 0.6. The culture was incubated on ice for 15 minutes and centrifuged at 4000 rpm for 10 minutes at 4 ^{o}C. The supernatant was discarded and the pellet resuspended in 10 mL 0.1 M $CaCl_2$. The mixture was centrifuged again for 10 minutes and the pellet resuspended in 4 mL 0.1 M $CaCl_2$, 15% glycerol solution. The resulting competent cells were stored at – 80 ^{o}C in 50 µl aliquots.

3.2.13. Transformation of Plasmids to Competent Cells

Competent cells were thawed on ice for 5 minutes and up to 10 µl of ligation reactions or plasmid DNA were added. The mixture was incubated 5 minutes on ice, incubated 1.5 minutes at 42 ^{o}C and 5 more minutes on ice. The mixture was then spread to ampicillin plates and incubated overnight at 37 ^{o}C. Colonies were picked the next day and grown in 3 ml LB overnight. The next day, plasmids were isolated from cells with QIAprep Spin Miniprep Kit (Qiagen) and analysed by analytical digests for success of cloning.

3.2.14. Analytical Digests

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

3.2.15. Injection of DNA into Zebrafish Embryos

Zebrafish eggs were collected in the morning and sorted to the side of a petri dish by pipetting out all the water. Glass capillary needles were filled with 2 µl injection solution which consisted of 50 ng / µl plasmid, 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® (Eppendorf).

3.2.16. Culture of Hk-2 Cells and Transfection of Plasmids

Hk-2 Cell were stored at -80 ^{o}C as aliquots. One aliquot was spread on a 10 cm petri dish, 5 ml DMEM Low Glucose (Gibco) with 100 U/ml penicillin and 100 µg/ml streptomycin was added and the cells were kept at 37 ^{o}C until they reach the desired density. The cells were lifted from the plate by 5 minute treatment with 0.25% HyClone® Trypsin (Thermo Scientific), resuspended in 10 ml medium. 10 µl aliquot was stained with tryphan blue and cells were counted under microscope. 50000 or 75000 cells were distributed into each well of 12-well plates and incubated 24 hours before transfection.

Transfection was done as follows: a mixture of 1 µg plasmid to be transfected was mixed with medium up to 47.5 µl. Then 2.5 µl FuGENE® HD Transfection Reagent (Promega) was added. This mixture was incubated for 15 minutes at room temperature and added to the well containing the cells to be tranfected.

3.2.17. Lysis of Cells and Luciferase Assays

Lysis of cells and the luciferase assays were performed using the Dual-Luciferase® Reporter Assay (Promega) according to the manufacturer's protocol. Briefly, 250 µl Passive Lysis Buffer was dispensed into each well and the 12-well plates were incubated at room temperature with gentle shaking for 15 minutes. 100 µl of the lysates were taken to 96-well plates and the plates were inserted to Fluoroskan Ascent FL luminometer (Thermo

Scientific). The luminometer was set to dispense 100 µl LARII, delay 2 seconds and measure luminescence for 10 seconds. After measurement, 100 µl Stop & Glo® Reagent was dispensed by the luminometer and luminescence was measured for 10 seconds.

4. RESULTS

Co-expression of OR genes by the same OSN in the OE is a rare phenomenon (Rawson *et al.*, 2000; Tian and Ma, 2008; Sato *et al.*, 2007) and cases of co-expression constitute an interesting exception to the general 'one neuron-one receptor' rule (Mombaerts, 2004). However, the molecular mechanisms that underlie this rule are not well understood. A better understanding of the nature of co-expression of OR genes may therefore help to shed some light on the mechanisms that usually ensure the exclusive expression of a single OR gene in the majority of OSNs. Thus, it is interesting to study cases in which OSNs systematically escape the one receptor – one neuron rule and how a low number of sensory neurons express more than one receptor, while the other OSNs do not. In addition, because the function of OR proteins is to detect odorant molecules, a systematic co-expression might have an important role in the perception of smell.

Here, we present a detailed analysis to examine the nature of a previously reported case of systematic co-expression between different OR genes from the OR103 family in zebrafish (Sato *et al.*, 2007). In this example, the OR103-1 gene was reported to be co-expressed with either one or both of OR103-5 and OR103-2 genes in every OR103-1-expressing cell. Because of the high degree of sequence similarity between the OR103-2 and OR103-5 genes, the study could not discriminate between those two genes. Interestingly, not all OR103-5/OR103-2-expressing OSNs were positive for the OR103-1 gene, suggesting that either OR103-5 or OR103-2 is co-expressed with OR103-1, or that different OSN populations exist, one that co-expresses the combination of OR103-1 and OR103-5/2 and one that expresses the OR103-5 and OR103-2 genes in a normal monogenic fashion.

Members of the OR103 family reside on chromosome 15 as a part of a larger cluster and is composed of five OR genes, which are all in the same transcriptional orientation. Of these genes, OR103-1 is the first gene among the family in terms of genomic position and the putatively co-expressed OR103-5 and OR103-2 genes are located directly downstream of OR103-1 (Figure 4.1). The intergenic region between OR103-1

and OR103-5 spans 1432 bp and the intergenic region between OR103-5 and OR103-2 spans 4692 bp, respectively. Two more members of the OR103 family are located downstream of OR103-2 and the entire OR103 family is flanked by members of the OR111 and OR102 families, respectively.

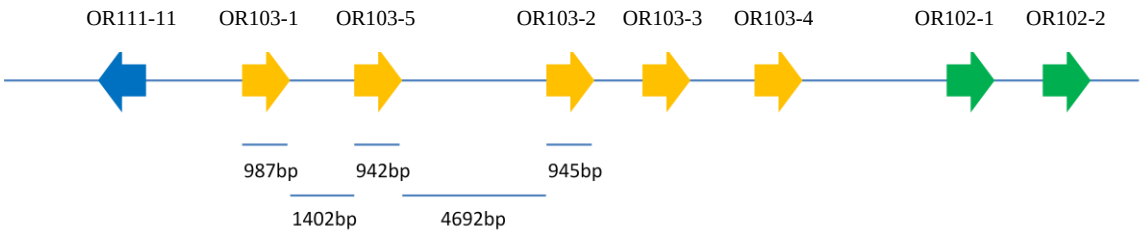


Figure 4.1. Genomic Locus of The OR103 Family

OR103 family is located between OR111 and OR102 families. There are five members, named in the order they were identified. OR103-1 is followed by OR103-5, OR103-2, OR103-3 and OR103-4, respectively.

In principal, the observed co-expression between OR103 family members could be due to at least two alternative mechanisms. In one scenario, the expression of OR103-1 could be independent of OR103-5 / OR103-2. In this case of 'true' co-expression, the observed co-detection of two OR transcript by in situ hybridization would be a result of the independent activation and expression from the OR103-1 and OR103-5 promoters. In an alternative scenario, co-expression between OR103-1 and OR103-5 could result from a failure of transcriptional termination giving rise to a single large transcript that harbors both OR coding sequences. We wanted to discriminate between these alternative models to gain insight into the nature of the observed co-expression phenomenon.

My starting hypothesis was that the co-expression might be between the OR103-1 and OR103-5 genes because this hypothesis is easy to address experimentally. Because OR103-1 and OR103-5 are close in terms of their genomic distance, which spans only 1432 bp, we designed experiments to test whether we could find any evidence for independent transcription of the two genes or whether OR103-1 and OR103-5 can be found on a single transcript.

4.1. Transcript Analysis of OR103-1 and OR103-5

To analyze transcripts that encode the OR103-1 and OR103-5 genes, we employed three independent methods: rapid amplification of cDNA ends (RACE), reverse transcription polymerase chain reaction (RT-PCR), and gene-specific RACE. To generate the required cDNA from olfactory tissues that can be used in these experiments, we collected olfactory epithelia from 10 zebrafish to extract total RNA. The extracted total RNA was used to synthesize cDNA and RLM-RACE 5' cDNA using oligo$(DT)_{12\text{-}18}$ primer. Figure 4.2a shows OE total RNA run on 1% agarose gel. B-actin, which is expressed in all cells, is used as a control of cDNA synthesis. PCR with B-actin-Forward and B-actin-Reverse gave expected 407 bp band from +RT and no band was amplified from –RT (Figure 4.2b). 5'RACE of B-actin with Outer Primer and B-actin-Reverse gave a 712 bp band as expected (Figure 4.3).

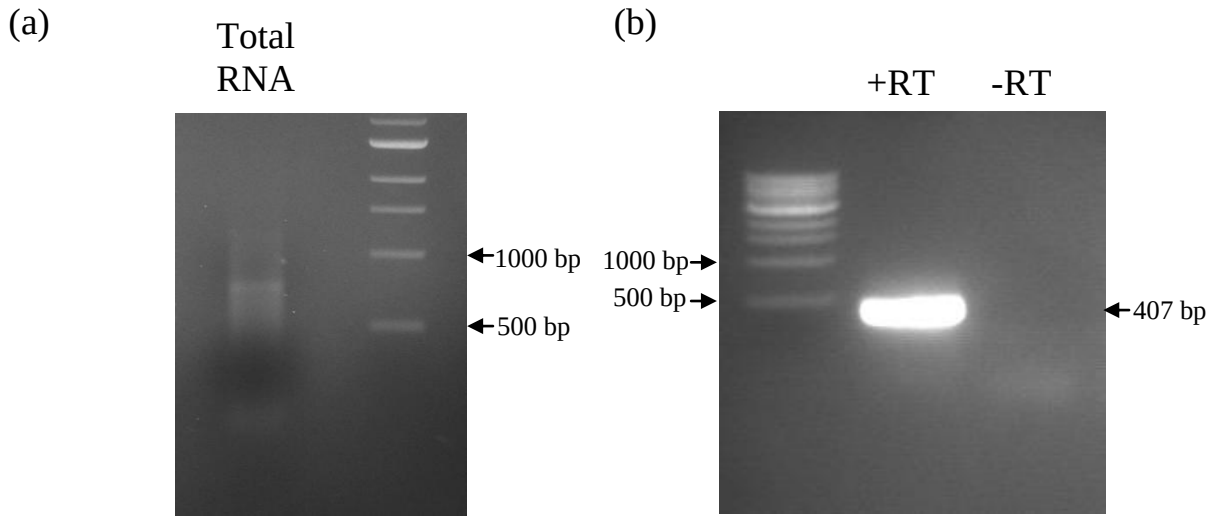


Figure 4.2. Extracted OE Total RNA and B-actin Control on Synthesized OE cDNA

3 µl of Extracted OE Total RNA run on 1% agarose gel (a). Control PCR with B-actin-Forward and B-actin-Reverse giving a 407 bp band from +RT but no band from –RT (b).

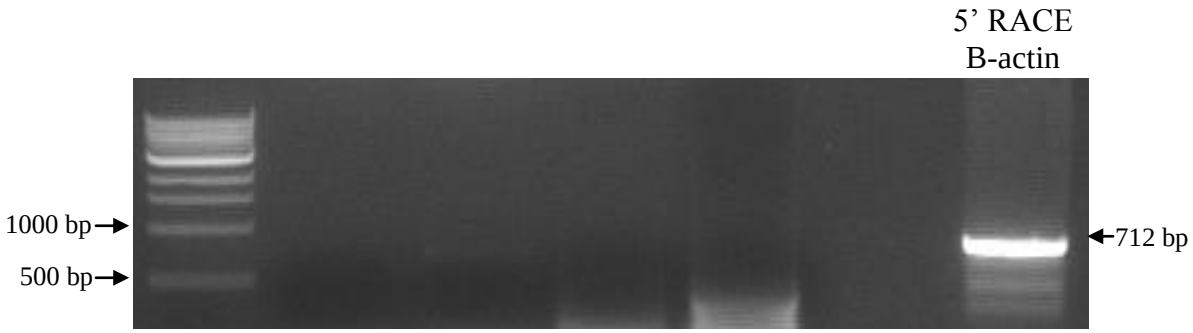


Figure 4.3. B-actin Control on Synthesized OE RLM RACE 5’ cDNA

1 µl of OE RLM-RACE 5’ cDNA was amplified with Outer Primer and B-actin-Reverse (right lane). The amplified band is the same size as expected (712 bp).

4.1.1. Rapid amplification of cDNA ends (RACE)

4.1.1.1. OR103-1. First we wanted to analyze the cDNA structure of the OR103-1, OR103-5, and OR103-2 genes and to map their transcription start sites. This analysis would be informative in the sense that it would reveal whether independent transcripts of the OR103-1 and OR103-5 genes exist. RLM-RACE 5’ cDNA was synthesized using the FirstChoice® RLM RACE Kit (Ambion). To obtain 5’ UTR information of OR103-1, two gene specific RACE primers, 103-1RACE1 and 103-1RACE2 and the Outer and Inner Primers matching to the ligated adapter sequences were used. In a first round of amplification, 1 µl of the RLM-RACE 5’ cDNA was amplified by PCR with 35 cycles using a combination of the Outer Primer and the 103-1RACE2 primer. In a second round of amplification by nested PCR, 1 µl of the product was amplified for 35 cycles with a combination of the Inner Primer and the 103-1RACE1 primer. The result of the amplification is shown in Figure 4.4a. Only a single band with a size of 168 bp was detected after nested PCR. To verify that the obtained PCR product corresponds to the OR103-1 transcript, the band was gel extracted and sequenced. The sequencing results showed that the amplified product matched to OR103-1 gene and included the OR103-1 coding sequence upstream of 103-1RACE1 primer (including the primer sequence itself), genomic sequence upstream of the OR103-1 gene, and the sequence of the Inner primer that was used for nested PCR in the RLM-RACE procedure. Thus, it is likely, that the obtained genomic sequence upstream of the OR103-1 gene corresponds to the 5’-

untranslated region (5'-UTR) of the OR103-1 transcript. The sequence was aligned against the recent version of the zebrafish genome (build Zv9) to map the transcription start site of the gene and its intron-exons structure. The corresponding structure of the 5'-part of the OR103-1 gene is schematically shown on Figure 4.4c. The OR 103-1 transcript consists of 20 bp of a non-coding 5'-exon, a 227 bp intron and a second exon that contains 13 bp immediately upstream of the OR103-1 coding sequence. Thus, the transcription start site of the gene is located 260 bp upstream of the OR103-1 coding sequence.

4.1.1.2. OR103-5. A similar approach was used to obtain 5' UTR information and the gene structure of the OR103-5 gene. In a first round of amplification, 1 µl of the RLM-RACE 5' cDNA was amplified for 35 cycles using the Outer Primer and 103-5-2-R primers. In a nested PCR, 1 µl of this product was used for 35 cycles of amplification with Inner Primer and 103-5RACE primers (Figure 4.4b left lane) or Inner Primer and 103-5-5UTR-R (right lane). The amplification with Inner Primer and 103-5RACE gave a total of 3 different products with sizes of 347 bp, 206 bp, and ~100 bp, whereas the use of the Inner Primer and 103-5-5UTR-R combination resulted in a single strong band of 69 bp. The 347 and 206 bp and 69 bp bands were subjected to DNA sequencing analysis.

The 206 bp and 69 bp sequenced bands corresponded to the OR103-5 gene. The resulting sequence information was matched to genomic sequence of the OR103-5 locus and revealed the presence of a 33 bp 5'-noncoding exon, a 118 bp intron and a second coding exon containing 38 bp of untranslated sequence in front of the 942 bp OR 103-5 coding sequence. A putative transcription start site of the OR103-5 gene is thus located 189 bp upstream of the gene's coding sequence.

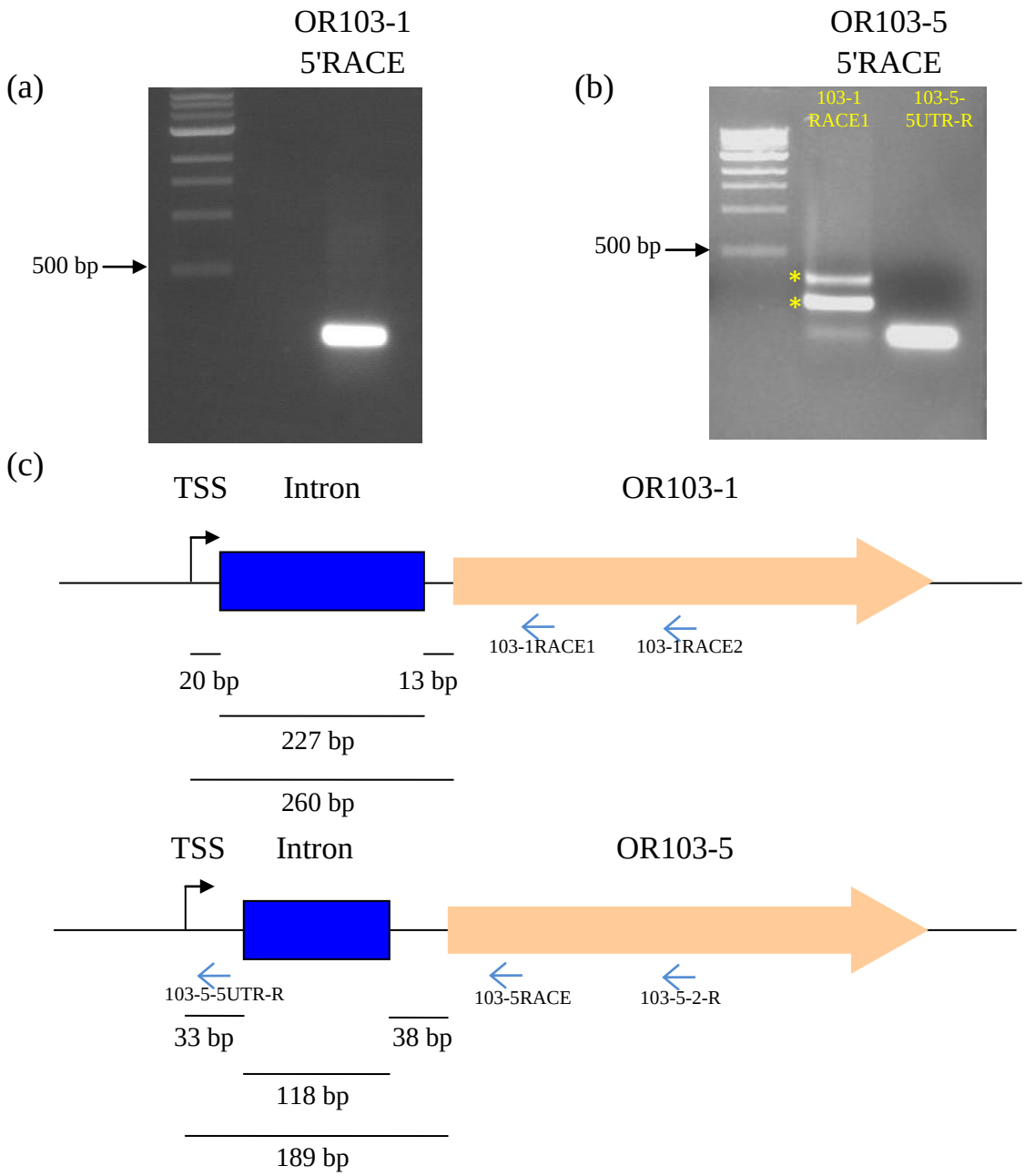


Figure 4.4. 5'UTR Information of OR103-1 and OR103-5

(a) RLM RACE result for 5'RACE of OR103-1 showing a 168 bp band. The band was obtained after 35 cycles of amplification with Outer Primer and 103-1RACE2 and nested amplification of the 1µl product with Inner Primer and 103-1RACE1. The band has been sequenced and shown to match OR103-1. (b) RLM RACE result for 5'RACE of OR103-5. Left lane is showing 35 cycles of amplification product of Inner Primer and 103-5RACE on 1 µl of 35 cycles of amplification with Outer Primer and 103-5-2-R on OE RLM 5' cDNA. Bands marked with asterisks have been sequenced. The lower band with the asterisk matched to OR103-5. Right lane is showing 35 cycles of amplification with Inner Primer and 103-5-5UTR-R on 1 µl of the product obtained by 35 cycles of amplification with Outer Primer and 103-5-2-R on OE RLM RACE 5' cDNA. Band obtained has been sequenced and shown to match OR103-5. (c) Diagram showing the 5' Transcript information for OR103-1 and OR103-5 obtained from the sequencing of the matching bands in (b). Arrows show positions of primers.

4.1.1.3. OR103-2. The 347 bp upper band in Figure 4.4b was matched to the genomic sequence of OR103-2 locus and revealed a single 174 bp 5'-noncoding exon directly upstream of the 945 bp coding sequence (Figure 4.5).

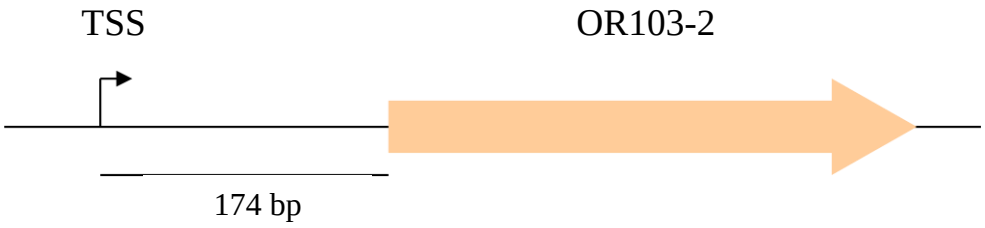


Figure 4.5. 5'UTR Information of OR103-2

Sequencing of the upper band marked with asterisk in Figure. 4.4b revealed a single 174 bp 5' nontranslated exon immediately upstream of OR103-2 coding sequence.

The experiments show that all three OR genes, OR103-1, OR103-5 and OR103-2, might be expressed from independent promoters and the RACE procedure did not reveal any products that would correspond to a single large transcript that contains the OR103-1 and OR103-5 coding sequences. However, such a product would be approximately 2.5 kb in size and the failure to detect such a band could be due to the specific PCR conditions that were used.

4.1.2. RT-PCR

Altough the RACE experiments did not reveal any experimental evidence that the OR103-1 and OR103-5 genes could be located on a single large transcript, we wanted to put this hypothesis to a direct test using RT-PCR to identify the existence of potentially longer transcripts. For this purpose we designed specific oligonucleotide primers that are located within the OR103-1 and OR103-5 coding sequences and would amplify a product that spans the intergenic region between the two genes.

Using OE-derived cDNA as a template and negative control samples in which the reverse transcriptase has been omitted during generation of the cDNA, we obtained two bands from the experimental samples and a faint band from the –RT control sample, probably due to some genomic contamination (Figure 4.6). The sizes of the amplified bands were 1906 bp and 1788 bp for the experimental sample and 1906 bp for the –RT sample. The size of 1906 bp that was seen in both samples (experimental and –RT) perfectly matches the distance of the two primers on the genomic sequence of the OR103-1 / OR103-5 locus. Interestingly, the smaller band that was observed for the experimental sample perfectly corresponds to a product that contains the region between the primer pair without the intronic sequence that was observed in the OR103-5 RACE analysis. The two bands that were obtained by RT-PCR in the experimental sample were not of equal intensity, suggesting that the 5'-intron of the OR103-5 gene is subject to alternative splicing in this transcript. It should be emphasized that a weak band of the same size was also obtained from the –RT control sample. However, the corresponding band in the experimental sample was much stronger, suggesting that the product that was obtained from the experimental sample was amplified from cDNA sequences and is not entirely due to genomic contamination.

To confirm the identity of the product, the two bands amplified in RT have been directly ligated to pGEMT-easy (Promega) by T/A cloning, transformed to competent cells, screened for presence of PCR products by colony PCR (Figure 4.7), isolated and sequenced.

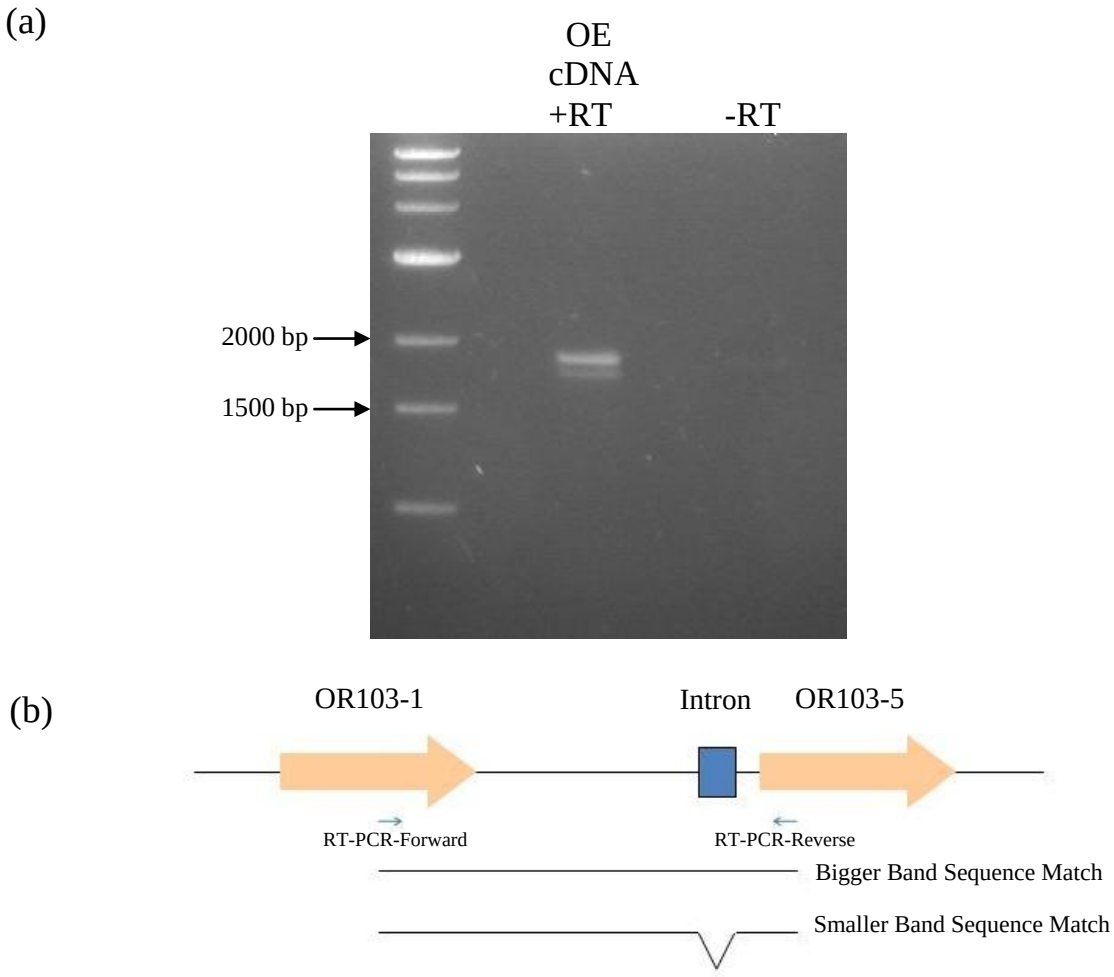


Figure 4.6. RT-PCR on OE cDNA to Test For a Single mRNA with OR103-1 and OR103-5

(a) RT-PCR giving a 1906 bp and a 1788 bp band with RT-PCR-forward and RT-PCR-reverse. -RT control only giving a very faint band due to minimal genomic DNA contamination. Both bands have been ligated to pGEMT-easy (Promega), transformed to competent cells and sequenced with T7-highTM and M13R-highTM primers. (b) Sequencing of the bands reveal that they match to the expected products; the bigger band (1906 bp) matching to the unspliced version of single transcript and the smaller band (1788 bp) matching to the spliced version of the single transcript. Arrows indicate positions of primers.

Sequencing results revealed that the upper band matched to the 1906 bp region between the primers, including the coding sequence of OR103-1 downstream of RT-PCR-forward primer (including the primer sequence), the 1432 bp intergenic region, and the part of coding sequence of OR103-5 upstream of RT-PCR-reverse (including the primer sequence), respectively (Figure 4.6b). The lower band matched to the same region with the

exception of a 118 bp intron 38 bp upstream of the start of OR103-5 coding sequence that was also identified from the RACE analysis. Thus, the presence of a spliced version of the long transcript is convincing proof that the results obtained by RT-PCR were not obtained from genomic contamination of the experimental sample. Genomic DNA is not subject to RNA splicing and consequently only a single band corresponding to the unspliced variant was seen.

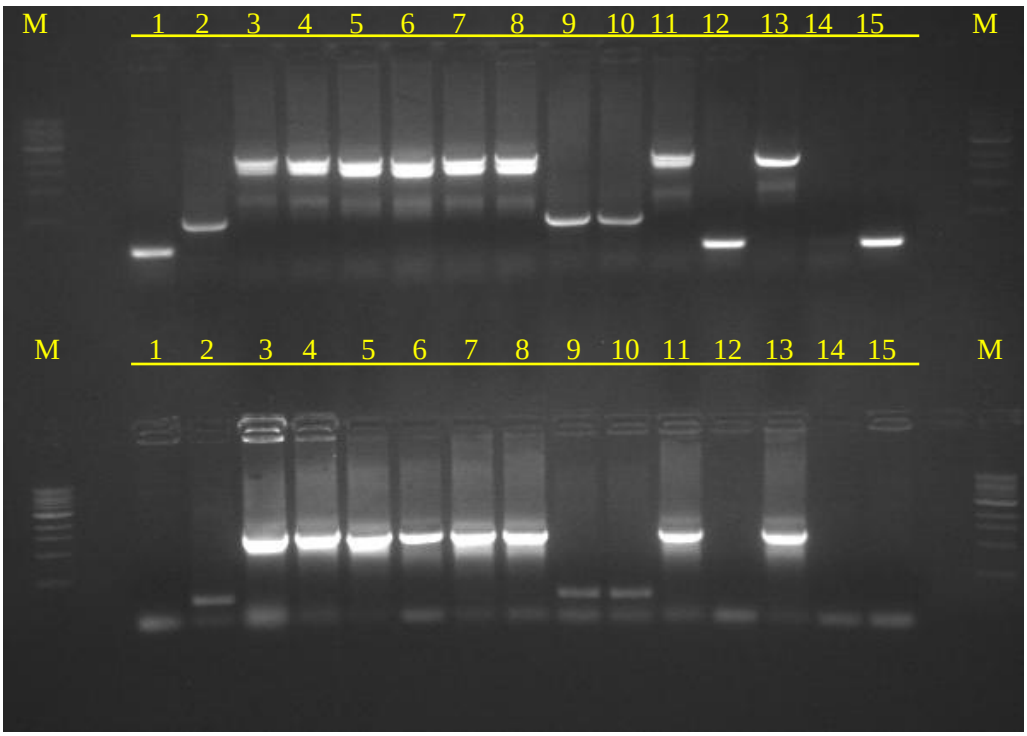


Figure 4.7. Cloned RT-PCR Products Identified By Colony PCR

Products from +RT cloned to pGEMT-easy (Promega) and colonies were screened for the presence of two different RT bands. Upper row: Colony PCR with T7-highTM and M13-R-HighTM. Bottom Row: Colony PCR with 103-1-5-Int-F and RT-PCR-Reverse. Colonies 3-4-5-6-7-8-11-13 gave expected sizes, indicating presence of a PCR product from RT. Colonies 3-5-6 have been sequenced. Colony 3 matched to unspliced version of the single transcript, whereas, 5-6 matched to spliced version. Lanes marked with M have been loaded with 1kb marker.

The results obtained for the RT-PCR experiments indicate that a single transcript might exist that contains both, OR103-1 and OR103-5 coding sequences. In addition, two different variants were detected, one with the unspliced OR103-5 5'-intron and one with the intron spliced. This is important because it's a strong proof that OR103-1 and OR103-5

coding sequences could be part of a single large transcript. Thus, a single transcript could be underlying reason for the co-expression previously observed (Sato *et al.,* 2007).

4.1.3. Gene-specific RACE

A shortcoming of the previous analysis by RT-PCR and RLM-RACE is that the oligonucleotide primers used for PCR were located in positions flanking the intergenic region and that the full 5'-extent of the long transcript could not be revealed. There was some evidence from sequences submitted to NCBI GenBank that the 3'-UTR and 5'-UTR of the OR103-1 and OR103-5 genes might overlap. Thus, theoretically, the observed long transcript could have been generated by overlap PCR during the process of RT-PCR on the complex OE-derived cDNA sample. To rule out this possibility, we designed a strategy to obtain independent proof of the existence of the long transcript using gene-specific RACE.

In this procedure, we synthesized RLM-RACE 5'-cDNA using a gene specific primer that binds specifically to the OR103-5 gene instead of the commonly used oligo-dT primer. The resulting cDNA only represents transcripts to which the 103-5-2-R can bind. Because there are 14 bp mismatches between OR103-1 coding sequence and the 103-5-2-R primer, it is not expected to bind to the OR103-1 coding sequence during cDNA synthesis.

However, if OR103-1 and OR103-5 are located on a single large transcript, it should be possible to amplify a specific RACE product using OR103-1 RACE primers on cDNA that has been synthesized with the OR103-5 specific oligonucleotide. Experimentally, 1 µl of OR103-5-specific 5'-RACE cDNA was used in the same conditions as described for the amplification of OR103-1 5'-UTR (Figure 4.4) with nested PCR with Inner Primer and 103-1RACE1 on the product of Outer Primer and 103-1RACE2, and the results are shown in Figure 4.8. A single strong band of the same size of 168 bp was amplified, similar to the results obtained for the OR103-1 RACE experiment. The band was subjected to direct DNA sequencing and the identical 5'-UTR structure was obtained for comprising 20 bp of 5'-noncoding exon, 227 bp of intronic sequence, and 13 bp of non-translated sequence upstream of the OR103-1 translation start codon. Therefore,

these results provide additional proof for the presence of a single transcript with OR103-1 and OR103-5 coding sequences.

4.2. IRES Activity of the Intergenic Region in *vivo*

The RT-PCR and RLM RACE experiments have shown that a single transcript might exist that encompasses the OR103-1 and OR103-5 coding sequences. However, the presence of a single transcript does not necessarily suggest that both coding sequences are translated into proteins. For both ORs on the single transcript to be translated into proteins, there has to be a sequence that would allow for independent initiation of translation. Such sequences are known as Internal Ribosome Entry Site (IRES) that are located upstream of the second coding sequence of bicistronic transcripts which allow independent assembly of a ribosomal complex (Pelletier and Sonenberg, 1988; Jang *et al.,* 1989). In order to understand if both genes on this single mRNA shown above are translated into proteins, specific antibodies would be required but were not available. Therefore, we designed transgenic experiments in zebrafish to demonstrate whether the intergenic region between OR103-1 and OR103-5 could serve as an IRES by driving co-translation of two fluorescent proteins in OSNs. We prepared a construct pOMP::CIG (pOMP::mCherry-OR103-1 / 5 Intergenic Region-GFP-pA) that consists of the olfactory marker protein gene promoter (pOMP), a sequence coding for the red fluorescent protein mCherry, the OR103-1 / OR103-5 intergenic region, a sequence coding for the green fluorescent protein GFP and a SV40 polyadenylation signal, which is shown in Figure 4.9a, along with a control construct pOMP::C (pOMP::mCherry), which contains the OMP gene promoter, a sequence coding for mCherry and the SV40 polyadenylation signal. The OMP gene promoter has previously been shown to drive tissue-specific expression in the OE (Celik *et al.,* 2002; Sato *et al*., 2005). The successful cloning of the constructs was verified by analytical digests with restriction endonucleases (Figure 4.9b) and direct DNA sequencing. The pOMP:CIG construct should, in theory, recapitulate the behavior of the OR103-1 / OR103-5 locus. If the intergenic region between the two OR genes has any intrinsic IRES activity, transgene-expressing cells showing mCherry expression should also be double-positive for GFP fluorescence. Both constructs, pOMP::C and pOMP::CIG have been injected into single-

cell zebrafish embryos and reporter gene expression has been examined by fluorescence 48 hours after fertilization.

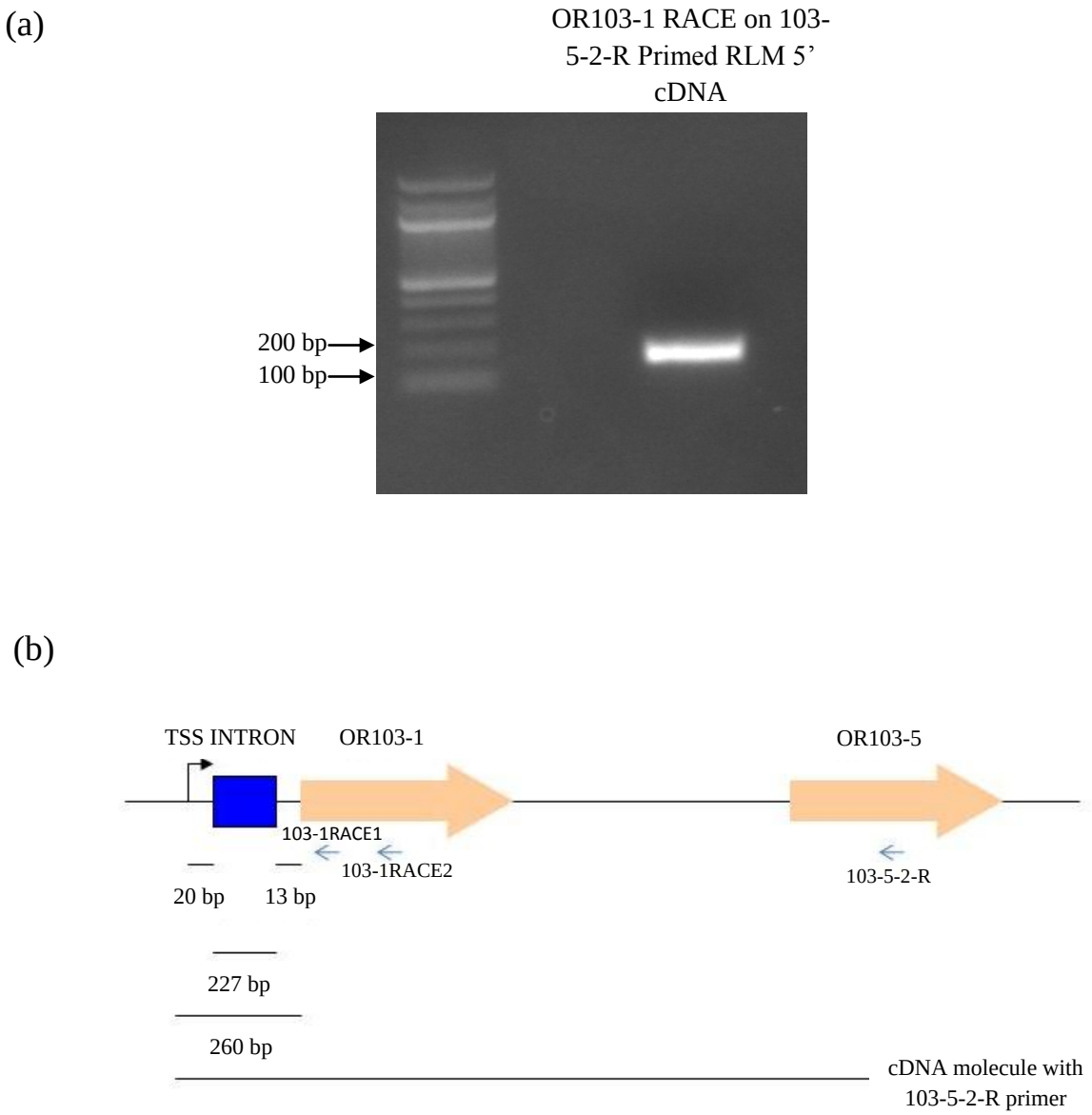


Figure 4.8. OR103-1 5'RACE on 103-5-2 Primed RLM-RACE 5' cDNA

RLM-RACE 5'RACE cDNA was prepared using 103-5-2-R primer instead of oligo$(DT)_{12-18}$. The band shown in (a) is the result of 35 cycles of amplification with Inner Primer and 103-1RACE1 on 1 µl of the 35 cycles amplification product with Outer Primer and 103-1RACE2. Marker: 100 bp DNA Ladder (NEB). The band was sequenced and the 5' UTR information of the transcript obtained is shown (b). Arrows show positions of primers.

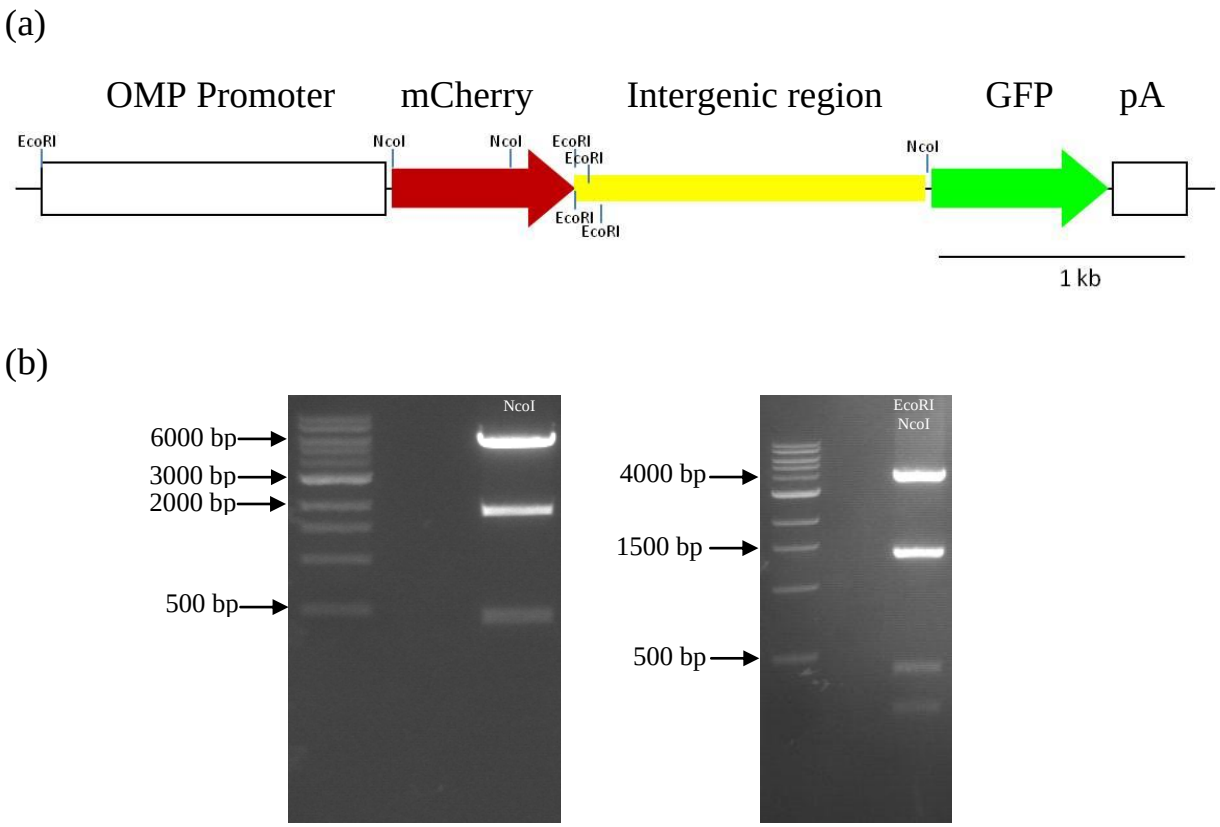


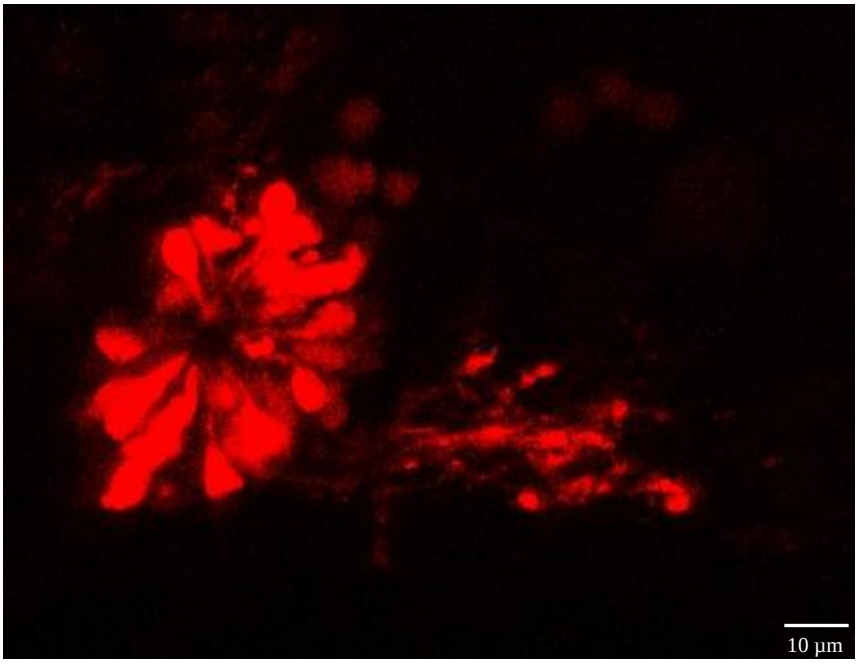
Figure 4.9. Construct To Test IRES Activity of Intergenic Region in OSNs

(a) DNA construct prepared for observation of possible IRES activity of 1432 bp intergenic region between OR103-1 and OR103-5. OMP promoter was coupled to mCherry coding sequence, intergenic region, GFP coding sequence followed by a polyadenylation signal. (b) NcoI analytical digest showing expected pattern (left) and EcoRI + NcoI analytical digest showing expected pattern (right). This construct was injected into one cell zebrafish embryos.

4.2.1. Expression of the pOMP::C Transgene

The OMP promoter transgene, as expected, resulted in specific mCherry expression in OSNs in the nose (Figure 4.10a). Clearly, large numbers of transgene expressing OSNs and their axonal projections to the olfactory bulb could be detected. Large numbers of OSNs labeled result in a display of the overall shape of OE and Selected OSNs have been labeled with arrows (Figure 4.10b). Multiple protoglomeruli in the OB are visible by axons converging onto each other, which will form functional glomeruli over the first eight days of development (Dynes and Ngai, 1998).

(a)



(b)

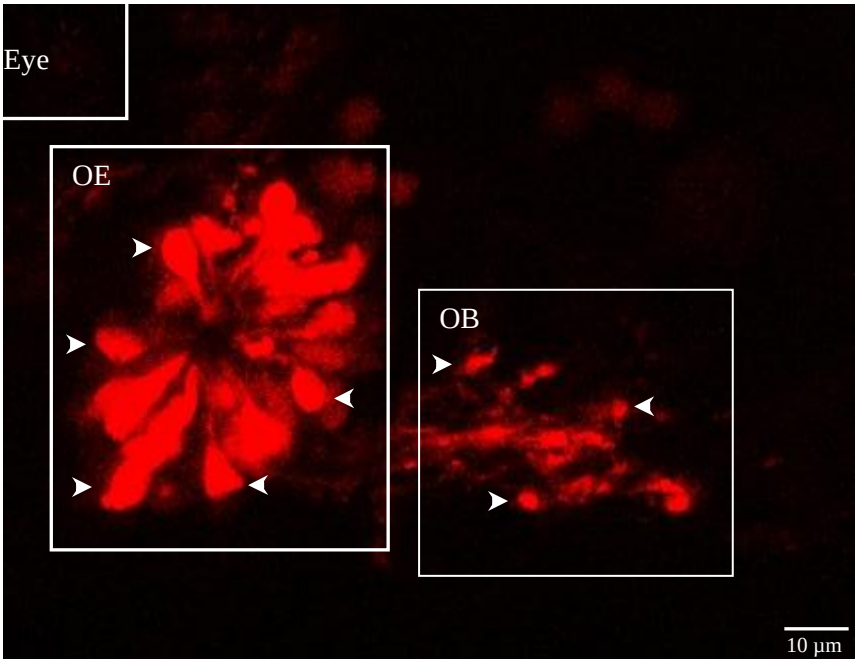


Figure 4.10. Expression of mCherry In Mature OSNs By The OMP Promoter

Numerous mature OSNs expressing mCherry in one OE of 2 day old pOMP::C injected embryo. (a) Unlabeled image taken with confocal microscope. (b) Boxes showing positions of the right eye, OE and OB. Arrows in OE show selected OSNs and arrows in OB show selected protoglomeruli.

4.2.2. Expression of the pOMP::CIG Transgene

Next, we injected the pOMP::CIG construct to test whether the OR103-1 / OR103-5 intergenic region contains any IRES activity in OSNs. If the intergenic region would contain any sequences that promote co-translation of two proteins, we would be able to observe OSNs that are double positive for mCherry and GFP expression. Among 60 embryos that were injected with the pOMP::CIG construct 30% survived 2 days post fertilization. Of the 18 surviving embryos, 2 embryos were positive for the transgene and could be identified by their mCherry fluorescence (Table 4.1). Interestingly, all mCherry-expressing OSNs were also positive for GFP (Figure 4.11a and b). Two epithelia from two different positive fish are shown in Figures 4.11a and 4.11b. The eyes show flourescence due to non-specific reflection of fluorescent light from iriodiophores, but OSNs that are double fluorescent for mCherry and GFP can clearly be distinguished. Out of 29 cells visualized, 28 were certainly positive for both GFP and mCherry, one cell could be positive for only GFP or both, which could not be determined decisively by confocal imaging. This shows a co-translation rate of 97-100%, strongly suggesting that the intergenic region contains IRES activity.

As observed for most labeled OSNs, the intensities of mCherry and GFP are consistent, with bright red-fluorescent cells also being bright green-fluorescent and cells that show fainter mCherry fluorescence also showing weak GFP fluorescence. However, we could also observe OSNs that displayed weak or absent mCherry but strong GFP signals. It might be that those cells have utilized the internal OR103-5 promoter -which we identified to be located within the intergenic region- in addition to the OMP promoter. The RLM-RACE identified a short OR103-5 transcript with a transcription start site 189 bp upstream of the OR103-5 coding sequence, thus the OR103-5 gene most likely possesses an independent promoter capable of driving a transcription that only contains the OR103-5 sequence. This sequence is located within the intergenic region and in the case of the pomp::CIG transgene, can function as an independent promoter as well, driving expression of GFP only in some OSNs of the injected fish. In those cells, more or exclusive GFP-fluorescence would be observed as compared to mCherry fluorescence. However, preliminary experiments using a transgenic construct that only contains the intergenic

region fused to a sequence coding for GFP does not promote GFP expression in a high number of OSNs (0 out of 42 surviving embryos of 65 injected). Thus, it might be that the most of the GFP-fluorescence that was observed from pOMP::CIG injected embryos arises from the activity of the OMP gene promoter. Most likely the transgenic construct gives rise to a single long transcript harboring both, the mCherry and GFP coding sequences, and the translation of GFP originates from sequences contained within the intergenic region. The observed independent promoter activity of the intergenic region in OSNs does not contradict the property of the intergenic region to contain IRES activity.

Using a synthetic transgenic construct, pOMP::CIG, we were able to recapitulate the co-expression phenotype that is observed for the OR103-1 / OR103-5 gene locus. By using a different but OSN-specific promoter, we were able to increase the number of transgene expressing OSNs. As observed for the OR103-1 and OR103-5 genes, the two sequences coding for different fluorescent proteins that were separated by the intergenic region showed an almost complete level of correlation in their expression. All OSNs that were positive for expression of the first cistron (mCherry) were positive for the second cistron (GFP) as well. In addition, just like observed for the OR103-1 and OR103-5 genes, a low proportion of OSNs were positive for the second coding sequence only. The high level of correlation between of mCherry and GFP signals suggests a strong IRES activity is present in the intergenic region between OR103-1 and OR103-5 in addition to independent promoter activity that drives expression of the second coding sequence.

Table 4.1. Injection Statistics for the pOMP::CIG Construct

Number of Injected	Background sd	1st Day Survival	2nd Day Survival sada	3rd Day Survival	Number of Positive Fish Observed
60	AB/AB	19	18	18	2

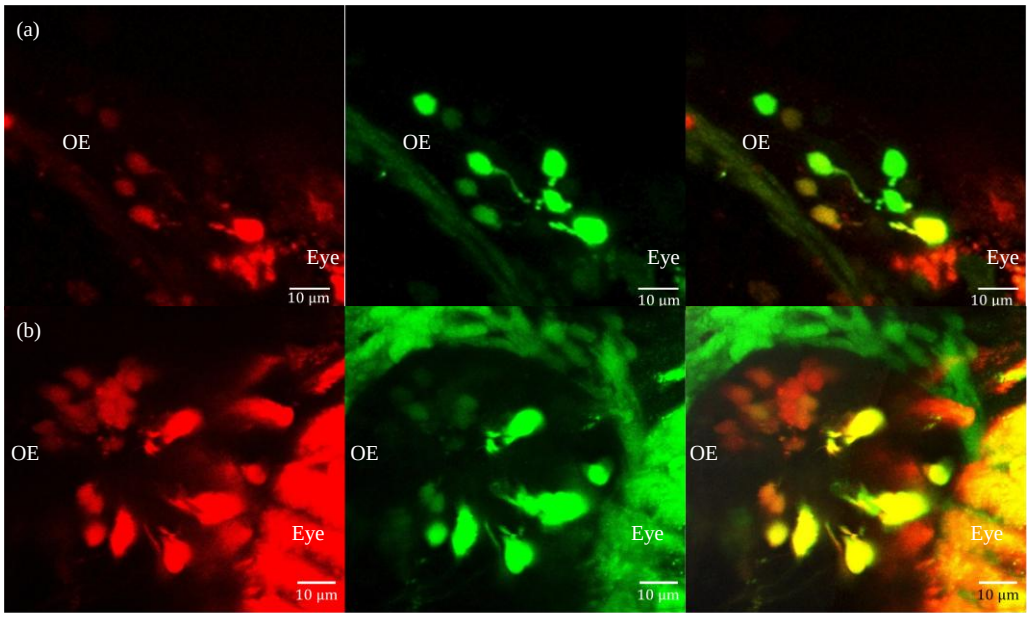


Figure 4.11. Co-expression of mCherry and GFP in OSNs

(a,b) Flourescence observed in OE of injected embryos, displaying co-expression of mCherry and GFP. The images were taken with Leica Confocal Microscope on the 2nd day after fertilization. (a) shows one OE from one positive fish, (b) another OE from another positive fish.

4.3. IRES Activity of the Intergenic Region in Other Tissues of the Zebrafish

Because we could show that the OR103-1 / OR103-5 intergenic region contains an independent transcription start site, which may contain an active OR103-5 promoter, these findings cast some doubt whether the observed co-expression of mCherry and GFP from the transgenic construct could be entirely due to the intrinsic promoter activity. Therefore, we wanted to test whether the intergenic region would promote co-expression of two coding sequences outside OSNs. This would be informative for two reasons. First, because the independent OR103-5 promoter is unlikely to be active outside the OE, a positive result would strengthen the hypothesis that the intergenic region acts as an IRES sequence, and second, it would show with some limitation whether the IRES is tissue specific.

To drive expression of a transgene outside the olfactory tissue we utilized the previously characterized growth-associated protein-43 (GAP-43) promoter, which is specific for developing neurons. A 1 kb rat sequence upstream of the GAP-43 gene was shown to recapitulate GAP-43 expression in zebrafish and drives expression in the brain and spinal cord (Reinhard *et al.*, 1994). The transgenic construct pGAP-43::CIG contains a sequence coding for mCherry, the full intergenic region between OR103-1 and OR103-5, a sequence coding for GFP and an SV40 polyadenylation signal (Figure 4.12). This construct is expected to drive transgene expression in developing neurons of the zebrafish. If the intergenic region has IRES activity outside the OE and in neurons other than OSNs, this construct is expected to result in neurons that are double positive for mCherry and GFP fluorescence.

Out of the 25 embryos that were injected, 2 survived (8%) to 3 days post fertilization (Table 4.2). Both of the two surviving embryos were positive for the transgene as observed with confocal microscope (Figure 4.13). Similar to the results obtained for the pOMP::CIG transgene, co-expression of mCherry and GFP was observed in all transgene expressing cells. The transgene expressing cells were located in the tail of embryos. As judged by this expression pattern and morphology, the transgene expressing cells were most likely notochord cells. A total of 6 double positive cells were detected 2 days post

fertilization. Unlike the transgene expression of the pOMP::CIG construct in the OE, no cells single positive for GFP-fluorescence were observed for the pGAP-43::CIG construct, supporting the hypothesis that the intrinsic OR103-5 promoter is not active outside the OE.

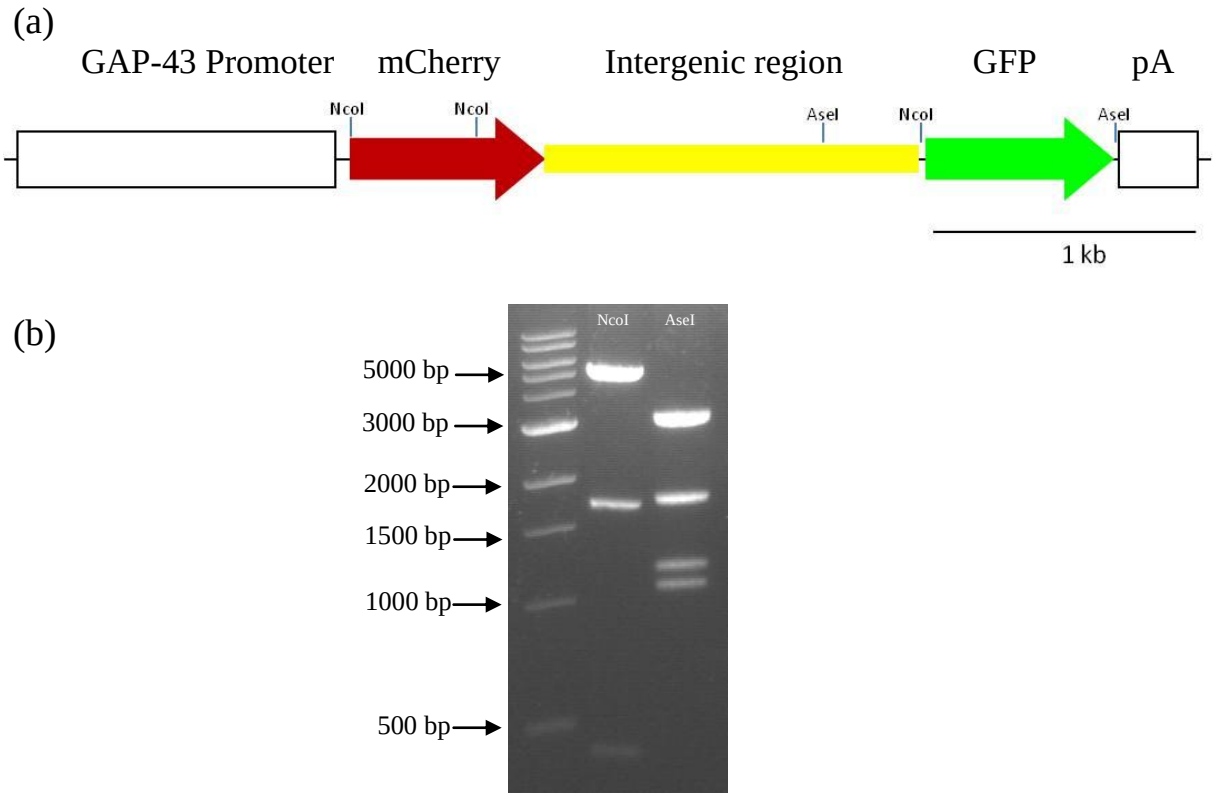


Figure 4.12. Construct To Test IRES Activity of The Intergenic Region In Zebrafish Neurons

(a) DNA construct prepared for observing possible IRES activity of 1432 bp intergenic region between OR103-1 and OR103-5 in nerve cells of the zebrafish. This construct is similar to the construct in Figure 4.10 except for the promoter present. In this construct, Rat GAP43 has been placed instead of OMP promoter. (b) NcoI and AseI analytical digests of the construct in (a). Bands match expected sizes.

The expression patern of the pGAP-43::CIG construct and the 100% co-expression between mCherry and GFP in notochord cells provides additional evidence that the intergenic region between OR103-1 and OR 103-5 possesses IRES activity. Furthermore, it also shows that the IRES activity is not restricted to OSNs, but can be observed in other neurons of the zebrafish.

Table 4.2. Injection Statistics for the pGAP-43::CIG Construct

Number of Injected	Background	1st Day Survival	2nd Day Survival	3rd Day Survival	Number of Positive Fish Observed
25	Tü/AB	3	2	2	2

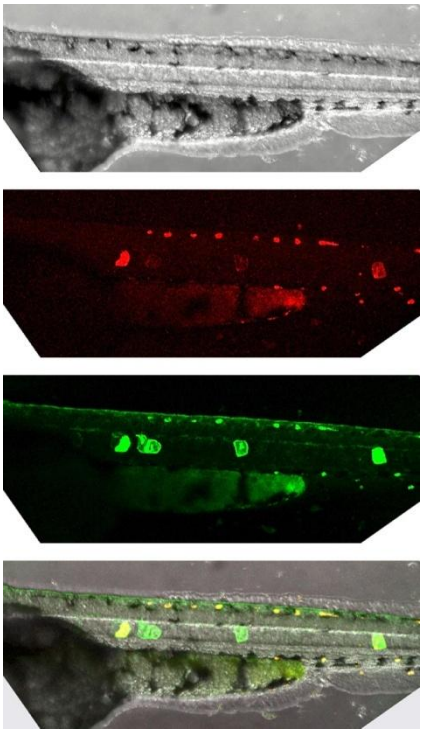


Figure 4.13. Co-expression of mCherry and GFP in Notochord of the Zebrafish

Flourescence observed in the notochord in pGAP-43::CIG injected embryos, displaying co-expression of mCherry and GFP. The images were taken with Leica Confocal Microscope. Co-expressing neurons were imaged on the 2nd after fertilization

4.4. IRES Activity of the Intergenic Region in vitro

To obtain quantitative data on the IRES acitivity of the intergenic region between OR103-1 and OR103-5 and to test the IRES activity of the intergenic region outside the zebrafish model, a dual luciferase reporter assay in cultured Hk-2 cells was performed. The Dual-Luciferase® Reporter Assay (Promega), which allows a quantitative measurement of renilla and firefly luciferase activity separately, was used. A variety of different constructs based on the basic pRF vector were constructed as shown in Figure 4.14. The basic pRF vector comprises the SV40 early promoter, a chimeric intron, sequences coding for renilla and firefly luciferase genes, followed by an SV40 polyadenylation signal and the SV40 enhancer (Coldwell *et al.*, 2001). The intergenic region between OR103-1 and OR103-5 has been cloned into the empty pRF vector's multiple cloning site inbetween the renilla and firefly luciferase genes (Figure 4.14d) and is referred to as pR-inter-F. A control construct containing the well characterized IRES sequence of the encephalomyocarditis virus was also used for comparison (pR-EMCV-F). To demonstrate that independent translation occurs for the second cistron, another construct was engineered that contains a sequence that forms a hairpin to block translation of the first cistron that codes for renilla luciferase and is referred to as hpR-inter-F (Coldwell *et al.*, 2001).

It is generally observed that transfection of the empty control vector pRF into cultured cells will result in high levels of renilla luciferase activity, but no or severly reduced firefly luciferase activity due to the lack of translation of the second cistron. If the intergenic region of OR103-1 and OR103-5 has IRES activity in Hk-2 cells, it's expected that high levels of firefly luciferase activity, in addition to renilla luciferse activity, can be observed when the pR-inter-F plasmid is transfected into cells. It's also expected that the hp-R-inter-F will give comparable levels of firefly luciferase activity when compared to pR-inter-F, but no or severly reduced renilla luciferase activity will be observed due to blockade of ribosomal scanning by the hairpin structure.

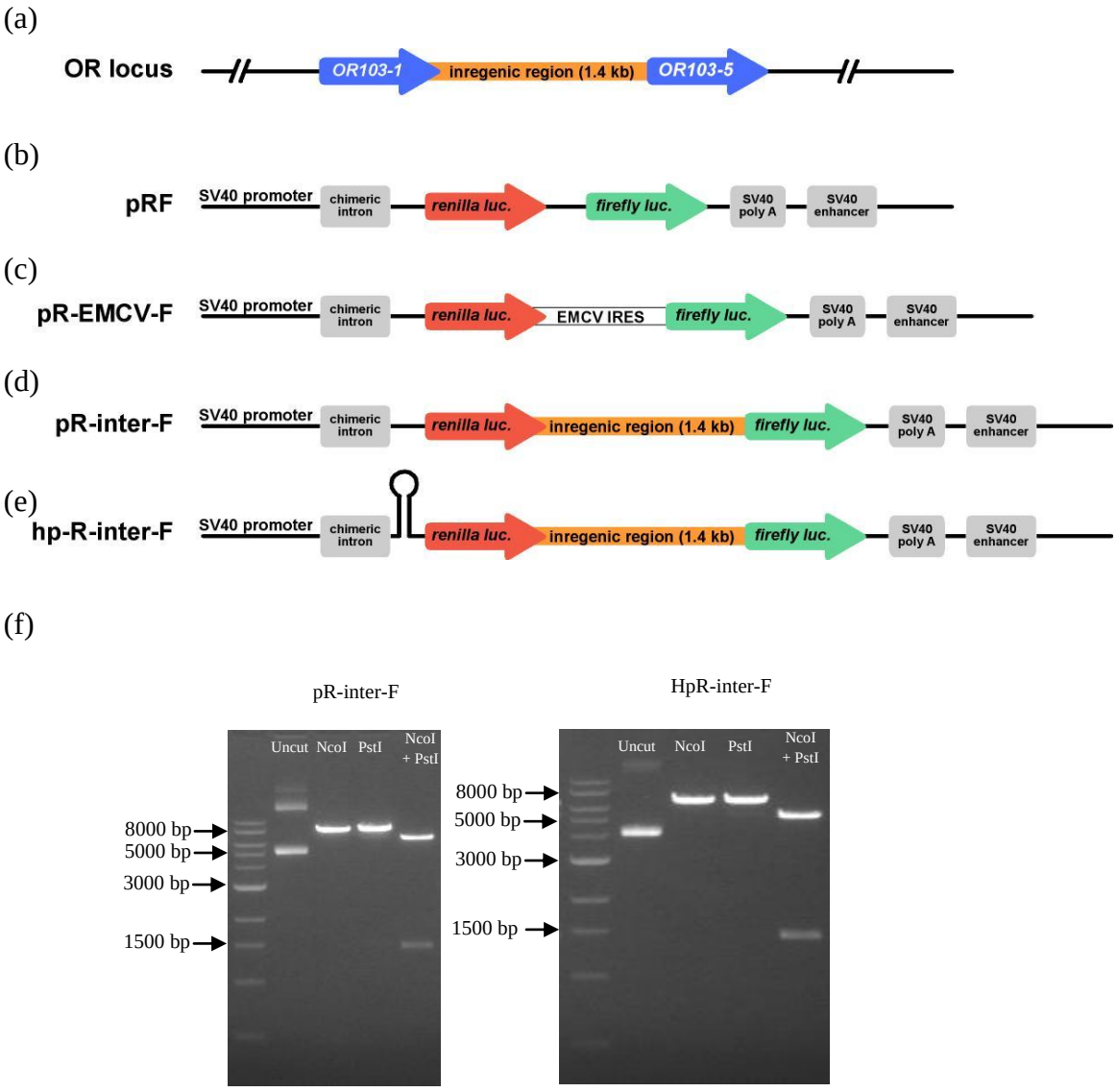


Figure 4.14. Constructs For Testing The Possible IRES Activity of The Intergenic Region Between OR103-1 and OR103-5 In Mammalian Cell Culture.

(a) OR 103-1 and OR103-5 loci. (b) OR Empty pRF vector containing SV-40 Promoter, Renilla Luciferase and Firefly Luciferase. (c) EMCV IRES in pRF vector. (d) 1432 bp intergenic region between OR103-1 and OR103-5 cloned into pRF between the two luciferases. (e) Hairpin after the SV-40 promoter to block ribosome scanning. (f) Analytical digests showing expected patterns.

All constructs were transfected into ~ 70% confluent Hk-2 cells in 3 replicates per experiment. Four independent experiments were performed for the comparison of pRF and pR-inter-F, and 3 independent experiments were performed for the comparison of pR-inter-F and hpR-inter-F. 100 ng (10:1) control plasmid was co-transfected to measure relative transfection efficiency and normalize results in experiments 3 & 4. 24 hours after transfection, cells were lysed using Passive Lysis Buffer and luciferase activity assays for renilla and firefly were performed (Table 4.3). As expected, pRF gave high levels of renilla but low levels of firefly luciferase activity. In the control plasmid containing the EMCV IRES, firefly luciferase activity raised about 23-fold (23.66 ± 4.35) in Hk-2 cells (Figure 4.15a). Interestingly, the OR103-1 / OR103-5 intergenic region increased the relative

Table 4.3. Luciferase Assay Data of Firefly-Renilla Constructs

Experiment Number	Number of Cells	Assay	pRF	pRF	pRF	pR-inter-F	pR-inter-F	pR-inter-F
1	50000	Firefly	180,7	252,9	282,5	57,95	242,2	113,4
		Renilla	5480	15300	16700	0	15,12	29,37
2	7500	Firefly	55,88	148,1	270,9	618,3	680,4	1105
		Renilla	121000	10200	9015	280	307	347,1
3		Firefly	1165	625,2	775,9	12900	17500	10400
		Renilla	19700	22000	22300	3203	5868	3314
		ß-Gal	8008	9202	10600	8392	9368	8052
4		Firefly	13500	13900	12300	169000	157000	163000
		Renilla	197000	200000	226000	18600	9693	20000
		ß-Gal	2413	2778	2790	1312	1723	1909
Experiment Number	Number of Cells	Assay	hp-R-inter-F	hp-R-inter-F	hp-R-inter-F	pR-EMCV-F	pR-EMCV-F	pR-EMCV-F
1	50000	Firefly	343,9	107,5	370,3	-	-	-
		Renilla	0	0	0	-	-	-
2	75000	Firefly	902,5	390,9	779,4	-	-	-
		Renilla	87,85	54,08	0	-	-	-
3		Firefly	-	-	-	10600	10200	14500
		Renilla	-	-	-	6818	12200	10900
		ß-Gal	-	-	-	19500	19200	21900
4		Firefly	94000	102000	101000	76200	70100	72900
		Renilla	69,04	99,47	43,42	105000	109000	69400
		ß-Gal	1527	2077	1825	4377	4197	4157

levels of firefly luciferase activity 172-fold (± 44.59) over the activity of renilla luciferase. Thus, the OR103-1 / OR103-5 intergenic region appears to have a strong capacity to increase firefly luciferase activity, which is in agreement with its function as an IRES element.

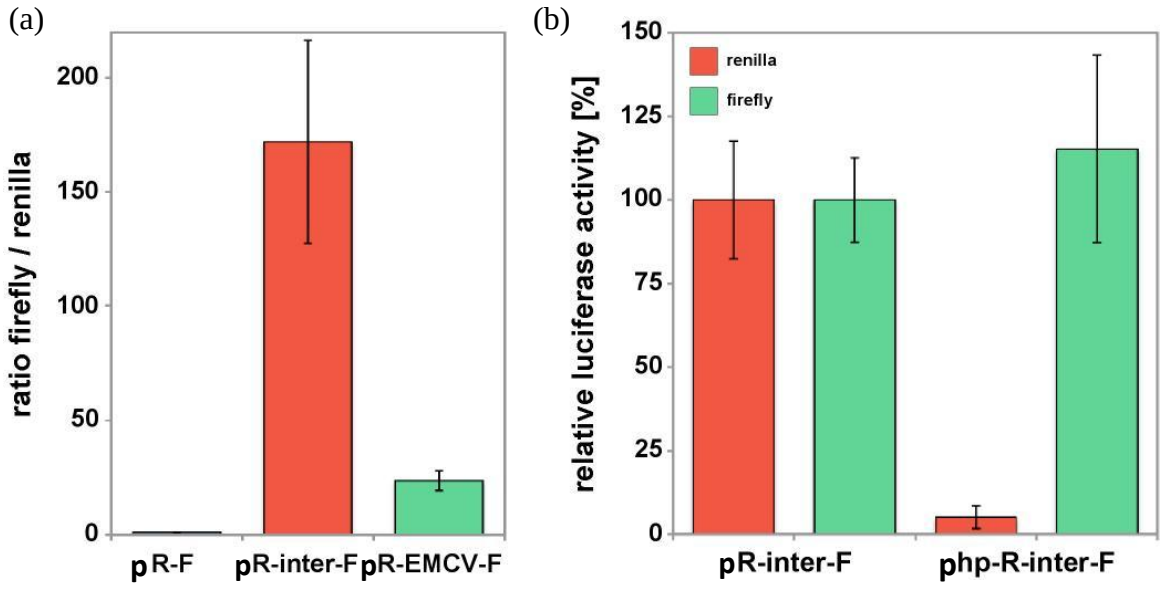


Figure 4.15. Firefly to Renilla Ratios and Relative Luciferase Activity In Transfected Cells.

(a) Firefly to renilla ratios of pRF, pR-inter-F and pR-EMCV-F. (b) Relative luciferase activities of pR-inter-F and hp-R-inter-F.

In the control expriment where translation of the renilla luciferase gene is blocked by a hairpin that interrupts scanning of the ribosome and assembly of a ribosomal translation complex, the activity of the gene encoded by the first cistron, renilla luciferase, was significantly reduced to 5.21% (± 3.44%) of the renilla luciferase activity in the pR-inter-F construct (Figure 4.15b). Activity of the firefly luciferase in the hp-R-inter-F construct remained virtually unchanged with 115.22% (± 28.03%), indicating that translation of the second cistron coding for firefly luciferase can initiate independently of the first cistron. This finding is again in agreement with IRES activity contained within the OR103-1 / OR103-5 intergenic region.

Thus, independent examination of the zebrafish OR103-1 / OR103-5 intergenic region in an in vitro cell culture system using standard assays performed to quantitatively study whether the sequence can promote translation of the second cistron from a bicistronic construct revealed that it might constitute or harbor an IRES element. The exact definition of the sequence within the OR103-1 / OR103-5 intergenic region that promotes this activity is not known at this point. It is also conceivable that the OR103-1 / OR103-5 intergenic region might possess strong promoter activity in Hk-2 cells. Experiments to discriminate between these possibilities are discussed below.

4.5. Summary of Experimental Results

Here we presented data that might explain the mechanism of the apparent co-expression of two OR genes in zebrafish that has been previously reported in the literature (Sato *et al*., 2007). We used conventional in vitro molecular biology assays to map the transcription start sites of the co-expressed genes and was able to find evidence for transcriptional initiation upstream of either one gene of the OR103-1 / OR103-5 gene pair. We were also able to gather evidence for the existence of a single long transcript that might comprise both genes and which may lead to translation of both genes in a bicistronic fashion. In functional experiments, utilizing transient transgenesis or luciferase reporter assays, we found evidence that the intergenic region between the OR103-1 and OR103-5 genes might function as an IRES element that promotes translation of two independent proteins from the same mRNA molecule.

5. DISCUSSION

Many studies have shown that the majority of OSNs express only one OR gene from the much larger genomic repertoire and that OR gene expression is limited to only one of the two parental alleles (Chess *et al.,* 1994; Malnic *et al.,* 1999; Touhara *et al.,* 1999; Serizawa *et al.,* 2000, 2003). This has been called the "one receptor-one neuron" rule. However, certain violations of this rule have also been found. Systematic co-expression of OR genes has been reported in rodents (Rawson *et al.,* 2000; Tian and Ma, 2008), in zebrafish (Sato *et al.,* 2007), and in *Drosophila* melanogaster (Larsson *et al.,* 2004). However, because of the very high number of up to 1500 OR genes in rodents, systematic co-expression is hard to identify. Because of the large number of OR genes it is almost impossible to examine systematically if any two ORs are co-expressed and therefore the actual number of co-expressed OR genes might be higher but have yet to be uncovered. Cases of systematic co-expression of two OR genes are interesting because examination of the mechanisms of co-expression might also shed light onto the mysterious process by which expression of more than a single OR per sensory neuron is prevented in most OSNs.

In zebrafish, co-expression of OR genes has previously been reported for members of the OR103 family (Sato *et al.,* 2007). The OR103 family comprises a total of five genes for which co-expression between OR103-1 and one or two linked OR genes has been shown. The study, however, could not experimentally distinguish between the two potentially co-expressed genes OR103-2 and OR103-5 because of the high sequence homology of 96% between the two genes, which is incompatible with discrimination by in situ hybridization. In studies presented in this thesis, we mapped the structure of the 5'-untranslated sequences of the OR103-5 and OR103-2 genes, but because of their relatively short size of 71 bp and 174 bp, respectively, they were not suitable for the design of specific riboprobes to resolve the issue whether co-expression exists between OR103-1 and OR103-5 or OR103-2, or both.

In this thesis, we have analyzed in some detail, the nature and aspects of the underlying mechanism of this co-expression. My starting assumption was that the underlying cause of this co-expression could result from two alternative mechanisms: OR103-1 could be truly co-expressed with either OR103-5 or OR103-2 by independent initiation of transcription from promoter sequences located upstream of each gene, or from co-translation of two neighboring genes from a single large transcript that contains the OR103-1 and OR103-5 coding sequences. We designed experiments to discriminate between those two, not necessarily exclusive, possibilities.

Using RLM-RACE to identify the transcription start sites of OR103-1 and OR103-5 genes, we could identify specific transcripts for each gene that initiate from sequences located 260 bp and 189 bp upstream of the two genes, respectively. The OR103-1 transcript starts 260 bp upstream of OR103-1 coding sequence and has a spliced 227 bp intron. The OR103-5 transcript initiates 189 bp upstream of the coding sequence and has a 118 bp intron. This indicates that OR103-5 probably has its own promoter, which is active in some OSNs of the OE. Thus, these results are in favor of the hypothesis that transcription of the two potentially co-expressed genes is an independent process and that the observed co-expression is due to a failure of the exclusion mechanism that usually prevents expression of more than one OR gene per OSN.

However, RACE experiments should always be regarded with caution since longer transcripts could always escape detection, because the procedure favors the detection of shorter transcripts. The RLM-RACE procedure, however, includes several important technical steps that ensure the detection of full-length transcripts. A crucial step of the RACE protocol is the initial de-phosphorylation of the RNA, followed by removal of the 5’-CAP structure. This ensures that only full-length transcripts with an intact 5’-CAP structure will be subsequently processed in the adaptor ligation step. Therefore, the short transcripts that only contains the OR103-5 gene coding sequence most likely is a truly expressed transcript. However, because a late step of the RLM-RACE procedure involves PCR amplification of the transcripts, long transcripts might be overlooked due to insufficient amplification of such long fragments. Thus, RLM-RACE could not fully rule out the existence of a much longer transcript of approximately 3500 bp that might contain

the OR103-1 and OR103-5 coding sequences on the same sequence. Regardless of the existence of such a long transcript, at least some OSNs express the OR103-5 gene from a short proximal promoter located 189 bp upstream of OR103-5.

Using a different approach, we asked whether the two OR sequences could be contained within a single long transcript by RT-PCR on OE cDNA. The results indicate that a single large transcript encompassing the OR103-1 and OR103-5 coding sequence might exist. Furthermore, we could detect two alternatively spliced versions: one with a 118 bp intron upstream of OR103-5 unpliced and one without the intron. Thus, in addition to two independent transcripts for OR103-1 and OR103-5, a single long transcript may also exist. The primer binding sites that we used for the RT-PCR approach were located within the two OR coding sequences flanking the intergenic region between them, amplifying a 1906 bp fragment. Unfortunately, we were not able to amplify a longer version of the transcript, using oligonucleotide primers that were located within the 5'-region of the OR103-1 gene and the 3'-regions of the OR103-5 gene, respectively. This might be due to the overall size of such a fragment and the fact that PCR conditions for such a long fragment were not optimal. However, some of the results provide good evidence that the long transcript is expressed. First, comparing the experimental RT-PCR reaction with the reaction for the negative –RT control cDNA, we could detect two specific bands of 1906 bp and 1788 bp size in +RT. We could also detect a faint band in the –RT control reaction resulting from genomic contamination, but the intensity of that band was much lower when compared to the experimental sample. Experimental cDNA template and –RT cDNA have been processed side by side from the same extracted total RNA. Thus, genomic contamination should be identical in both samples. The fact that the band was much stronger is in favor of the existence of a specific RNA in addition to genomic sequences spanning this region. In addition, we could detect a shorter band in the experimental cDNA sample which corresponds to a variant of the transcript in which the 118bp intron has been spliced out. Since splicing is specific for RNA, at least the alternatively spliced variant is good proof of the existence of the long transcript harboring both genes.

However, another possibility by which such a result could have been obtained exists. If the OR103-1 and OR103-5 genes had overlapping mRNAs, i.e. the 3'-UTR of

OR103-1 overlapped with the 5'-UTR of OR103-5, the long fragment could have been amplified by overlap PCR, giving rise to a hybrid product and therefore a false result. To rule out this possibility, we performed gene specific RACE for the OR103-5 gene. In this experiment only cDNA from OR103-5 transcripts has been generated through the use of a specific OR103-5 oligonucleotide primer during first-strand cDNA synthesis. Using this template, we were able to amplify sequences 5' of OR103-1. Overlapping OR103-1 3'-sequences and OR103-5 5'-sequences should not have resulted in the generation of a hybrid product during the generation of first-strand cDNA, thus indicating the existence of a transcript that contains OR103-1 and OR103-5.

Thus, we have obtained some contradictory results providing evidence for the existence of individual and combined transcripts, thus either of the two hypotheses, respectively. However, the true situation might be a combination of both findings. OR103-1 is always co-expressed with either OR103-5 or OR103-2 but only a fraction of OR103-5 / OR103-2 are double positive for OR103-1. Thus different OSN populations might exist that express different transcripts originating independently from the OR103-1 and OR103-5 promoter. Double positive OSNs might express a transcript from the OR103-1 promoter and this transcript might always be a long transcript including the OR103-5 sequence, thereby accounting for the observed co-expression. In addition, some OSNs might express only from the OR103-5 promoter thereby accounting for the population of OSNs that are single positive.

Further experiments are needed to unequivocally confirm this model. Two experimental approaches would be particularly informative: analysis by northern blot and in situ hybridization. Northern blot analysis would allow for the detection of different transcripts contained within the complex pool of mRNAs transcribed by the OE. If transcripts of different size exist that are expressed by different OSN cell populations specific probes could be designed to reveal the actual size of such transcripts. We made an attempt by using "virtual northern blot" analysis using amplified cDNA of the OE and hybridized gel separated samples with probes specific for OR103-1, Or103-5, and the intergenic region between them. Unfortunately, we could not reach the necessary sensitivity to detect any transcript except from control plasmids that were blotted onto the

same membrane. It might be that radioactive detection would provide the required sensitivity. We also tried to establish in situ hybridization for the intergenic region but did not have good success with this technique. It would be informative to demonstrate that co-expressing OSNs are positive for expressed sequences of the intergenic region and that OSNs single positive for OR103-5 do not.

The single transcript encompassing two OR genes might result from failure of the RNA polymerase to terminate in the 3'-region of OR103-1. In order for RNA polymerase to terminate transcription specific poly-A signals downstream of a given coding sequence exist. It's probable that a poly-A site is missing after OR103-1 coding sequence, allowing the polymerase to further transcribe until it reaches the first instructive poly-A signal after the OR103-5 gene.

So far, my experiments provide some evidence that OR103-1 and OR103-5 can be located on a single mRNA transcribed from the OR103-1 promoter, which would explain the double labeling that was observed in in situ experiments (Sato *et al.*, 2007). However, a single transcript encompassing the two coding sequences does not directly imply that both genes are translated. In order for the two coding sequences on a single mRNA transcript to be both translated into proteins, the sequence between the two OR coding sequences should allow for independent initiation of translation. Such sequences have been identified as IRES sequences that will present the second coding sequence to the ribosome (Pelletier and Sonneberg, 1988; Jang *et al.*, 1989).

We wanted to directly test whether both OR genes could be translated and whether the intergenic region between OR103-1 and OR103-5 contains IRES activity. We designed transgenic constructs to be injected into zebrafish oocytes for transient expression to test whether two sequences linked by the intergenic region can be co-translated. We changed the coding sequences of OR103-1 and OR103-5 with mCherry and GFP, respectively, keeping the intergenic region, the candidate for IRES sequence, and put it under control of the OMP gene promoter, which drives expression in ciliated OSNS in the OE. When injected to zebrafish embryos the pOMP::CIG construct resulted in 97-100% co-

translation. This has two implications. First, the sequence between those coding sequences, the intergenic region, has IRES activity in the OE. Secondly, OR103-1 and OR103-5 could both be translated from the single mRNA observed.

The identification of a novel IRES sequence active in the olfactory system might have an important technological consequence for zebrafish studies. The IRES sequence from the EMCV, which is commonly used in mammalian systems, has been reported to have low or no efficiency in the olfactory system (Sato *et al.,* 2007) although it has been shown to be active in the zebrafish model (Fahrenkrug *et al.,* 1999). So far, there have been only two reports of sequences with IRES activity that were identified from the zebrafish (Ul-Hussain *et al.,* 2008; Lekven *et al.,* 2001). One is on the mRNA of Connexin-55.5, which is expressed in the developing embryo and important in formation of the body plan, and the second is *wnt8,* which is expressed during development in the mesoderm and neural crest. However, the IRES activity of the first one is specific to only neurons expressing polyprimidine tract binding protein and the second one leads to splicing of the first coding sequence, which make them unsuitable for being used to tag proteins of interest in researches using zebrafish as a model.

In order to understand if the intergenic region between OR103-1 and OR103-5 has IRES activity in other zebrafish tissues, we used the rat GAP43 promoter to drive expression in maturing neurons in the entire central nervous system. The construct pGAP-43::CIG was injected into zebrafish embryos and notochord cells double positive for mCherry and GFP were observed in the tails up to the mid-dorsal parts of the embryos. 100% of the cells labeled (6 in total) were positive for both fluorescent proteins. The cells labeled are probably notochord cells because of morphology and location. This result shows that the intergenic region between OR103-1 and OR103-5 has IRES activity in the notochord of the zebrafish, thus outside the olfactory tissues.

However, to reach fully conclusive levels, certain control experiments need to be conducted. As indicated by the RACE results, the intergenic region contains the OR103-5 promoter and promoter activity could be exceptionally high in olfactory or other tissues. So

far, results on the injection of a construct containing only the intergenic region driving reporter gene expression are very limited and no positive cells have been found. The combination of both promoters, the OMP gene promoter and the OR103-5 gene promoter, on the same plasmid could result in synergistic effects and the otherwise less efficient OR103-5 promoter could "hitchhike" on enhancer elements contained within the OMP promoter, rendering it more effective. Thus, pomp::C and intergenic region::GFP should be co-injected as separate plasmids to check whether a similarly high level of co-labeling can be observed. It should also be examined whether a single transcript encompassing mCherry and GFP can be detected by RT-PCR in transgenic fish. However, these experiments will be very difficult as the number of transgenic OSNs and transgene-expressing embryos is quite low and would need to await the construction of a stable transgenic zebrafish line. However, preliminary results are available from independent studies that utilized the co-injection of pomp::Cherry with OR-specific promoters driving GFP (X. Bayramli, I Tastekin, unpublished results), where co-expression of the two transgenes was seen in ~40% of OSNs for the OR101-1 and OR111-7 promoters.

However, the preliminary results obtained so far are consistent with the notion that both OR genes, OR103-1 and OR103-5 could be translated from the same transcript and that the intergenic region has IRES activity. Interestingly in the case of the pOMP::CIG transgene occasional OSNs expressing GFP only were observed. This closely resembles the situation in the olfactory epithelium, where different cell populations were detected that do co-express both genes and those that only express OR103-5 (Sato *et al*., 2007).

The results also indicate that OR103-1 and OR103-5 are most likely both translated from the single mRNA observed, however this fact would need additional experimentation to be verified unequivocally. Specific antibodies would be needed against the two different ORs, but the generation of OR-specific antibodies has been extremely difficult because of the high numbers of very similar OR proteins (Strotmann and Breer, 1991; Barnea *et al.,* 2004).

The candidate IRES within the OR103-1 / OR103-5 intergenic region was tested in an in vitro system to quantify its activity in Hk-2 cells. Cell cultures provide easy means to measure IRES activity quantitatively and critical control construct can be obtained easily (Coldwell *et al.,* 2001). We observed high levels of renilla an firefly luciferase activity when the intergenic region was present between the two reporter genes. The intergenic region increased the activity of the second cistron, firefly luciferase about 170-fold over the empty control plasmid. The well established and extensively used EMCV IRES resulted in a 20-fold increase, thus, the OR103-1 / OR103-5 intergenic region has a 8-fold higher IRES activity than EMCV IRES. Consistent with its IRES function, a critical control experiment that blocks translation of the first cistron reduced renilla activity but did not alter activity of the firefly luciferase.

However, additional control experiments are necessary. For instance, the intergenic region contains the OR103-5 promoter, which in combination with the SV40 enhancer could lead to high expression levels of firefly luciferase. In fact, many sequences suspected to have IRES activity in mammalian systems have turned out to have cryptic promoter activity rather than IRES activity (Bert *et al.,* 2006). This could be tested by removing the SV40 promoter upstream of the first cistron, which should abolish expression of both luciferase genes even in the presence of an active IRES. We engineered a construct that lacks the SV40 promoter and tested it in a limited set of experiments. Due to difficulties with normalization of results using control plasmids, the results are not included in the scope of this thesis, but remain to be performed as an essential control to verify the IRES activity of the intergenic region in Hk-2 cells. In preliminary studies we have observed that some cryptic promoter activity might contribute to the observed IRES activity.

The minimal IRES sequence contained within the rather long intergenic fragment of 1.4 kb has also not been identified. Experiments could be designed to narrow down the active IRES elements through serial deletions of sequences from the intergenic region.

All these difficulties withstanding the IRES activity from zebrafish origin that was discovered in this thesis has strong potential to be used as a tool in other zebrafish studies.

In particular in our studies on the regulation of OR gene expression powerful IRES sequences are essential. However, the commonly available IRES sequences have not proven to give reliable results. Alternative tools have been developed. Recently T2 viral peptides have been widely used to generate fusion proteins between the gene of interest and a reporter gene. The reporter gene gets liberated from the nascent protein through ribosomal skipping at the T2 sequence. However, the use of these sequences has not been successful in the zebrafish olfactory system. It would be interesting to test, whether the intergenic region, or a part of it, could be effectively used to overcome these experimental difficulties.

In summary, my results indicate that the intergenic region between the OR103-1 and OR103-5 genes might promote translation of the downstream sequence from a single RNA molecule that is expressed in a fraction of OSNs that express a single bicistronic transcript from the OR103-1 promoter. Another cell population that is single positive for the OR103-5 or OR103-2 genes might express a different transcript from independent promoters located upstream of the OR103-5 and OR103-2 genes. This mechanism could explain the observed co-expression of two OR genes in the zebrafish olfactory system. Even though the situation at the OR103-5 / OR103-2 locus constitutes a violation of the one neuron – one receptor rule, it does not constitute a true phenomenon of independent co-expression of two OR genes, which would be informative for the mechanism that ensures the rule in most OSNs. However, it has been speculated that aberrantly specified OSNs that express more than one OR gene undergo elimination during development (Tian and Ma, 2008; Mombaerts 2004). The reason might be that a low number of wrongly specified OSNs would not be able to maintain stable synaptic connections in olfactory glomeruli through an activity dependent mechanism (Tian and Ma, 2008). Interestingly, the OR103-1 / OR103-5 co-expressing OSNs do not seem to be pruned away during development, probably because of the high number of OSNs that express both genes from the time they are born. Thus, elimination of wrongly specified OSNs might be context dependent, which is in agreement with an activity dependent mechanism. A similar example has been observed in transgenic experiments designed to express two OR gene in the same OSN directly from the earliest time point of transgene expression (Nguyen *et al.*, 2007). At least in terms of axonal wiring in the olfactory bulb, the olfactory system does

not care how OSNs adapt their specific odorant response profile and co-expression of two ORs, even with different ligand specificities, would not prevent glomerulus formation, if a large enough homogeneous population of such OSNs is present.

The identified IRES activity of the intergenic region might have significance as an experimental tool beyond the olfactory system and further experiments should be conducted to establish its core functional sequences.

APPENDIX A: EQUIPMENT

Table 6.1. Equipment

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)
Fluorescence Microscope:	Leica Microsystems, USA (MZ16FA)
Freezer:	-20 ^{o}C Arçelik, Turkey
	-80 ^{o}C Thermo Electron Corp., USA (Thermo Forma 723)
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

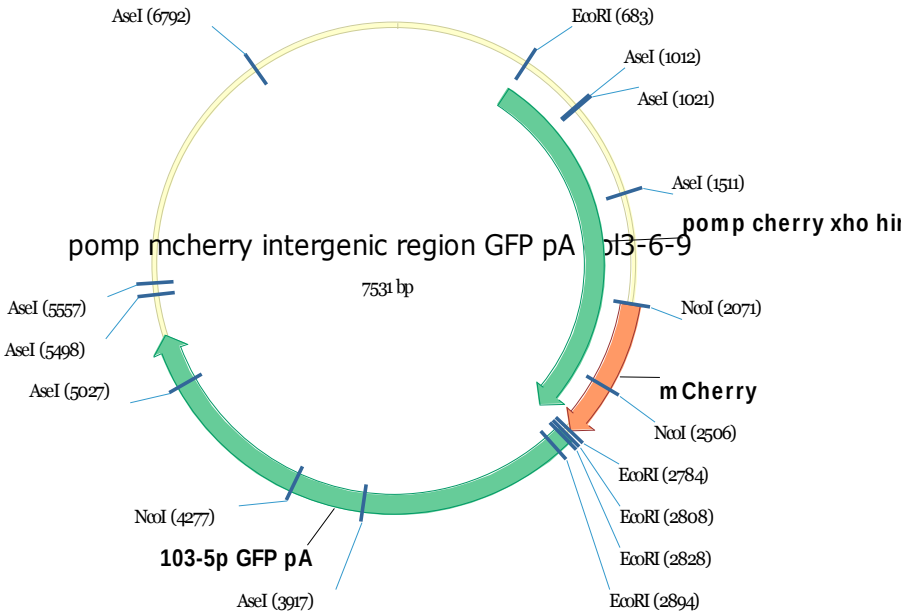
Disposable Labware	
14 ml Culture Tubes:	Greiner Bio-One,Belgium (187261)
CELLSTAR® Centrifuge Tubes,15 ml:	Greiner Bio-One,Belgium (186161)
CELLSTAR® Centrifuge Tubes,50 ml:	Greiner Bio-One,Belgium (227261)
Filtered Pipette Tips:	Greiner Bio-One,Belgium (771288, 772288, 740288)
Micro-centrifuge tubes:	Citotest, China (34730015)
PCR Tubes:	Bio-Rad, U.S.A. (TBS0201)
Chemical Supplies	
1 kb DNA Ladder:	New England Biolabs, U.S.A. (N3232)
100 bp 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).
AseI:	New England Biolabs, U.S.A (R0526 M).
BamHI:	Promega, U.S.A (R602A).
Bovine Serum Albumin:	Promega, U.S.A (R3961).
DMEM Low Glucose	Gibco, U.S.A. (11880).
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).
FuGENE® HD Transfection Reagent	Roche, Germany (04709691001)
Glycerol, for molecular biology:	Sigma-Aldrich, U.S.A. (G5516-500M L).
GoTaq® Flexi DNA Polymerase:	Promega, U.S.A (M830B).
HyClone® Trypsin	Thermo Scientific, U.S.A. (SH30042.01)
LB Agar:	Sigma Life Sciences, U.S.A. (SL08394).
LB Broth:	Sigma-Aldrich, U.S.A. (L7658-1 KG).
Magnesium Chloride,25mM:	Promega, U.S.A (A3511).

Table 6.2. Supplies (continued).

Magnesium Sulfate:	Sigma-Aldrich, U.S.A. (M7506)
Nco I:	Promega, U.S.A (R6515).
Not I:	Promega, U.S.A (R6435).
pGEM®-T Easy Vector System:	Promega, U.S.A (A1360).
Phenol : Chloroform : Isoamyl alcohol:	Sigma-Aldrich, U.S.A. (P2069).
Potassium Chloride:	Sigma-Aldrich, U.S.A. (P9541).
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).
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)
Xho I:	New England Biolabs, U.S.A (R0146).
Commercial Kits	
Dual-Luciferase® Reporter Assay	Promega, U.S.A (E1910).
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)

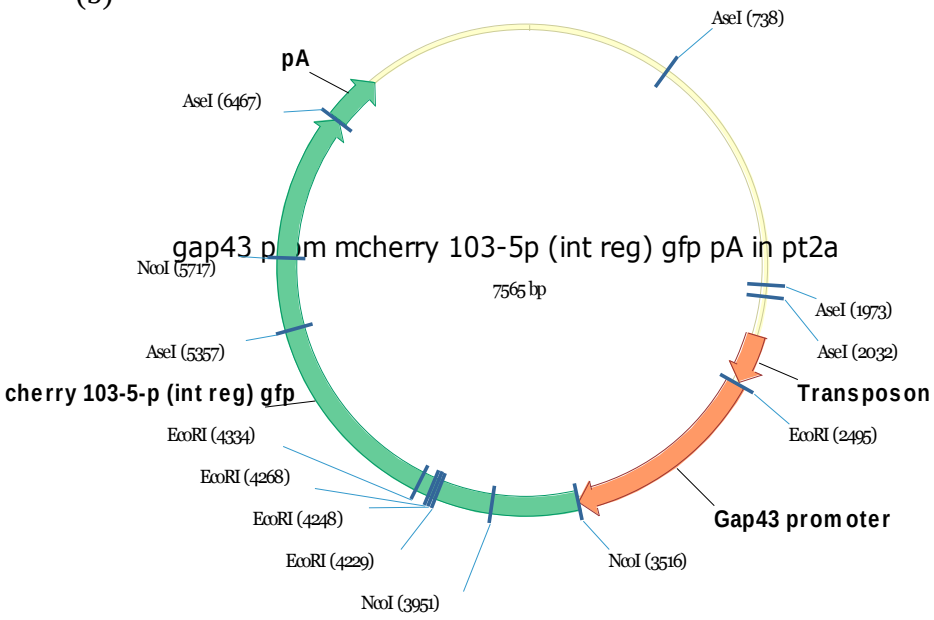


Figure 6.1. Vector Maps

(a) vector map of pOMP::CIG (b) vector map of pGAP-43::CIG

REFERENCES

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

Baier, H. and S. I. Korsching, 1994, "Olfactory Glomeruli in The Zebrafish Form an Invariant Pattern and Are Identifiable Across Animals", *Journal of Neuroscience*, Vol. 14, No. 1, pp. 219-230.

Barnea, G., S. O'Donnell, F. Mancia, X. Sun., A. Nemes, M. Mendelsohn and R. Axel, 2004, "Odorant Receptors on Axon Termini in the Brain", *Science*, Vol. 304, No. 5678, p. 1468.

Bert A. G., R. Grepin, M. A. Vadas, 2006, "Assessing IRES activity in the HIF-1a and Other Cellular 5' UTRs", *RNA*, Vol. 12, pp. 1074-1083.

Bozza, T., A. Vassalli, S. Fuss, J. J. Zhang, B. Weiland, R. Pacifico, P. Feinstein and P. Mombaerts, 2009, "Mapping of Class I and Class II Odorant Receptors to Glomerular Domains by Two Distinct Types of Olfactory Sensory Neurons in the Mouse", *Neuron*, Vol. 61, No. 2, pp. 220-233.

Celik, A., S. H. Fuss and S. I. Korsching, 2002, "Selective Targeting of Zebrafish Olfactory Receptor Neurons by the Endogenous OMP Promoter", *European Journal of Neuroscience*, Vol. 15, No. 5, pp. 798-806.

Chess, A., I. Simon, H. Cedar and R. Axel, 1994, "Allelic Inactivation Regulates Olfactory

Receptor Gene Expression", *Cell*, Vol. 78, No. 5, pp. 823-834.

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*, Vol. 57, pp. 505-516.

Coldwell M. J., M. L. DeSchoolmeester, G. A. Fraser, B. M. Pickering, G. Packham and A. E. Willis, 2001, "The p36 Isoform of BAG-1 is Translated by Internal Ribosome Entry Following Heat Shock", *Oncogene*, Vol. 20, pp. 4095-4100.

Dynes, J. L., J. Ngai, 1998, "Pathfinding of Olfactory Neuron Axons to Stereotyped Glomerular Targets Revealed by Dynamic Imaging in Living Zebrafish Embryos", *Neuron*, Vol. 20, No. 6, pp. 1081-1091.

Fahrenkrug S. C., K. J. Clark, M. O. Dahlquist and P. B. Hackett, Jr., 1999, "Dicistronic Gene Expression in Developing Zebrafish", *Marine Biotechnology*, Vol. 1, pp. 552-561.

Freitag, J. J. Krieger, J. Strotmann and H. Breer, 1995, "Two Classes of Olfactory Receptors in Xenopus Laevis", *Neuron*, Vol. 15, No. 6, pp. 1383-1392.

Friedrich R. W. and S. I. Korsching, 1997, "Combinatorial and Chemotopic Odorant Coding in the Zebrafish Olfactory Bulb Visualized by Optical Imaging", *Neuron*, Vol. 18, No. 5, pp. 737-752.

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. 101-112.

Jang S. K., M. V Davies, R. J. Kaufman and E. Wimmer, 1989, "Initiation of Protein Synthesis by Internal Entry of Ribosomes Into the 5' Nontranslated Region of Encephalomyocarditis Virus RNA In Vivo", *Journal of Virology,* Vol. 63, No. 4, pp. 1651-1660.

Kozak, M., 1989, "The Scanning Model for Translation: An Update", *Journal of Cell Biology*, Vol. 108, pp. 229-241.

Larsson, M.C., A. I. Domingos, W. D. Jones, M. E. Chiappe, H. Amrein and B. Vosshal, 2004, "OR83b Encodes a Broadly Expressed Odorant Receptor Essential For Drosophila Olfaction", *Neuron,* Vol. 43, pp. 703-714.

Lekven A. C., C. J. Thorpe, J. S. Waxman and R. T. Moon, 2001, "Zebrafish Wnt8 Encodes Two Wnt8 Proteins on a Bicistronic Transcript and is Required for Mesoderm and Neuroectoderm Patterning", *Developmental Cell,* Vol. 1, No. 1, pp. 103-114.

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

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. 393-399.

Macejak D. G., P. Sarnow, 1991, “Internal Initiation of Translation Mediated by the 5’ Leader of a Cellular mRNA”, *Nature*, Vol. 353, No. 6339, pp. 90-94.

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

Meister, M. and T. Bonhoeffer, 2001, “Tuning and Topography In an Odor Map On The Rat Olfactory Bulb”, *Journal of Neuroscience*, Vol. 21, No. 4, pp. 1351-1360.

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

Mombaerts, P., 2004, “Odorant Receptor Gene Choice in Olfactory Sensory Neurons: The One Receptor-One Neuron Hypothesis Revisited”, *Current Opinion in Neurobiology*, Vol. 14, No. 1, pp. 31-36.

Mori, K., H. Nagao and Y. Yoshihara, 1999, “The Olfactory Bulb: Coding and Processing of Odor Molecule Information”, *Science*, Vol. 286, No. 5440, pp. 711-715.

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

Niimura, Y., and M. Nei, 2005, “Evolutionary Dynamics of Olfactory Receptor Genes in Fishes and Tetrapods”, *Proceedings of the National Academy of Sciences*

of the United States of America, Vol. 102, No. 17, pp. 6039-6044.

Pelletier, J. and P. Sonenberg, 1988, "Internal Initiation of Translation of Eukaryotic mRNA Directed by a Sequence Derived From Poliovirus RNA", *Nature*, Vol. 334, No. 6180, pp. 320-325.

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. 185-195.

Reinhard E., E. Nedivi, J. Wegner, J. H. P. Skene and M. Westerfied, 1994, "Neural selective Activation and Temporal Regulation of a Mammalian GAP-43 Promoter In Zebrafish", *Development,* Vol. 120, pp. 1767-1775.

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

Sambrook, J., D. W. Russell, 1989, *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York. 2nd Edition.

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

Sato, Y., N. Miyasaka and Y. Yoshihara, 2007, "Hierarchical Regulation of Odorant Receptor Gene Choice and Subsequent Axonal Projection of Olfactory Sensory Neurons in Zebrafish", *Journal of Neuroscience*, Vol. 27, No. 7, pp. 1606-1615.

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. 2088-2094.

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", *Nature Neuroscience*, Vol. 3, No. 7, pp. 687-693.

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

Stoneley, M, F. E. M. Paulin, J. P. C. Le Quesne, S. A. Chappell and A. E. Willis, 1998, "C- myc 5' Untranslated Region Contains an Internal Ribosome Entry Segment", *Oncogene*, Vol. 16, pp. 423-428.

Strotmann, J. and H. Breer, 1991, "Generation of Monoclonal Antibodies Detecting Specific Epitopes in the Olfactory and Respiratory Epithelia", *Cell and Tissue Research*, Vol. 266, No. 2, pp. 247-258.

Tian, H. and M. Ma, 2008, "Activity Plays a Role in Eliminating Olfactory Sensory Neurons Expressing Multiple Odorant Receptors in the Mouse Septal Organ", *Molecular and Cellular Neuroscience*, Vol. 38, No. 4, pp. 484-488.

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 of the United States of America*, Vol. 96, No. 7, pp. 4040–4045.

Ul-Hussain, M., G. Zoidl, J. Klooster, M. Kamermans and R. Dermietzel. "Ires Mediated Translation of the Carboxy-Terminal Domain of The Horizontal Cell Specific Connexin Cx55.5 In Vivo and In Vitro", *BMC Molecular Biology*, Vol. 9, No. 52.

Vassar, R., J. Ngai and R. Axel, 1993, "Spatial Segregation of Odorant Receptor Expression in the Mammalian Olfactory Epithelium", *Cell*, Vol. 74, No. 2, pp. 309-318.

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

Watt, W. C., H. Sakano, Z. Y. Lee, J. E. Reusch, K. Trinh and D. R. Storm, 2004, "Odorant Stimulation Enhances Survival of Olfactory Sensory Neurons via MAPK and CREB", *Neuron*, Vol. 41, No. 6, pp. 955-967.

Westerfield, M., 2007, *The Zebrafish Book: A Guide For The Laboratory Use of Zebrafish (Danio rerio)*, University of Oregon Press: Oregon. 5th Edition.

Young J. M., C. Friedman, E. M. Williams, J. A. Ross, L. Tonnes-Priddy and B. J. Trask, 2002, "Different Evolutionary Processes Shaped Mouse and Human Olfactory Receptor Gene Families", *Human Molecular Genetics*, Vol. 11, No. 5, pp. 535-546.

Zhao H. And R. R. Reed, 2001, "X Inactivation of the OCNC1 Channel Gene Reveals a Role for Activity Dependent Competition in the Olfactory System", *Cell*, Vol. 104, pp. 651-660.