Investigation of alternative polyadenylation of odorant receptor mRNA and its potential role in outgrowing olfactory sensory neurons

(Untersuchung alternativer Polyadenylierung der Geruchsrezeptor mRNA und ihrer möglichen Rolle in auswachsenden olfaktorischen sensorischen Neuronen)

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und Biowissenschaften

des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Yue He

aus Jiangsu, China

Dekan: Prof. Dr. Reinhard Fischer

Referent: Prof. Dr. Martin Bastmeyer

Korreferent: Prof. Dr. Jörg Kämper

Tag der mündlichen Prüfung: 14.10.2019

Eidesstattliche Erklärung

Der experimentelle Teil der vorliegenden Arbeit wurde in der Zeit von Juni 2015 bis Juni 2019 am Zoologischen Institut in der Abteilung für Zell- und Neurobiologie, des Karlsruher Instituts für Technologie (KIT) durchgeführt.

Hiermit versichere ich, dass ich diese Arbeit selbstständig verfasst und nur die angegebenen Quellen und Hilfsmittel verwendet habe. Wörtlich oder inhaltlich übernommene Stellen sind als solche gekennzeichnet. Die Satzung des Karlsruher Instituts für Technologie (KIT) zur Sicherung guter wissenschaftlicher Praxis habe ich in der gültigen Fassung beachtet. Die Arbeit wurde in keiner Form einer anderen Prüfungsbehörde vorgelegt. Ich versichere, dass die beigelegte, elektronische Version der Arbeit mit der schriftlichen übereinstimmt.

Karlsruhe, den 15.11.2019   Yue He

Abstract

In the mouse, more than 1000 functional odorant receptor (OR) genes are expressed in a monoallelic manner by the olfactory sensory neurons (OSNs) residing in the olfactory epithelium (OE). All axons of OSNs expressing the same type of OR converge stereotypically in one or few of the 1800 glomeruli in the olfactory bulb (OB), forming a highly specific receptotopic map. The mechanisms of olfactory axon guidance have remained largely elusive. It has been established, however, that the OR has a necessary instructive role in the guidance process but that it is insufficient.

We previously performed a thoroughly controlled differential single-cell cDNA screen to find additional receptor-specific guidance components. Surprisingly we found for the exemplary receptor mOR37A that a non-coding 3'-extension of the OR mRNA, presumably generated by alternative polyadenylation (APA), is exclusively expressed during axon guidance.

The major findings of this thesis were based on the odorant receptor mOR37A in a mOR37A-IRES-tauEGFP mouse line. First of all, I ensured that the IRES-tauEGFP insertion in this strain does otherwise not alter the transcript structure by sequencing the insertion site and by a detailed characterization by 3' Rapid Amplification of cDNA Ends (3' RACE) of the mOR37A short and long transcript isoforms generated from this locus. This confirms and extends our previous results in wild type mice. Next, the interplay between alternative splicing in the 5' untranslated region (UTR) and APA was investigated. However, no direct correlation was discovered. In addition, an OSN maturation-dependent regulation of the long isoform was addressed by quantitative RNA-fluorescence *in situ* hybridization (FISH), where it was found to be regulated from a 1:1 ratio with the short isoform in the immature population to a complete switching off in the mature population. By Droplet Digital PCR (ddPCR), a global developmental downregulation of the long isoform was confirmed in the OE. Notably, preliminary evidence from other receptors, mOR37C and olfr701, also demonstrated APA and a similar developmental isoform regulation. After switching the experimental focus to the olfactory axons, I found that the proximal amplicon of the mOR37A mRNA was consistently suppressed from ddPCR amplification. This suppression showed a highly specific effect by most likely affecting only the short isoform and by occurring exclusively in the axon terminals during the guidance phase. Series of loss-of-function experiments with RNase treatment proved that the inhibitor was an RNA molecule. A gain-of-function experiment, aiming to reconstitute the effect with the non-inhibited *in vitro* synthesized short and long RNAs together with natural small RNA species isolated from the OB, hinted at a small RNA as the potential interaction partner. Its identity remains to be clarified experimentally.

In sum, a model is proposed in which the short isoform might be suppressed in the outgrowing axons, since the same protein, only when translated from the long isoform, might recruit OR-specific guidance factors via an UTR-dependent protein localization (UDPL) mechanism.

Zusammenfassung

Das Mausgenom umfasst mehr als 1000 funktionelle Geruchsrezeptor-(OR)-Gene, wobei jedes olfaktorische sensorische Neuron (OSN) des Riechepithels (OE) stochastisch nur eines zur Expression auswählt. Dabei konvergieren Axone von OSNs, die den gleichen Rezeptor exprimieren, in einen oder wenige der 1800 stereotypen Glomeruli im olfaktorischen Bulbus (OB) und erzeugen somit eine rezeptotope Karte. Die Mechanismen der olfaktorischen Verdrahtung sind noch weitgehend unklar. Es ist aber etabliert, dass der OR eine notwendige aber nicht ausreichende Rolle bei der Axonlenkung spielt.

Wir haben zuvor einen streng kontrollierten differenziellen Einzelzell-cDNA-Screen durchgeführt, um zusätzliche rezeptorspezifische Lenkungskomponenten zu finden. Überraschenderweise haben wir bei dem exemplarischen Rezeptor mOR37A festgestellt, dass eine nicht-kodierende 3'-Extension der OR-mRNA, vermutlich erzeugt durch alternative Polyadenylierung (APA), ausschließlich während der Axonlenkung exprimiert wird.

Die Ergebnisse dieser Arbeit wurden auf der Grundlage des Rezeptors mOR37A in einem mOR37A-IRES-tauEGFP-Stamm gewonnen. Zunächst wurde sichergestellt, dass die IRES-tauEGFP-Insertion in diesem Stamm die Transkriptstruktur ansonten nicht verändert, indem die Insertionsstelle sequenziert wurde und die kurzen und langen Transkriptisoformen von diesem Locus detailliert durch *3' Rapid Amplification of cDNA Ends* (3' RACE) charakterisiert wurden. Dies bestätigt und erweitert unsere bisherigen Ergebnisse bei Wildtyp-Mäusen. Als Nächstes wurde der Zusammenhang von alternativem Spleißen in der 5'-untranslatierten Region (UTR) und APA untersucht. Es wurde jedoch kein direkter Zusammenhang festgestellt. Darüber hinaus wurde die OSN-reifungsabhängige Regulation der langen Isoform durch quantitative RNA-Fluoreszenz-in-situ-Hybridisierung (FISH) analysiert, wobei gefunden wurde, dass sie von einem 1:1-Verhältnis mit der kurzen Isoform in der unreifen Population bis hin zu einer vollständigen Abschaltung in der reifen Population reguliert wurde. Durch *Droplet Digital PCR* (ddPCR), wurde eine globale Reduktion der langen Isoform im OE während der Entwicklung bestätigt. Bemerkenswerter Weise zeigten vorläufige Resultate von anderen Rezeptoren (mOR37C und olfr701) ebenfalls APA und eine ähnliche entwicklungsabhängige Isoformregulation. Nachdem ich den experimentellen Fokus auf die olfaktorischen Axone verlagert hatte, stellte ich fest, dass das proximale Amplikon der mOR37A-mRNA bei der ddPCR-Amplifikation konsistent unterdrückt wurde. Diese Unterdrückung war sehr spezifisch, da sie höchst wahrscheinlich nur die kurze Isoform betraf und nur während der Lenkungsphase exklusiv in den Axonterminalien auftrat. Eine Reihe von *loss-of-function* Experimenten mit RNase-Behandlung bewies, dass der Inhibitor ein RNA-Molekül war. Ein *gain-of-function* Experiment, das darauf abzielte, den Effekt mit nicht inhibierten *in vitro* synthetisierten kurzen und langen RNAs und natürlichen kleinen RNA-Spezies des OB zu rekonstruieren, deutete auf eine kleine RNA als möglichen Interaktionspartner hin. Seine Identität muss noch experimentell geklärt werden.

Zusammenfassend wird ein Modell vorgeschlagen, in dem die kurze Isoform in den auswachsenden Axonen unterdrückt werden könnte, weil das gleiche Protein nur dann, wenn es aus der langen Isoform translatiert wird, OR-spezifische, lenkungsrelevante Faktoren über einen UTR-abhängigen Proteinlokalisierungsmechanismus (UDPL) rekrutieren könnte.

CONTENTS

1. Introduction

The olfactory system represents one of the most ancestral sensory modalities in the evolutionary history of animals. It exerts influences on both innate and learned behavioral responses, from the simple avoidance of repellents to the complicated recall of smell-associated memories. Olfaction is essential for both invertebrates and vertebrates, but the degree of its importance varies between different species. We humans, like many other higher primates, have reduced olfactory capabilities compared to other macrosmats, such as rodents, probably due to our highly developed visual system. Nevertheless, our sense of smell has substantial hedonistic value. For instance, we would appreciate the sweet smell of lilacs in a summer garden and be able to recall associated memories whenever being exposed to it.

The olfactory system is unique in several aspects. The olfactory sensory neurons are the only primary sensory neurons in vertebrates whose cell bodies are exposed to the external world. This on the one hand largely enables the processing of volatile odorant molecules, on the other hand renders this population vulnerable to injuries. To cope with that, the olfactory epithelium possesses a lifelong regenerative capacity from its stem cells. Another observation is that olfactory axons are not myelinated by the peripheral Schwann cells, but are ensheathed by the olfactory specific ensheathing glial cells. Furthermore, the olfactory projection into higher cortical regions lacks the thalamic relay, which is typical for other peripheral sensory systems.

The fundamental work describing the basic organizational principles of the olfactory system was performed by Richard Axel and Linda Buck in the 1990s, who were awarded a Nobel Prize in Physiology and Medicine a decade later. Simply speaking, odor detection is mediated by a huge family of G-protein coupled receptors called odorant receptors expressed by the olfactory sensory neurons in the nasal epithelium. These cells send axonal processes to innervate the olfactory bulb of the brain, making excitatory connections with secondary neurons there within distinct and stereotypic microdomains called glomeruli, thereby establishing a spatial representation of the odor. From there, the information is further relayed to higher brain regions of olfactory cortex to elicit cognitive and behavioral responses. The mammalian olfactory system conforms to two governing rules, the "one neuron-one receptor" and the "one receptor-one glomerulus" rule, meaning each olfactory sensory neuron expresses only one receptor in a monogenic and monoallelic manner and those with the same receptor, although being scattered throughout the sensory surface, project their axons to the same glomerulus. Recent advances in single-cell transcriptomics as well as genetic manipulations have greatly promoted the basic understanding of the cellular and molecular landscape of the olfactory system. We are particularly interested in the olfactory axon guidance mechanism, which seems to differ from what is utilized by other sensory systems, as a non-topographic projection is established between the peripheral olfactory

epithelium and the central olfactory bulb. In this thesis, the unexpected role of odorant receptor mRNA alternative polyadenylation in the olfactory axon guidance process is studied.

1.1 The mouse olfactory system

The mouse olfactory system contains a number of subsystems, each of which serves distinct functions (Fig.1.1). Two major chemosensory systems are the main olfactory epithelium (MOE) in the posterior nasal cavity and the vomeronasal organ (VNO) along the base of nasal septum. It is generally believed that the MOE mediates discrimination of volatile odorants, while the VNO is specialized for pheromone detection. The olfactory information originating from these two systems is also processed in separate pathways. Sensory signals collected in the MOE are relayed through the main olfactory bulb (MOB) to the olfactory cortex, whereas sensory neurons in the VNO project to the accessory olfactory bulb (AOB), from which signals are further transmitted towards the amygdala and hypothalamus (reviewed in Buck, 2000). The sensory neurons in the MOE express a huge diversity of odorant receptors. In addition, the MOE contains a small population of sensory neurons expressing a distinct subtype of receptor guanylyl cyclase and they project their axons to 12 glomeruli in the caudal MOB, forming the "necklace glomeruli" (Fülle at al., 1995; Juilfs et al., 1997). It has been demonstrated that the guanylyl cyclase subsystem is involved in CO_2 detection at near-atmospheric concentrations (Hu et al., 2007). A third olfactory subsystem, the septal organ, is a small island of sensory epithelium lying near the ventral base of the nasal septum (reviewed in Breer et

Figure 1.1. Subsystems of the mouse olfactory system (midsagittal section of the mouse nasal cavity and forebrain). The two major subsystems are the main olfactory epithelium (MOE) with its neural connection to the main olfactory bulb (MOB), and the vomeronasal organ (VNO) with its neural connection to the accessory olfactory bulb (AOB). Guanylyl cyclase expressing neuorns are located in the MOE and project their axons to the necklace glomeruli. Other subsystems include the septal organ and the Grueneberg ganglion. Figure modified from Zufall and Leinders-Zufall, 2007.

al., 2006). The septal organ resembles the MOE to a great extent, with regard to the cellular organization, the odorant response properties as well as the projection to the MOB (Ma et al., 2003; Levai and Strotmann, 2003; Tian and Ma, 2004). It is therefore considered as a "mini-nose". Another specialized subsystem is the Grueneberg ganglion, located at the rostral tip of the nose (Gruneberg, 1973). Its lifelong expression of olfactory marker protein (OMP) and direct innervation of the caudal MOB confirms its olfactory nature (Fuss et al., 2005; Koos and Fraser, 2005; Fleischer et al., 2006). The Grueneberg ganglion glomeruli are near the necklace glomeruli

at the junction of MOB and AOB (Fuss et al., 2005; Koos and Fraser, 2005). It has been identified as a chemodetector of alarm pheromones (Brechbühl et al., 2008).

1.2 The cellular organization of the mouse olfactory system

1.2.1 Cellular composition of the olfactory epithelium

The olfactory epithelium (OE) is a ciliated pseudostratified columnar epithelium (Morrison and Costanzo, 1992). The cellular constituents of the OE include several types of both non-neuronal and neuronal cells (Fig.1.2).

Apically, the non-neuronal supporting sustentacular cells are found. They are capped by microvilli, and contain abundant mitochondria and endoplasmic reticulum (Farbman, 1992). Analogous to neural glial cells, they provide physical and metabolical support to the OE (Getchell and Getchell, 1992).

Closest to the basal lamina are two populations of stem cells, the horizontal basal cells (HBC) and the globose basal cells (GBC). Unlike most other neural tissues, the olfactory epithelium displays a pronounced lifelong regenerative capacity, implying an unlimited self-renewal of the olfactory stem cells (Schwob et al., 1994). Among these two populations, HBSs are mitotically quiescent under basal conditions, whereas GBCs are the active progenitors that give rise to both olfactory sensory neurons (OSN) and supporting cells during ongoing neurogenesis (Holbrook et al., 1995; Huard and Schwob, 1995). Recent experimental data have also demonstrated a more direct role of the HBCs in actively reconstituting all major cells types in the OE upon targeted destruction of sensory neurons (Iwai et al., 2008; Fletcher et al., 2017).

Situated in the intermediate layers are the olfactory sensory neurons, which form the bulk of the OE (Graziadei and Monti Graziadei, 1979). OSNs are bipolar neurons with one single dendrite and one single axon. The apical dendrite ends in a knob form with 10-20 fine cilia splaying out over the OE. The surface membrane of these cilia contains odorant receptors for odor detection (Menco, 1980; Farbman, 1992). The single unmyelinated axon exits the basal epithelium to join the olfactory nerve targeting the olfactory bulb (Farbman, 1992). Though the olfactory axons are not myelinated by Schwann cells, they are enveloped by the olfactory ensheathing cells (Chuah and West, 2002). Based on the differentiation levels, OSNs are roughly subdivided into mature and immature states. Several neuronal marker genes are commonly used to identify them. For example, Growth associated protein (GAP43) is a marker for the immature OSNs during outgrowth (Verhaagen et al., 1989), whereas olfactory marker protein (OMP) defines the mature OSNs with functional signal transduction cascade (Farbman and Margolis, 1980). The ratio of immature to mature OSNs varies throughout the life of the animal. Furthermore, benefiting from single-cell transcriptomics and lineage tracing data, intermediate neuronal precursors have been

identified in olfactory cell lineage trajectories between neuronal progenitor cells and differentiated OSNs (Hanchate et al., 2015; Fletcher et al., 2017). It has also been nicely demonstrated that the odorant receptor expression starts in the late precursor stage and the developing OSNs initially express multiple receptors before final lock in of one receptor in the mature stage (Hanchate et al., 2015; Fletcher et al., 2017).

In addition, Bowman's glands extend from the lamina propria to the OE's surface. The Bowman's glands consist of cells with abundant secretory granules, serving mainly as a protection layer of the olfactory mucosa against infection and dehydration (Solbu and Holen, 2011).
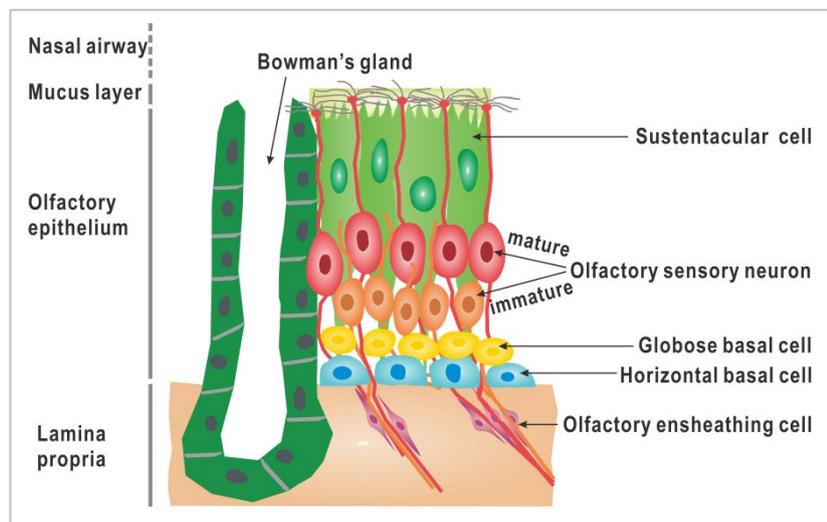
Figure 1.2. Cellular constituents of the olfactory epithelium.
Olfactory epithelium is a pseudostratified epithelium. The major cell types are sustentacular cells, olfactory sensory neurons (OSN), basal progenitor cells including globose basal cells (GBC) and horizontal basal cells (HBC), and Bowman's glands. OSNs are bipolar neurons, extending one basal axon and one apical dendrite. Their axons are non-myelinated but enveloped by the olfactory ensheathing cells. Figure modified from Schwob et al., 2017.

1.2.2 Cellular composition of the olfactory bulb

To reach the first processing center of the olfactory pathway, axons of OSNs exit the OE in nerve bundles, pass through the cribiform plate and reach the surface of the olfactory bulb (OB). These axon bundles defasciculate once contacting the bulb, and those from the same receptor-specific population are resorted in the olfactory nerve layer to converge into distinct spherical neuropils of 100-200μm in diameter called glomeruli on the OB surface (Whitesides and LaMantia, 1996). The glomeruli form the glomerular layer. Within each glomerulus, OSN axons make excitatory synaptic connections with the apical dendrites of the OB projection neurons, the mitral/tufted cells, and with local-circuit interneurons called periglomerular cells (Kasowski et al., 1999). From here, the olfactory input from the periphery can be further relayed to the olfactory cortex.

The OB can be divided into several layers apart from the aforementioned glomerular layer, namely the external plexiform layer, the mitral cell layer and the granule cell layer (Fig.1.3).

Each layer is characterized by morphologically and functionally distinct cell types. In the mitral cell and the external plexiform layer, cell bodies of both types of projection neurons, mitral and tufted cells are found (Price and Powell, 1970b). They share certain morphological as well as biophysical properties. For example, they both extend a single apical dendrite into one glomerulus in the glomerular layer to form synaptic contacts with olfactory nerve axons, and they both extend their secondary dendrites tangentially in the external plexiform layer, though those of the tufted cells are much more superficially positioned than those of the mitral cells (Price and Powell, 1970b; Orona et al., 1984). As the most abundant GABAergic interneuron type in the bulb, granule cells reside in the granule cell layer in the center of the bulb and extend an apical dendrite towards the external plexiform layer. Within the external plexiform layer, granule cells synapse with the lateral dendrites of mitral cells (Price and Powell, 1970a). Such reciprocal synapses mediate the lateral inhibition in the OB, which results in slow firing patterns of the mitral cells and finally refines the representation of each odor molecule over time (Friedrich and Laurent, 2001; Nusser et al., 2001).

One of the fundamental rules that the olfactory system conforms to is the so called "one receptor-one glomerulus" rule. It refers to the fact that the OSNs expressing the same type of OR converge their axons onto one or two glomeruli at stereotyped positions per bulb, which follow a mirror-symmetry with one on the medial side and the other on the lateral side. Moreover, this mirror-symmetry is reflected in the other bulb, making a bilaterally symmetrical glomerular map within the OB with little variability between animals (Ressler et al., 1994; Vassar et al., 1994; Mombaerts et al., 1996).
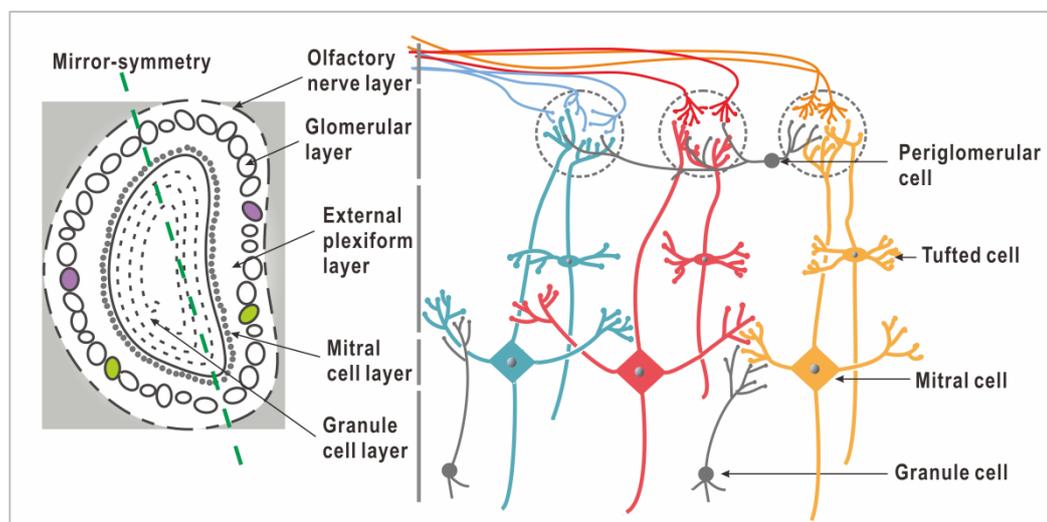
Figure 1.3. Cellular constituents in the olfactory bulb.
Olfactory bulb is a multi-layered structure, with different types of neurons residing in each layer. The major cell types include projection neurons, mitral cells and tufted cells, and interneurons such as periglomerular cells and granule cells. The somata of the mitral cells are found in the mitral cell layer, and those of the tufted cells are in the external plexiform layer. The granule cell layer is largely composed of granule cells, which lies in the center of the OB. More superficially is the glomerular layer, where the axons of OSNs synapse with mitral/tufted cells and periglomerular cells in a spherical structure called glomerulus. The afferent projection from the OE forms the olfactory nerve layer. Each OB contains a mirror-symmetrical glomerular map with two glomeruli innervated by OSNs expressing the same OR (color coded here). Figure modified from Haines and Mihailoff, 2018.

1.3 Odorant receptors

Discovered in the 1990s, odorant receptor (OR) genes represent the largest multigene family in the mammalian genome (Buck and Axel, 1991). The significance of olfaction might eventually be mirrored by the abundance of ORs expressed in one species. There are roughly 1300 OR genes expressed in the mouse olfactory system, including about 20% pseudogenes (Zhang and Firestein, 2002). In contrast to mouse, humans have approximately only 640 OR genes and 47% of them are likely to be pseudogenes (Malnic et al., 2004).

In general, the mouse OR genes are spread in 27 clusters on nearly all chromosomes except chromosomes 12 and Y (Zhang and Firestein, 2002). Based on phylogenetic analysis, they can be separated into two broad classes and numerous subfamilies (Zhang and Firestein, 2002; Godfrey et al., 2004). Class I ORs are highly related to fish ORs, which are expected to be more sensitive towards water-soluble compounds, whereas class II ORs are mammalian specific and might favor more hydrophobic compounds (Ngai et al., 1993; Zhang and Firestein, 2002). There are 147 Class I OR genes in the mouse, with 120 being potentially functional, and they are located in a large cluster on chromosome 7 (Zhang and Firestein, 2002). The OR gene family can further be divided into more than 200 subfamilies with sequence homology and functional similarity. Despite a wide distribution throughout the whole genome, most subfamilies are found in the same region and encoded by a single locus (Zhang and Firestein, 2002; Godfrey et al., 2004). Mouse OR genes show a distinctive zonal expression pattern in the MOE. OSNs expressing one particular OR are dispersed in one of the four topographically distinct zones in the MOE, but within a given zone the distribution is random (Ressler et al., 1993; Vassar et al., 1993). The four zones are organized in a dorsal-ventral manner in the OE, and such zonal organization in the MOE correlates coarsely with the dorsal-ventral positioning of glomeruli in the MOB, giving rise to a coarse zone-to-zone topography overlaying the non-topographic map (Mori et al., 2000; Miyamichi et al., 2005).

On the mRNA level, the OR transcripts also display some unique features. The coding region of the OR genes is typically of 1kb in length and is devoid of introns (Buck and Axel, 1991). Recent studies suggest that more than 860 OR genes contain a 5' non-coding exon (Ibarra-Soria et al., 2014), which could be located up to 11kb upstream from the coding exon (Glusman et al., 1996; Sosinsky et al., 2000). Furthermore, OR mRNA is subject to intensive post-transcriptional modifications (Young et al., 2003; Shum et al., 2015). Screen of OE cDNA library indicates that more than two thirds of the OR genes exhibit transcriptional isoforms, including 5' alternative splicing of the non-coding exon and 3' alternative polyadenylation (Young et al., 2003). OR 3'UTRs tend to be AU-rich and consequently rich in AU-rich binding elements, which raise the possibility of regulatory functions from untranslated regions (Shum et al., 2015).

On the protein level, ORs belong to the rhodopsin-type of G-protein coupled receptors (GPCR) like many other sensory receptors. They have typical hallmarks of GPCRs, including an extracellular N-terminus, followed by seven helical transmembrane domains (TM), and an intracellular C-terminus (Buck and Axel, 1991). Among them, the TM4, TM5 and the central region of TM3 are highly variable, whereas the TM7 shows the highest level of amino acid conservation (Liu et al., 2003). Such sequence diversity in these helices is assumed to be associated with ligand-receptor interactions (Singer and Shepherd, 1994; Katada et al., 2005). Like many other GPCRs, ORs can switch between different conformations from a basal conformation in the absence of ligand to an active conformation upon ligand binding (reviewed in Kobilka and Deupi, 2007).

1.3.1 Mechanism of odorant receptor choice

In mouse olfactory system, each OSN expresses only one functional OR gene from a repertoire of more than 1000 in a monogenic and monoallelic manner, which is also summarized as the "one neuron-one receptor" rule (Chess et al., 1994; Malnic et al., 1999; Serizawa et al., 2000; Ishii et al., 2001). The singular OR expression is believed to be regulated by a slow chromatin-mediated activation of one OR followed by a fast negative feedback loop that suppresses the activation of additional ORs.

It has been shown that the OR gene repertoire is largely silenced prior to functional OR choice by exhibiting molecular hallmarks of the constitutive heterochromatin, H3K9me20 and H4K20me3 (Magklara et al., 2011). The epigenetic silencing is reinforced by the OSN nuclear organizations, where the OR genes are condensated into a few foci at the edge of heterochromatin clusters in the nucleus (Clowney et al., 2012). The expression of the histone demethylase LSD1 demethylates H3K9, thus desilencing and initiating OR transcription. Indeed, the chosen OR allele has been demonstrated to be free from heterochromatin marks but bear the histone modifications for active transcription (H3K4me3) (Magklara et al., 2011; Clowney et al., 2012). A successful transcription of the chosen OR allele further depends on the convergence of multiple enhancers, both in *trans* and *cis*, which promotes the formation of olfactory receptor compartments and interchromosomal interactions on the active receptor gene (Markenscoff-Papadimitriou et al., 2014; Monahan et al., 2019). Once the choice is made, OR expression activates ER-resident kinase Perk, which phosphorylates the translation initiation factor eif2a. Phosphorylation of eif2a has been shown to halt global translation of most transcripts (Ron and Walter, 2007), but in this case it selectively increases the transcription of the activating transcription factor 5 (Atf5) in the nucleus (Dalton et al., 2013). Subsequently, Atf5 initiates the transcription of adenylyl cyclase 3 (Adcy3), which relieves the stress in the ER and restores the translation (Dalton et al., 2013). As a regulator of OSN maturation, Adcy3 downregulates LSD1 expression to prevent other

repressed loci from being unpacked and finally stabilizes the singular OR transcription (Lyons et al., 2013). Thus, the OR chromatin state and the OSN nuclear architecture together with the downstream OR-induced feedback loop by triggering the unfolded protein response largely lays the molecular foundation of the singular OR expression.

1.3.2 Olfactory signal transduction cascade

Odorant signal transduction is initiated by odorant molecules binding to the ORs on the cilia surface of OSNs (Fig.1.4). As GPCRs, ORs are coupled to heterotrimeric G-proteins, among which the subunit G_{olf} mediates the mature signal transduction cascade (Jones & Reed, 1989; Belluscio et al., 1998). Upon OR activation, the GTP-bound G_{olf} dissociates from the complex and subsequently activates adenylyl cyclase type III, generating elevated concentration of intraciliary cAMP from ATP, which opens the cyclic nucleotide-gated ion channels, allowing the influx of calcium and sodium ions (Brunet et al., 1996; Wong et al., 2000). Ultimately, the depolarization is further achieved by chloride efflux through a Ca^{2+}-gated Cl^- channel and action potentials are fired to relay the signal to the OB (Pifferi et al., 2006).
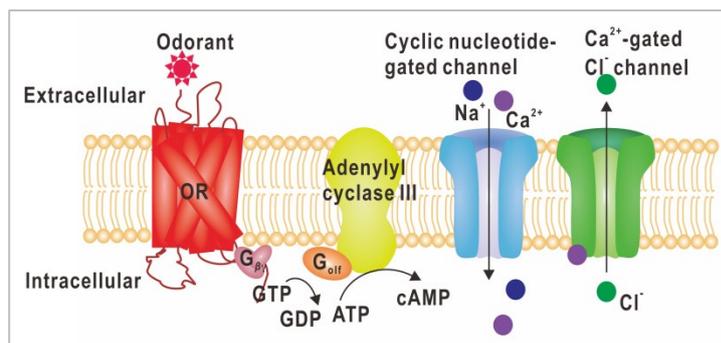
Figure 1.4. Signal transduction cascade in OSN for odorant recognition.
Upon odorant binding, the olfactory signaling cascade is activated through olfactory specific G-protein. The activated G_{olf} disassociates from the $\beta\gamma$ subunits and stimulates adenylyl cyclase III, resulting in the generation of cAMP. The cyclic nucleotide-gated channels and the calcium-gated chloride channels are subsequently opened, carrying the major transduction current. The OSN is then depolarized and action potential is transmitted to the OB. Figure modified from DeMaria and Ngai, 2010.

The discriminatory power of the olfactory system is achieved by combinatorial coding, where an individual odorant can activate multiple ORs and an individual OR can recognize multiple odorants (Malnic et al., 1999). The olfactory information is reflected in an epitope map in the OB (Ressler et al., 1994). Each OR type, therefore each glomerulus which is innervated by that OR type is responsible for one particular epitope present on the odorant. Each odorant contains a variety of different epitopes, and is thereby recognized by different OR types and different glomeruli. In this scheme, each OR and each glomerulus can function as one component of the unique combinatorial coding for various odorants, which allows the discrimination of a large number of odorants.

1.3.3 Odorant receptors from the *mOR37* family

The representative receptor investigated in this thesis is from the mOR37 subfamily, which is also known as the mOR262 subfamily according to the nomenclature of Zhang and Firestein (Zhang and Firestein, 2002) or olfr155-159 according to the nomenclature of the Mouse Genome Informatics. Despite being expressed in the MOE, odorant receptors from the mOR37 subfamily display a variety of unique features. In total, five *mOR37*-like genes have been identified in a cluster on mouse chromosome 4 and are designated as *mOR37A (Olfr155)* to *mOR37E (Olfr159)*, with *mOR37D* being a pseudogene (Strotmann et al., 1999). All of them share a high degree of amino acid similarity of more than 85% (Strotmann et al., 1999). The uniqueness of this family is reflected by its clustered expression pattern, its monoglomerular projection and its protein structure. OSNs expressing the mOR37 receptors are organized in a small patch within the MOE, neglecting the zonal expression rule (Strotmann et al., 1992; Strotmann et al., 1994). Their axons wire to the MOB and converge onto only one glomerulus per bulb in the anterior ventromedial part (Strotmann et al., 2000), where social cues such as urine are perceived (Schaefer et al., 2001). Notably, the projection stereotype of the mOR37 glomeruli is not precise down to each individual glomerulus, but exhibits local permutations in their relative positions (Strotmann et al., 2000). Defined mOR37 glomeruli are connected to the paraventricular nucleus of the hypothalamus instead of the olfactory cortex (Bader et al., 2012). Moreover, six additional amino acids are inserted in the third extracellular loop, marking a unique extension in this domain, which is not seen in other receptors (Strotmann et al., 1999).

The *mOR37* locus contains two other receptor genes *mOR17 (Olfr71)* and *mOR6 (Olfr70)*, which are not members of the narrow *mOR37* gene subfamily, but are phylogenetically closely related to it. Both receptors also show a clustered expression pattern within the mOR37-typical region of the MOE (Strotmann et al., 1999).

Comparative sequence analysis reveals the presence of the *mOR37* family only in mammals, such as rodents and opossum, and there is a high degree of conservation across species (Kubick et al., 1997; Hoppe et al., 2003, 2006). It is therefore assumed that they are tuned to identify mammal-related odors. OR deorphanization experiments have identified fatty aldehydes with different chain lengths as ligands for different mOR37 receptors, with mOR37A preferably activated by C15al, mOR37B by C16al and mOR37C by C17al (Bautze et al., 2012). In searching for the natural source of such odorant compounds, bodily secretions from conspecifics are found to activate mOR37A-C. C16al has been confirmed to be present in mouse fecal pellets, exposure of which can elicit a response from the mOR37B glomerulus, indicating a role in social communication (Bautze et al., 2014; Klein et al., 2015). This might be in line with the aforementioned mammalian origin of the mOR37 receptors and their mapping position in the OB.

1.4 Axon guidance in mouse olfactory system

Unlike many other sensory systems, the spatial organization of the olfactory map is not topographical, in other words, the nearest-neighbor relationships are not preserved between the sensory surface and the central mapping target. On the contrary, the OSN axons are organized by a "typographic" principle, where OSNs with the same type of receptor converge into their target glomeruli in the bulb. It has always been an intriguing question, how could 1000 different types of neurons, which exhibit a random distribution across the entire olfactory epithelium, form such highly stereotyped connections with the olfactory bulb.

Classic axon guidance systems were long believed to utilize some of the prominent families of axon guidance cues, such as netrins, semaphorins, ephrins and slits on the target to achieve appropriate mapping (reviewed in Flanagan, 2006). Contrasting this, target derived cues seem to play a minor role in the olfactory system. In fact, deletion of the secondary projection neurons or interneurons in the OB does not influence the convergence of P2 axons into a glomerulus (Bulfone et al., 1998). Even the complete removal of OB does not prevent axons from sorting and forming glomerulus-like structures (St John et al., 2003). These observations demonstrate an important feature of the olfactory map formation, that is it relies little on fiber/target interactions. Then, what does it rely on?

1.4.1 Proposed models of olfactory axon guidance

The olfactory receptor plays a vital role in axon guidance via fiber/fiber interactions.
The key feature that distinguishes these 1000 types of neurons from each other is the odorant receptor they express. Indeed, an OR-directed guidance mechanism would greatly simplify the system. The initial evidence for the instructive role of ORs in axon guidance originates from the OR swap experiment (Fig.1.5). Deletion of the receptor P2 coding region is seen with failure in convergence into a glomerulus in the OB (Wang et al., 1998). Replacement of the coding region of one OR (recipient CDS) by that of another OR (donor CDS) leads to axonal coalescence into a novel, distinct glomerulus (Mombaerts et al., 1996; Wang et al., 1998; Feinstein and Mombaerts, 2004). Notably, axons of the swapped receptors could not be rewired to either the donor or the recipient glomerulus. Only substitution with the coding region from highly related receptors (e.g. the same chromosomal cluster, the same expression zone in the MOE, high amino acid identity) results in glomerular convergence spatially similar to the donor glomerulus (Wang et al., 1998). These observations strongly suggest that the odorant receptors are one determinant in the olfactory guidance process, yet not the sole determinant. In 2004, Mombaerts et al. proposed a "contextual model of axon sorting" mainly based on the experimental observations made from M71/M72 hybrid receptors (Feinstein and Mombaerts, 2004; Feinstein et al., 2004). The model

suggests that axon sorting is achieved through OR-mediated homophilic and heterophilic interactions. Axons with the same axon identity would coalesce, whereas those with different identities would segregate. This type of axon sorting can be achieved with little involvement of the chemical cues derived from the bulb; in other words, or as the authors summarized "axons do not look for targets – they are the targets!" (Feinstein and Mombaerts, 2004). The absolute positioning of the glomeruli, however, will still need target derived cues.
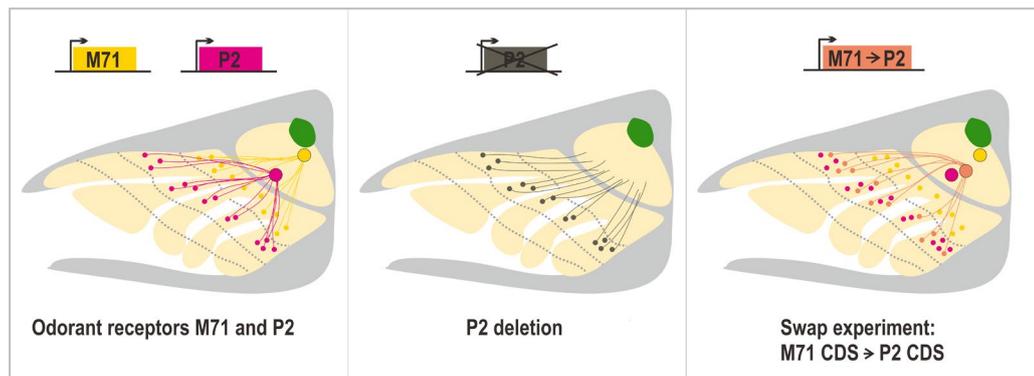
Figure 1.5. Role of odorant receptor in olfactory axon guidance.
Left. OSNs expressing the receptors M71 (yellow) and P2 (magenta) are distributed in zone IV and III and project to distinct glomeruli in the posterior and anterior part of the OB, respectively.
Middle. After substitution of the P2 coding sequence with *IRES-tau-lacZ*, the P2 neurons do not converge into a glomerulus (Wang et al.,1998).
Right. When replacing the coding sequence of P2 by that of M71, a topographically fixed yet distinctive from both donor and recipient glomerulus (orange) is formed in the OB (Feinstein and Mombaerts, 2004).

What supports the direct involvement of ORs in axon guidance is the presence of OR mRNAs as well as OR proteins in the olfactory axon termini prior to and during glomeruli formation (Ressler et al., 1994; Vassar et al., 1994; Barnea et al., 2004; Feinstein et al., 2004). The undoubtable necessity of ORs in axon guidance is elegantly demonstrated by these genetic manipulations, however the systematic failure of rewiring to the donor glomerulus in the swap experiments clearly indicates that OR itself is insufficient and additional OR-specific determinants are needed.

OR-derived cAMP signals specify transcription levels of axon guidance molecules.
How could the necessary role of ORs in axon guidance be mechanistically mediated? Given that deficiency in G_{olf} or CNG channels does not interrupt with normal axonal convergence in the OB, the canonical olfactory signal transduction is not assumed to be involved in the axon guidance process (Belluscio et al., 1998; Lin et al., 2000). However, in 2006 Imai et al. generated a DRY-motif (Asp-Arg-Tyr) mutant, in which the conserved tripeptide G-protein coupling site was abolished and so was any downstream signaling cascade. They have observed that OSN axons expressing the mutant OR wandered in the anterior part of the OB instead of converging onto their target glomerulus (Imai et al., 2006). Rescue experiments with constitutively active Gs, protein kinase A (downstream component of adenylyl cyclase), CREB (protein kinase A-

regulated transcription factor) could partially restore the glomerular projection (Imai et al., 2006). Their findings indicate that OR-derived cAMP signaling although via non-canonical pathways does influence the glomerular map formation.

Indeed, a number of guidance-related molecules have been identified responding to the cellular cAMP level, *e.g.*, Neuropilin1 (Nrp1) and its repulsive ligand Semaphorin3A (Sema3A) correlate positively or negatively with the cAMP level in the OSNs, exhibiting a complementary expression pattern. It has been observed that overexpression or knockout of *Nrp1* leads to a glomerular shift along the anterior-posterior axis (Imai et al., 2009). Similarly, conditional knockout of *Sema3A* also perturbs the glomerular map formation (Imai et al., 2009). The authors therefore proposed a model in which the OR-derived cAMP signals determine the anterior-posterior glomerular positioning via regulation of the transcription levels of axon guidance molecules; in short, OSNs with high cAMP level converge into glomeruli on the posterior OB, whereas OSNs with low cAMP level on the anterior OB (reviewed in Mori and Sakano, 2011). Later, they suggested that it is the basal ligand-independent OR activity that regulates the A-P axis targeting (Natashima et al., 2013). The origin of such endogenous basal activity is the spontaneous toggling of ORs between their active and inactive conformations in the absence of ligand binding.

For the dorsal-ventral glomerular projection, a substantially different mechanism has been proposed. The spatial organization along the dorsal-ventral axis in the OB is correlated with the corresponding positioning of the OSNs along the dorsomedial/ventrolateral axis in the OE (Miyamichi et al., 2005). The guidance receptor Neuropilin2 (Nrp2) and its repulsive ligand Semaphorin3F are found to be expressed in a complementary manner in the OE (Takeuchi et al., 2010). Based on these observations, Sakano et al. proposed a model for the dorsal-ventral targeting, in which the positional information within the OE decides both the receptor choice and its *Nrp2* expression level and thus the d/v mapping (reviewed in Mori and Sakano, 2011).

In sum, their models suggest an anterior-posterior patterning via the OR-specific spontaneous cAMP signaling and a dorsal-ventral patterning via positional information of the OSNs in the OE. Attractive as it seems to be, the Sakano-model cannot explain the systematic miswiring in the swap experiments. If OR-specific cAMP level truly determines the a/p mapping, OSNs with the swapped receptor should coalesce into the donor glomerulus, which was however never the case. The actual spontaneous activity level of different ORs is in fact not linearly related to the a/p mapping position (Natashima et al., 2013). There are many other unsolved questions, such as how could the intrinsic signaling machinery distinguish 1000 levels of activity? Or how might the external ligand-induced firing influence the postnatal cAMP level, since axon guidance takes place during the entire lifespan of the animal? Several publications so far have contradicted an anterior-posterior shift, when the *Nrp1/Sema3A* expression is impaired or when the basal G-protein signaling is interrupted (Assens et al., 2016; Movahedi et al., 2016; Zapiec et al., 2016).

Local sorting and refinement

It is known that a coarse olfactory map is already established during the prenatal stages. However, the local sorting and refinement, in terms of glomeruli maturation and removal of mixed OR innervations, takes place during the postnatal development (Zou et al., 2004). A variety of molecules has been identified for this process. Axel and colleagues first suggested that ephrin-As and EphAs could be possible candidates (Cutforth et al., 2003). OSNs expressing different receptors show different levels of ephrin-A expression in their axon termini, and the cognate receptors EphA3 and EphA5 are detected in target cells in the OB. Gain or loss of function of ephrin-As leads to an anterior or posterior shift of the P2 glomerulus in the OB, hinting at their involvement in local targeting. Subsequently, two cell adhesion molecules Kirrel2 and Kirrel3 from the immunoglobulin superfamily have been identified in the axon termini with a complementary expression pattern (Serizawa et al., 2006). Mosaic upregulation of Kirrel2 generates duplicated glomeruli, suggesting its role in glomerular segregation (Nakashima et al., 2019). Other molecules such as BIG-2/contactin-4, a glycoprotein from the immunoglobulin superfamily, have been shown to facilitate local sorting because of its mosaic expression pattern among different glomeruli and altered OB targeting upon BIG-2 knockout (Kaneko-Goto et al., 2008). And the protocadherin cell surface proteins, which are expressed by OSNs in distinct combinations, has been demonstrated to provide individual OSNs with surface diversity required for axon sorting and glomeruli formation (Mountoufaris et al., 2017). When the surface diversity is reduced by either deletion of all *Pcdh* gene clusters or overexpression of one isoform over the rest, defects of axon circuit assembly are observed. Interestingly, the olfactory marker protein (OMP), a marker for mature OSNs (Farbman and Margolis, 1980), seemingly is also needed for the fine-tuning of the local glomerular map, as in OMP null mice, OSNs expressing one receptor terminate at multiple glomeruli in close vicinity to their target glomerulus (Albeanu et al., 2018).

1.4.2 Searching for other olfactory guidance–related molecules

During olfactory axon guidance, the necessity of OR is fairly clarified and relates to its involvement in homophilic interactions. However, the constant failure in rewiring to the donor glomerulus with the swapped OR locus strongly suggests that OR itself is necessary but not sufficient to establish the correct olfactory mapping, and there must be other OR-specific guidance proteins.

We have aimed to identify such OR-specific guidance-associated molecules. After differential screening of single-cell cDNA library of outgrowing mOR37A-expressing OSNs by subtracting the transcripts expressed by mature mOR37A-expressing OSNs and outgrowing other receptor-expressing OSNs, we found a profound overrepresentation of a 3'-extended mOR37A transcript exclusively in the outgrowing mOR37A-expressing neurons (Haag, 2009). This 3'UTR extension

constitutes the only OR-specific transcriptomic difference between outgrowing and mature mOR37A neurons and is therefore likely to play a pivotal role in the developmental differentiation of these neurons. As the major process of developmental differentiation in the analyzed time window (GAP43$^+$.OMP$^-$ *vs.* GAP43$^-$.OMP$^+$) is axon targeting, it is assumed to be potential contributor in the guidance process. Further experimental data revealed that this 3'-extended transcript is likely to be part of the mOR37A mRNA, generated by alternative polyadenylation, in parallel to a much shorter transcript (Helisch, 2014). Such developmental regulation of different isoforms has been corroborated by fluorescence *in situ* hybridization and qPCR (Falk, 2015).

In the following section, some essential aspects of alternative polyadenylation and its post-transcriptional importance are summarized.

1.5 mRNA polyadenylation

First observed around 60 years ago (Edmond and Abrams, 1960), polyadenylation modifies nearly all protein encoding transcripts (mRNAs) and most long non-coding RNAs, by adding a stretch of non-templated adenosines to their 3' ends. It is a two-step nuclear reaction: an endonucleolytic cleavage of the nascent mRNA and the addition of poly(A) tail on its 3' terminus (reviewed in Tian and Manley, 2017). Typically the poly(A) tails have a more or less defined length, that differs between different species, *e.g.*, 250-300 adenines are synthesized in mammalian cells, whereas only 70-80 in yeast cells (Brawerman, 1981; Brown and Sachs, 1998). The sequence elements that direct 3' polyadenylation are present both upstream and downstream of the cleavage site (reviewed in Proudfoot, 2011). In vertebrates, the key signal component is a hexanucleotide motif called poly(A) signal (PAS), which comes in the canonical form of A[A/U]UAAA together with other close variants and is located 15-30nt upstream of the cleavage site. Apart from this, downstream U- and GU-rich sequences as well as upstream U-rich and UGUA sequences enhance the 3' end formation (Fig.1.6). The polyadenylation machinery includes some 20 core proteins (Fig.1.6) (reviewed in Colgan and Manley, 1997; Zhao et al., 1999; Mandel et al., 2008). Major components include four protein complexes and some additional proteins (Shi et al., 2009). The complexes are the cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF), cleavage factor I and II (CFI/CFII), additional proteins are poly(A) polymerase (PAP), poly(A)-binding protein II (PABII), scaffold protein simplekin and RNA polymerase II (RNA Pol II). In brief, CPSF recognizes and binds to the poly(A) signal, and initiates the cleavage. In a cooperative interaction with CstF and CFI, where CstF mainly binds to the downstream U- or GU-rich sequence and CFI recognizes the upstream elements, effective binding of the pre-mRNA substrate and cleavage is then achieved. Cleavage is also aided by CFII. After cleavage, polyadenylation starts by adding adenosine

residues to the RNA via the interaction of CPSF and PAP. An additional protein, PABII, supports the rapid elongation of the poly(A) tail and regulates its length to a maximal of 300 residues. Some other proteins such as symplekin and the regulatory C-terminal domain (CTD) of the RNA Pol II function as scaffold to recruit the 3' end processing machinery. Alternative polyadenylation (APA), a process in which a gene generates distinct mRNA isoforms by utilizing one of its multiple PASs, is emerging as an essential and conserved regulatory mechanism of gene expression
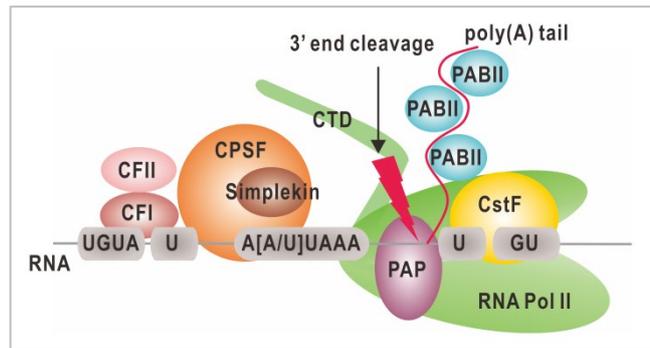
Figure 1.6. Sequence elements and core proteins involved in 3' end processing.
The canonical poly(A) signal is A[A/U]UAAA. In addition, there are other sequence elements including upstream U-rich and UGUA sequences, and downstream U- and GU-rich sequences. The major proteins for 3' cleavage and polyadenylation include cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF), cleavage factor I and II (CFI/CFII), poly(A) polymerase (PAP), poly(A)-binding protein II (PABII), simplekin and RNA polymerase II (RNA Pol II) and its C-terminal domain (CTD). Figure modified from Elkon et al., 2013.

(reviewed in Tian and Manley, 2017). In fact, 70% or more of the mammalian protein encoding genes undergo APA (Derti et al., 2012; Hoque et al., 2013). There are two types of APA, one with PAS in the internal exon/intron region, generating distinct protein isoforms (coding region-APA), and the other with PAS in the 3'UTR, producing transcripts with varying 3'UTR lengths but encoding the same protein (3'UTR-APA). Given the unique transcript structure of OR genes, only the 3'UTR-APA will be addressed in detail in this section.

1.5.1 Mechanism of 3'UTR alternative polyadenylation

Biased APA usage has been reported in different biological processes or among various tissues. In general, 3'UTR shortening has been observed in proliferating cells across many tissues and of different cell types (Sandberg et al., 2008; Elkon et al., 2012), whereas 3'UTR lengthening has been seen during embryonic development and differentiation (Ji and Tian, 2009). For instance, transcripts in the central nervous system tend to utilize distal PAS to generate longer mRNAs (Zhang et al., 2005). Similarly, neural-specific lengthening of 3'UTR is also reported during *Drosophila* embryogenesis and mouse embryonic stem cell differentiation (Hilgers et al., 2011; Shepard et al., 2011). On the contrary, activation of primary murine T lymphocytes increases the usage of upstream PAS and generates shorter 3'UTRs (Sandberg et al., 2008). In recent years, mechanisms underlying the active regulation of poly(A) site choice are emerging and will be briefly introduced in the following section.

Regulation by poly(A) machinery activity

Globally speaking, 3'UTR length is reversely correlated with the strength of the polyadenylation machinery in form of the expression level of the 3' end processing proteins. In other words, stronger polyadenylation machinery often leads to shorter 3'UTR and *vice versa*. For example, accompanying the 3'UTR lengthening during myoblast cells differentiation is the downregulation of genes which are part of the 3' processing complex, especially those from the CtsF protein complex (Ji et al., 2009). Similarly, in proliferating cells with 3'UTR shortening, the expression of this set of genes is upregulated (Elkon et al., 2012). Other observations include that a lower concentration of CstF2 is related to increased usage of the distal site and *vice versa* during B cell differentiation (Takagaki et al., 1996). As opposed to the example of CtsF, another cleavage-related protein CFI_m has the opposite effect, namely lower concentration of CFI_m results in elevated cleavage at the proximal site in cell cultures (Brown and Gilmartin, 2003). Thus, the activity of the poly(A) machinery could regulate APA choice by modulating the expression level of APA-specific proteins.

Regulation by transcriptional activity

A general correlation between transcription and APA usage is observed from some human tissues. Highly expressed genes tend to have short 3'UTR, whereas lowly expressed genes preferentially choose long 3'UTR (Ji et al., 2011). Transcriptional activity can be intertwined with APA regulation, given that, on the one hand, the CTD of RNA Pol II often functions as scaffold to recruit other poly(A) factors (Ahn et al., 2004; Adamson et al., 2005), and on the other hand, several transcription elongation factors such as ELL2 and Cdc73 have been proved to promote the usage of proximal site of specific genes due to their direct interactions with the CstF complex/CPSF complex (Martincic et al., 2009; Rozenblatt-Rosen et al., 2009). Another aspect of transcription-APA interaction deals with transcriptional kinetics. It has been shown that a reduced elongation rate of RNA Pol II in *Drosophila* leads to selection of proximal APA sites in a few transcripts (Pinto et al., 2011).

Interplay with alternative splicing

The interplay between alternative splicing and alternative polyadenylation, especially 3'UTR APA, has been elegantly demonstrated by RNA-seq in 15 human tissues and cell lines, in which the pattern of both is highly correlated (Wang et al., 2008). It might be attributed to the observation that the binding motifs of some splicing regulators such as FOX-1/FOX-2 are identified in the 3'UTR, hinting at a dual role of these proteins (Wang et al., 2008). Other examples of splicing factors involved in APA regulation include the Nova protein, which is among the first identified neuron-specific alternative splicing factors (Jensen et al., 2000). Apart from regulating neuronal pre-mRNA splicing, it has been shown to bind the 3'UTR of some brain

transcripts and regulate their APA. Notably, Nova does not possess regulation over the APA complex, but binds directly onto the poly(A) region, where it might antagonize the processing of polyadenylation machinery (Licatalosi et al., 2008). Another similar example is the neuronal ELVA protein, an mRNA processing regulator best known in *Drosophila* neurons (Robinow and White, 1991; Koushika et al., 2000). ELVA protein is needed for not only the splicing profile of erect wing mRNA in photoreceptors (Koushika et al., 2000), but also the poly(A) choice by binding in close proximity to the PAS (Hilgers et al., 2012). Similar to Nova, it has been assumed that ELVA promotes 3'UTR extension by masking the proximal APA site (Hilgers et al., 2012).

1.5.2 Function of 3'UTR alternative polyadenylation

In 3'UTR-APA, the polyadenylation machinery can utilize proximal or distal PAS, which might differ in their composition, to generate mRNA isoforms. Although the coding frame is not affected in 3'UTR-APA, significant changes regarding post-transcriptional regulations still occur due to the variation of the 3'UTR length. Such changes have been reported with regard to mRNA stability and translation efficiency, mRNA localization, protein complex formation and protein localization (Fig.1.7, reviewed in Tian and Manley, 2017).
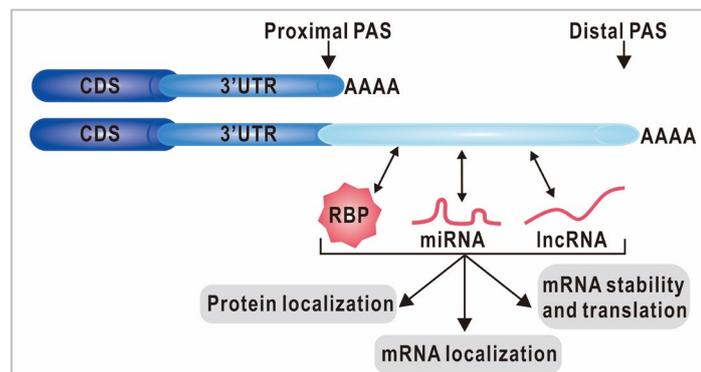
Figure 1.7. 3'UTR-APA and its functional relevance.
3'UTR-APA utilizes proximal or distal PAS to produce mRNA isoforms with distinct 3'UTR length. Interaction between the long isoform and RNA-binding proteins (RBP), microRNAs (miRNA) and long non-coding RNAs (lncRNA) can influence various aspects of gene expression, such as mRNA stability and translation, mRNA localization and protein localization. Figure modified from Tian and Manley, 2017.

mRNA stability and translation

The best studied regulatory function of 3'UTR-APA possibly arises from the interaction with microRNAs (miRNAs). miRNAs are short (~22nt) endogenous RNA molecules, which typically pair with target regions in the 3'UTR to deliver post-transcriptional repression (reviewed in Bartel, 2018). Notably, more than 50% miRNA targets are located after the proximal PAS in human (Legendre et al., 2006). Despite being a key component in the evolutionarily conserved RNA regulatory machinery, the exact silencing mechanism varies between plants and animals (reviewed in Carrington and Ambros, 2003; Millar and Waterhouse, 2005). Depending on the degree of complementarity, miRNA-mediated repression can be seen in two forms, mRNA

cleavage or translational arrest through the Argonaute protein silencing complex. Plant miRNA shares a near-perfect pairing with its target, leading to target cleavage by the Argonaute protein, whereas animal miRNA has imprecise pairing, resulting in poly(A) tail shortening and subsequent mRNA destabilization (reviewed in Bartel, 2018).

Moreover, 3'UTR harbors other sequence elements that could positively or negatively regulate mRNA stability, such as AU-rich elements, GU-rich elements or other stabilizing elements (reviewed in Garneau et al., 2007). Among them, as the best characterized example, AU-rich elements are known to be able to recruit RNA-decay machinery to specific transcripts, and thus mediate mRNA destabilization (reviewed in Barreau et al., 2005). In addition, 3'UTR Alu repeat has also been shown to interact with other Alu element in the cytosolic long non-coding RNA, which results in the formation of Staufen1 binding site and leads to the degradation of active mRNA (Gong and Maquat, 2011).

mRNA localization

The targeting of mRNA to distinct subcellular compartments can efficiently establish spatially restricted protein synthesis. A general mechanism of active and directed transport includes several steps: first, the recognition of *cis*-acting signals (zipcodes) mostly in the 3'UTR by specific RNA-binding proteins; second, the formation of ribonucleoprotein particles and finally the transport along the microtubules and actin filaments in the cytosol to its destination (reviewed in Martin and Ephrussi, 2009). One of the most cited and well understood cases of mRNA localization directed by 3'UTR-APA in the nervous system comes from the mRNA encoding the brain-derived neurotrophic factor (BDNF) (An et al., 2008). BDNF mRNA is polyadenylated at two different sites, generating mRNA isoforms with short or long 3'UTR. In hippocampal neurons, the short isoform is restricted to the soma while the long isoform also localizes in the dendrites. Long 3'UTR directed dendritic localization regulates local BDNF protein synthesis, and is important for dendritic spine morphogenesis.

Protein localization and protein complex formation

Interestingly, the 3'UTR can also act as a scaffold to regulate protein localization. One prominent example is the transmembrane protein CD47 (Berkovits and Mayr, 2015). *CD47* gene produces transcripts with alternative 3'UTRs. The long 3'UTR can recruit a protein complex including RNA-binding proteins HuR and SET to the translation site, which leads to the membrane localization of CD47 through a set of protein-protein interactions. The short 3'UTR, however, lacks the core sequence for HuR-SET assembly and the CD47 protein translated from it stays primarily in the ER.

Another example of 3'UTR-mediated protein complex formation comes from the BIRC3 protein in malignant B cell leukemia (Lee and Mayr, 2019). BIRC3 mRNA is alternatively

18

polyadenylated. Only the BIRC3 protein translated from the mRNA with long 3'UTR can form a functional protein complex with other interactors recruited by its 3'UTR. The specific protein complex in this case enables additional 3'UTR-dependent functions, such as trafficking of the receptor CXCR4 to cell surface and eventually the promotion of B cell migration.

1.6 Aims of the thesis

Before this work was started, a rigorously controlled differential single-cell cDNA library screen had been conducted for transcripts obeying the logic "expressed in mOR37A$_{immature}$ *not* in (mOR37A$_{mature}$ *or* ORX$_{immature}$)" in order to find OR-specificity associated axon guidance molecules. To assure the specificity of the screen, each individual step of the procedure was thoroughly optimized. This included the optimization of the synthesis of length-controlled cDNA-tags, their efficient tailing, unbiased PCR-amplification of the tag-library, exhaustive cloning and sensitive radioactive screening. As OSNs are extremely similar, presumably differing only in a handful out of more than ten thousand transcripts, the probability of retrieving non-specific, shared gene-products in such screen is extraordinarily high. Nonetheless, the screen yielded a single transcript in >86% of the clones, which was expressed exclusively in mOR37A$_{immature}$ neurons, testifying to its exquisite specificity (Haag, 2009). Notably, this single transcript extended in 5' direction towards mOR37A coding region and overlapped with a short transcript isoform of the mOR37A mRNA. An upstream primer inside the transcript for 5' RACE only yielded the 5' end of the mOR37A mRNA, indicating that it might be a long transcript isoform from the mOR37A mRNA, generated by alternative polyadenylation (Helisch, 2014). Examined by RNA-FISH, this transcript was never expressed without the mOR37A coding region, further corroborating the assumption that it was an elongated mRNA isoform (Wunderlich, 2016). Furthermore, it exhibited a maturation-dependent regulation in the mOR37A-expressing cells, namely its expression was highly restricted to the GAP43$^+$ immature population. During development, the cells expressing this transcript slowly vanished (Falk, 2015). Based on these findings, the major aims of this thesis are to characterize the exact structure of mOR37A transcripts, and to elucidate any potential functional relevance of different isoforms during the olfactory axon guidance process.

Detailed aims of this thesis are to,

❖ clarify the transcript structure of receptor mOR37A, especially its 3' ends.

❖ verify the spatial and temporal expression profile of different mOR37A isoforms.

❖ investigate any potential link between alternative polyadenylation and alternative splicing in mOR37A transcripts, since the activity between both machineries is often intertwined and co-regulated.

❖ explore potential roles of different mOR37A isoforms during the axon outgrowth phase in the axon terminal.

❖ extend the knowledge gained from mOR37A to other receptors of mOR37C, olfr701 and olfr702, with regard to the transcript structure and expression profiles of different isoforms.

2. Materials and Methods

2.1 Materials

2.1.1 Animals

Both wild type C57BL/6 and transgenic mOR37A-IRES-tauEGFP mouse strains were bred and maintained under conventional husbandry conditions by the Mouse Facility at Campus North, Karlsruhe Institute of Technology. Mice of three different developmental stages were used in this thesis, namely postnatal day 1 (P1), postnatal day 7 (P7) and adult. All experiments complied with the German Animal Welfare Act.

2.1.2 Bacterial strain and plasmid

The electrocompetent *E. Coli* used in this work was the XL1-Blue strain (Agilent Technologies). To prepare the electrocompetent cells, steril LB broth was inoculated with small aliquot of bacteria and cultured overnight. After harvesting the bacterial mass, it was resuspended and washed multiple times in ice-cold steril H_2O. Finally, the cell pellet was suspended in ice-cold 10% glycerol, aliquoted, snap frozen in liquid nitrogen and stored at -80ºC.

The cloning plasmid used in this work was the pBluescript I KS vector (Agilent Technologies). It contains a polylinker of unique restriction enzyme recognition sites flanked by T3/T7 RNA polymerase promoter sequences. For cloning purposes, target DNA fragment was inserted into the SacI and KpnI digested vector. For colony selections, the vector-encoded ampicillin resistance marker gene was used.

2.1.3 Chemicals

All standard chemicals were, if not stated otherwise, purchased in analytical grade from Carl Roth (Carl Roth GmbH & Co. KG, Karlsruhe, Germany). Other experiment-specific chemicals, especially for RNA-related work, are listed in Tab.2.1.

Table 2.1. Other experiment-specific chemicals

Name	Details	Manufacturer
3-(N-Morpholino)propanesulfonic acid (MOPS)	≥ 99.5%	Fluka
4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI)	≥ 98%	Carl Roth
Acetic anhydride	≥ 99%	Fluka
Agarose, LE, Analytical Grade	Powder	Promega
Agarose NEEO ultra-quality	Powder	Carl Roth
Bolton Hunter Reagent	Sulfosuccinimidyl-3 – (4-hydroxypheynyl) propionate	Thermo Fisher Scientific
Bovine serum albumin (BSA), Fraction V	approx. 99%	Sigma Aldrich

Table 2.1. Other experiment-specific chemicals (continued)

Name	Details	Manufacturer
Blocking Reagent	Powder	Roche
CDP-*Star* Chemiluminescent Substrate	0.25mM in H_2O	Sigma Aldrich
Diethyl pyrocarbonate (DEPC)	$\geq 97\%$	Carl Roth
Diethylenetriaminepentaacetic acid (DETAPAC)	$\geq 99\%$	Sigma Aldrich
Formaldehyde solution	37%	Carl Roth
Formamide	$\geq 99.5\%$	Sigma Aldrich
n-Propyl gallate	$C_{10}H_{12}O_5$	Sigma Aldrich
Paraformaldehyde	Powder, 95%	Sigma Aldrich
Phenol/Chloroform/Isoamyl alcohol	25/24/1, pH 7.5-8.0	Carl Roth
Polyvinylpyrrolidone K 15	Powder	Fluka
Repel-silane ES	2% dimethyldichlorosilane in octamethylcyclooctasilane	GE Healthcare
Roti-GelStain	n.a.	Carl Roth
Sodium dodecyl sulfate (SDS)	$\geq 99\%$	Fluka
Sodium N-lauroylsarcosinate	Powder	Fluka

2.1.4 Solutions, media and buffers

Distilled deionized water (ddH$_2$O) was purified with the Veolia Water System (Veolia Water Technologies). H$_2$O was used as solvent, if not explicitly stated otherwise. For RNA applications, all solutions were either directly treated with 0.1% DEPC (diethyl pyrocarbonate) and subsequently autoclaved, or dissolved with DEPC treated H$_2$O. All solutions, media and buffers are listed in Tab.2.2.

Table 2.2. Solutions, media and buffers

Name	Concentration	Components
Agarose gel	0.5%-2% (w/v)	Agarose NEEO in 1x TAE
BHR	20mM	Bolton Hunter Reagent in DMSO
Blocking reagent	2% (w/v)	Blocking Reagent in PBST pH 7.4
Borate buffer (1x)	0.1M	Boric acid pH 8.5
Denhardt's (50x)	1% (w/v) 1% (w/v) 1% (w/v)	Ficoll 400 Polyvinylpyrrolidone K15 Bovine serum albumin Fraction V
DETAPAC	10mM	Diethylenetriaminepentaacetic acid pH 7.4
EDTA	0.5M	Disodium EDTA • 2H$_2$O pH 8.0
Glucose (mutarotated)	1.5M	D(+)-Glucose monohydrate in Borate buffer let stand at RT O/N
HCl	0.2M	Hydrochloric acid (32%)
LB agar	1.5% (w/v) 100µg/ml	Agar-Agar in LB medium Ampicillin
LB medium	1% (w/v) 0.5% (w/v) 0.5% (w/v) 100µg/ml	Trypton NaCl Yeast extract pH 7.4 Ampicillin

Table 2.2. Solutions, media and buffers (continued)

Name	Concentration	Components
LiCl	4M	Lithium chloride
MaBS (1x)	0.1M	Maleic acid
	0.15M	NaCl
		pH 7.5
MOPS (10x)	0.2M	3-(N-Morpholino)propanesulfonic acid
	0.05M	NaAc
	0.01M	EDTA
		pH 7.0
Mowiol	0.1g/ml	Mowiol 4-88
	25% (w/v)	Glycerol
	0.1M	Tris/HCl (pH 8.5)
	trace	n-Propyl gallate
NaAc	3M	Sodium acetate
		pH 5.2
PBS (1x)	137mM	NaCl
	2.7mM	KCl
	8.1mM	$Na_2HPO_4 \cdot 2H_2O$
	1.8mM	KH_2PO_4
		pH 7.4
PBST (1x)	0.05% (v/v)	Tween-20 in 1x PBS
PFA	4% (w/v)	Paraformaldehyde in 1x PBS
Pre-hybridization solution	50% (v/v)	Formamide
	5x	SSC
	5x	Denhardt's
	0.5% (v/v)	SDS
Reaction buffer (1x)	0.1M	Tris
	0.15M	NaCl
	5mM	$MgCl_2$
		pH 9.5
SDS	10% (w/v)	Sodium dodecyl sulfate
SOC medium	2% (w/v)	Trypton
	0.06% (w/v)	NaCl
	0.5% (w/v)	Yeast extract
	0.02% (w/v)	KCl
	10mM	$MgCl_2$
	10mM	$MgSO_4$
SSC (20x)	3M	NaCl
	0.3M	Trisodium citrate
		pH 7.0
Sucrose	30% (w/v)	Sucrose in 1x PBS
TAE (10x)	400mM	Tris
	200mM	Acetic acid
	10mM	EDTA
		pH 8.0
TBE (5x)	450mM	Tris
	450mM	Boric acid
	10mM	EDTA
		pH 8.0
TEA	0.1M	Triethanolamine
		pH 8.0
Transfer solution	1M	Ammonium acetate
Wash buffer	0.3% (v/v)	Sodium N-lauroylsarcosinate in 1x MaBS

2.1.5 Fluorescent labels

The fluorescent labels Alexa488/Cy3/Cy3B/-methyldopamine were synthesized from Alexa Fluor 488-NHS (Thermo Fisher Scientific), Cy3-NHS (Lumiprobe), Cy3B-NHS (Thermo Fisher Scientific) and 3-Methoxytyramine (Sigma Aldrich) in our lab, with a final concentration of 10mM (unpurified). They were stored at -20ºC and protected from light.

2.1.6 Enzymes, antibodies and (ribo)nucleic acids

All enzymes, antibodies and (ribo)nucleic acids (Tab.2.3) were reconstituted (if needed) and stored referring to manufacturer's guidelines.

Table 2.3. Enzymes, antibodies and (ribo)nucleic acids

Name	Stock	Manufacturer
Anti-Digoxingenin-AP, Fab fragments	150 U (200µl)	Roche
Anti-Digoxigenin-POD, Fab fragments	150 U/ml	Roche
Anti-Fluorescein-POD, Fab fragments	150 U/ml	Roche
ddPCR Supermix for Probes (2x)	ddPCR reaction cocktail	Bio-Rad
DreamTaq DNA Polymerase	5 U/µl	Thermo Fisher Scientific
Glucose oxidase	1000 U/ml	Sigma Aldrich
HotStart Taq DNA Polymerase with Antibody	5 U/µl	Genaxxon
KAPA SYBR FAST qPCR Master Mix (2x)	qPCR reaction cocktail	Sigma Aldrich
Phusion High-Fidelity DNA Polymerase	2 U/µl	Thermo Fisher Scientific
Proteinase K	Powder	Roche
Restriction enzymes (SacI, KpnI etc.)	10 U/µl	Thermo Fisher Scientific
RiboLock RNase Inhibitor	40 U/µl	Thermo Fisher Scientific
RNase A	10 mg/ml	Thermo Fisher Scientific
RNase H	5 U/µl	Thermo Fisher Scientific
Shrimp Alkaline Phosphatase	1000 U/ml	New England Biolabs
SMARTScribe Reverse Transcriptase	100 U/µl	Clontech
SuperScript IV Reverse Transcriptase	200 U/µl	Thermo Fisher Scientific
T3 RNA Polymerase	20 U/µl	Roche
T4 DNA Ligase	5 U/µl	Thermo Fisher Scientific
T4 DNA Polymerase	5 U/µl	Thermo Fisher Scientific
T7 RNA Polymerase	20 U/µl	Roche
TURBO DNase	2 U/µl	Thermo Fisher Scientific
Ribonucleotides (A,U,G,C)	100mM	Carl Roth
DIG RNA Labeling Mix	10x conc.	Roche
Fluorescein RNA Labeling Mix	10x conc.	Roche
Ribonucleic acid from torula yeast, Type VI	Powder	Sigma Aldrich
RiboRuler High Range RNA Ladder	n.a.	Thermo Fisher Scientific
Ribonucleic acid, transfer from baker's yeast	Powder	Sigma Aldrich
Deoxyribonucleotides (A,T,G,C)	100mM	Carl Roth
Oligo dT (20) Primer	50µM	Invitrogen
peqGOLD DNA Ladder Mix	0.5mg DNA/ml	VWR
Random hexamers	3 µg/µl	Invitrogen

2.1.7 Molecular biology kits

The storage and usage of the following molecular biology kits (Tab.2.4) was in accordance with manufacturer's protocols.

Table 2.4. Molecular biology kits

Name	Manufacturer
DNA Clean & Concentrator^{TM}-5	Zymo Research
miRNeasy Mini Kit	Qiagen
QIAprep Spin Miniprep Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
Qubit^{TM} RNA BR Assay Kit	Thermo Fisher Scientific
RNeasy MinElute Cleanup Kit	Qiagen
SMARTer RACE 5'/3' Kit	Clontech

2.1.8 Other consumables

Table 2.5. Other consumables

Name	Details	Manufacturer
Coverslips	24 x 60 mm	Carl Roth
Nylon membranes	positively charged	Roche
Parafilm	n.a.	Bemis
SuperFrost Plus Adhesion slides	n.a.	Thermo Fisher Scientific
Whatman 3MM filter paper	n.a.	GE Healthcare

2.1.9 Hard- and software

Table 2.6. Hard- and software

Name	Manufacturer
Amersham Imager 600	GE Healthcare
C1000 Touch Thermal Cycler	Bio-Rad
Cryostat CM3050 S	Leica
E.coli Pulser	Bio-Rad
Dissection tools	Fine Science Tools
Hybridization incubator 7601	GFL
Nanodrop 1000 Spectrophotometer	Peqlab
Qubit 3 Fluorometer	Thermo Fisher Scientific
QX200 Automated Droplet Generator	Bio-Rad
QX200 Droplet Reader	Bio-Rad
Rotor-Gene Real Time PCR Thermocycler	Qiagen
Thermocycler peqSTAR 2x	Peqlab
UV Crosslinker Stratalinker 1800	Stratagene
Zeiss ApoTome	Carl Zeiss
Zeiss LSM 800 with Airyscan	Carl Zeiss
CorelDRAW 2018	Corel
Fiji is just ImageJ	Fiji
Origin 2016	OriginLab
QuantaSoft Software	Bio-Rad
Rotor-Gene 6000 Series Software	Qiagen
SnapGene Viewer	SnapGene
ZEN (blue edition)	Carl Zeiss

2.2 Methods

2.2.1 Tissue and RNA sample preparation

For the purpose of tissue dissection, juvenile animals (P1, P7) were directly decapitated, and adult animals were first deeply anesthetized with CO_2 before decapitation. The skin was carefully removed with scissors to expose the entire skull.

Dissection for total RNA isolation

For olfactory epithelium (OE) total RNA isolation, the nose was first cut midsagittally into two halves through the nasal bone. The septum was then carefully removed if it was still attached to the underlying tissue. The OE in the nasal cavity could be easily identified by its slightly yellow color and unique folding structure. Nasal turbinates II and II' (for receptor subfamily mOR37) or all four turbinates (for other receptors) were cut along the border and underneath with a pair of fine scissors, and were carefully transferred to a petri dish filled with ice-cold 1x PBS, where the turbinate bones were removed. The same procedure was repeated with the epithelium from the other half. The two pieces of epithelia from one animal were pooled in a Potter-Elvehjem homogeniser and homogenized thoroughly in 1ml TRIzol reagent (Thermo Fisher Scientific). The homogenate could be directly subject to RNA isolation or be stored at -20ºC before further use.

For olfactory bulb (OB) total RNA isolation, the head was first cut open at the posterior end (approximately between cerebellum and medulla). The skull was incised on both lateral sides along the superficial temporal vein, till the incision reached the dorsal side. The dorsal surface of the skull was then slightly scratched with a scalpel at the level of the anterior end of the olfactory bulb. This entire piece of skull was then carefully uplifted and removed, exposing the olfactory bulb beneath it. The olfactory bulb together with the rest of the brain was carefully removed from the skull base to ice-cold 1x PBS. Both bulbar halves were harvested with one single cut, and were transferred to a Potter-Elvehjem homogeniser and homogenized thoroughly in 1ml TRIzol reagent (Thermo Fisher Scientific). The homogenate could be directly subject to RNA isolation or be stored at -20ºC before further use.

Total RNA isolation and DNase treatment

Total RNA isolation was performed with TRIzol reagent according to manufacturer's protocol (Thermo Fisher Scientific). The final RNA pellet was dissolved in 30-50µl RNase-free H_2O (Qiagen), and the concentration was assessed with the Nanodrop spectrophotometer (Peqlab).

The isolated RNA was then subject to DNase treatment to remove any possible genomic DNA contamination. TURBO DNA-*free* kit (Thermo Fisher Scientific) was used for this purpose, referring to manufacturer's protocol. To further concentrate the RNA, ethanol precipitation was

done with 0.1 volume of 3M NaAc (pH 5.2) and 3 volumes of ice-cold 100% ethanol. RNA was either precipitated at -80ºC for over 1h or at -20ºC overnight. It was then centrifuged, washed with 70% ethanol, air dried and dissolved in 30µl RNase-free H$_2$O (Qiagen). The final concentration was measured with Qubit 3 Fluorometer with the Qubit RNA BR Assay Kit (Thermo Fisher Scientific), following manufacturer's instructions. For long term storage, RNA was snap frozen in liquid nitrogen and stored at -80ºC.

To control the integrity of purified RNA, gel electrophoresis was performed. Under most circumstances, non-denaturing agarose gels in 1x TAE with Roti-GelStain (Carl Roth) were used. RNA samples were denatured in 50% formamide for 3min at 99ºC, chilled on ice and electrophoresed at 7.5V/cm for 45min. For specific applications, especially when the RNA samples should be well separated or the size should be precisely determined, formaldehyde denaturing gel electrophoresis was performed. The agarose powder was dissolved with 1.5% formaldehyde in 1x MOPS buffer, and the RNA samples as well as the RNA ladder were denatured in 50% formamide and 5% formaldehyde in 1x MOPS buffer for 15min at 55ºC. The gel was loaded and run in the chemical hood with 1x MOPS/0.6% formaldehyde buffer at 5.5V/cm for 2.5h.

Preparation of miRNA-enriched fractions (<200nt)

To separate the miRNA-enriched fraction from OB total RNA, both the miRNeasy Mini Kit and the RNeasy MinElute Cleanup Kit (Qiagen) were used in accordance with manufacturers' protocols. The isolated small RNA molecules (<200nt) were snap frozen in liquid nitrogen and stored at -80ºC.

Dissection for RNA *in situ* hybridization

For RNA *in situ* hybridization on juvenile animals, the tip of the nose of the skinned skull was first sliced off with a pair of scissors, so that the entry of the nasal cavity was exposed. The olfactory epithelium was cryoprotected by injecting 30% sucrose into both nasal cavities. Afterwards, the whole head was immersed in 30% sucrose solution in a Falcon tube for 30min at 4ºC. The head was deep frozen with liquid nitrogen vapor before slowly immersion in liquid nitrogen. Long term storage was at -80ºC.

The preparation for adult animals differed slightly. The maxillary incisors were first removed. The whole head was then fixed in 4% PFA overnight at 4ºC, followed by a two-day incubation in 0.5M EDTA (pH 8) for decalcification and one overnight incubation in 30% sucrose at 4ºC (including sucrose injection into the nasal cavities). The head was deep frozen with liquid nitrogen vapor before slowly immersion in liquid nitrogen. Long term storage was at -80ºC.

Cryosection of tissue slices

Coronal sections of the OE were acquired from frozen P7/adult mouse heads with a cryostat (Leica). The sectioning thickness was 16µm and the cutting temperature was set to -28ºC. The sections were transferred to microscopic slides (SuperFrost Plus, Thermo Fisher Scientific) in three alternating series and baked at 55ºC for 3h before RNA *in situ* hybridization. For long term storage, the sections were placed in a -80ºC freezer.

2.2.2 Reverse transcription (RT) and polymerase chain reaction (PCR)

cDNA synthesis was carried out following the protocol of SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific). For different applications, random hexamers (Invitrogen), Oligo dT(20) primer (Invitrogen) or gene specific primers were used. As a negative control, reverse transcription was carried out with all components except the reverse transcriptase. In conventional PCR and quantitative real-time PCR applications, the first strand cDNA was directly used as template without further purification. In quantitative ddPCR applications, especially for OB cDNA samples, spin column purification (Zymo Research) was performed to remove dNTPs and other reaction remnants from reverse transcription. cDNA concentration after purification was measured with a Nanodrop spectrophotometer (Peqlab). For different research purposes in ddPCR, other treatments such as RNase treatment or Proteinase K treatment was also administrated, which are described in detail in section 2.2.3.

In this work, three different DNA polymerases were used for conventional RT-PCR applications, namely the Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific), the DreamTaq Hot Start DNA Polymerase (Thermo Fisher Scientific) and the HotStart Taq DNA Polymerase with Antibody (Genaxxon). PCR reaction setup was in accordance with the manufacturers' guidelines. PCR products were separated and visualized by agarose gel electrophoresis. If not stated otherwise, agarose gel electrophoresis was conducted in 1x TAE and stained with Roti-GelStain (Carl Roth), with agarose concentration from 0.5% to 2%.

Quantitative real-time PCR (qPCR)

To quantify the relative abundance of different transcript isoforms, qPCR was performed with KAPA SYBR FAST qPCR Master Mix (2x) Kit (Sigma Aldrich) in a real-time PCR cycler (Rotor-Gene 6000, Qiagen). Typically two sets of primers were designed in the proximal (before 1st polyA) and the distal part (after 1st polyA) of the transcripts, respectively. The relative abundance of different isoforms can be reflected by the distal/proximal ratio.

Random hexamer primed cDNA without purification was used as template and the input amount referred to the corresponding RNA amount in reverse transcription, assuming an efficiency of RNA to cDNA conversion of 100%. As standards, linearized plasmids containing the

corresponding target sequence were made into dilution series with known copy numbers (10^7, 10^5, 10^4, 10^2). Negative controls with no reverse transcriptase, as well as non-template controls with H_2O instead of cDNA template were prepared. For both cDNA samples and plasmid standards, typically five replicates were made. qPCR setup was summarized in Tab.2.7. The reaction was performed according to manufacturer's instructions in the absence of ROX reference dye, with a 3-step protocol: 95ºC for 5min, 40 cycles of 95ºC for 5s, T_m[1)] for 20s, and 72ºC for 10s. This was followed by the dissociation cycle (72ºC-95ºC) for melt curve analysis. Final data analysis was achieved with Rotor-Gene 6000 Series Software (Qiagen). The *slope correction* function was activated to determine the background fluorescence noise level and *Auto-Find Threshold* was applied to calculate the Ct value. The standard curve derived from plasmid serial dilutions was used to quantify the absolute copy number of the unknown targets. Both the proximal and distal targets were quantified from the same cDNA set independently and the ratio of them was calculated.

Table 2.7. Typical setup for SYBR FAST qPCR

Sample number	Sample type	Sample concentration / reaction
1-5	cDNA	30ng
6-8	Negative control	/
9-11	Non-template control	/
12-16	pBluescript with target sequence	10^7 calculated copies
17-21	pBluescript with target sequence	10^5 calculated copies
22-26	pBluescript with target sequence	10^4 calculated copies
27-31	pBluescript with target sequence	10^2 calculated copies

[1)] Melting temperature was calculated by the common formula: Tm=2*(numbers of A's and T's)+4*(numbers of G's and C's). Final annealing temperature was set based on gradient PCR with temperature range of Tm $\pm$ 10ºC.

Droplet Digital PCR (ddPCR)

To achieve sensitive detection of different transcript isoforms especially in the axon termini, ddPCR was performed (ddPCR Supermix for Probes, Bio-Rad), using the QX200 Droplet Digital PCR system (Bio-Rad). Similar to qPCR, two sets of primers were chosen in the proximal (before 1st polyA) and the distal part (after 1st polyA) of the transcripts, respectively. Two TaqMan probes (biomers.net) labeled with FAM or HEX reporter fluorophores were designed within each amplicon, so that both targets could be quantified in one reaction. The relative abundance of different isoforms can be reflected by the distal/proximal ratio.

Random hexamer primed cDNA was used as template. As a general guideline, cDNA input from OE was 20ng-100ng per reaction and from OB was 200ng-1000ng per reaction. For cDNA samples without purification, the input amount referred to the corresponding RNA amount in reverse transcription and for samples with purification, it referred to the Nanodrop measured concentration. Negative controls with no reverse transcriptase were included as well. No replicates were performed. The reaction setup was in accordance with manufacturer's instructions, with a 3-step protocol: 95ºC for 5min, 40 cycles of 94ºC for 30s, 54ºC for 1min and

58ºC for 1min, 98ºC for 10min and final hold at 4ºC. The ramp rate was set to 2ºC/s. After PCR, data acquisition and analysis were accomplished with QuantaSoft Software (Bio-Rad). The fluorescence amplitude threshold, separating the positive and negative droplets was set manually. Similarly, the proximal and distal targets were quantified from the same sample and the ratio of them was calculated.

2.2.3 Experimental procedure for the analysis of potential 3'UTR interaction partners

RNase treatment (section 3.3)

RNase treatment was carried out in some cases to remove the RNA remnants after cDNA synthesis. RNase treatment included RNase H (Thermo Fisher Scientific) and RNase A (Thermo Fisher Scientific). A typical protocol would be, cDNA from 4µg OB total RNA was incubated with 5U RNase H at 37ºC for 10min followed by 1µg RNase A for another 10min. RNase-treated cDNA was purified with spin columns (Zymo Research).

Cloning of mOR37A short and long 3'UTR into a cloning vector (section 3.3.2)

In order to establish a synthetic system with both proximal and distal targets present in a controllable manner, the short and long 3'UTR of mOR37A were cloned into a cloning vector (plasmid backbone: pBluescript KS) via cohesive-end ligation. The short 3'UTR was 1438nt and the long 3'UTR was 4727nt (Tab.2.8). For the short 3'UTR, the same cloning procedure was performed as in producing RNA antisense probes (see section 2.2.4 for details), whereas for the long 3'UTR, minor modifications were adopted due to its internal SacI and KpnI restriction sites (Fig.2.1). The PCR product of the long 3'UTR was first digested with EcoRI (Thermo Fisher Scientific) and the two resulting fragments were separated by agarose gel electrophoresis, excised from the gel (QIAquick Gel Extraction Kit, Qiagen) and purified (Zymo Research). Fragment 1 and 2 were then digested with SacI and KpnI (Thermo Fisher Scientific) respectively, and purified subsequently (Zymo Research). Both fragments were ligated into the vector in a 5:5:1 molar ratio, and the ligation reaction was incubated at 4ºC for 20min, 16ºC for 2h and RT for 20min. The correct insertion was confirmed by sequencing (LGC Genomics). *In vitro* transcription was driven by the T7 promoter with unlabeled ribonucleotides (Carl Roth), following otherwise the same protocol as with producing DIG/Fluorescein labeled RNA probes. The final RNA products were examined by RNA formaldehyde denaturing gel.

For ddPCR applications with pure synthetic RNA, random primed and purified cDNA from the short and the long RNA was mixed at equal mass ratio of 0.1fg or 2fg per reaction. For ddPCR applications with zebrafish total RNA, cDNA was generated by reverse transcription of 4µg zebrafish total RNA together with 1.3fg long RNA and 2fg short RNA, and 100ng-400ng purified cDNA was used as template per reaction. ddPCR was carried out as described.
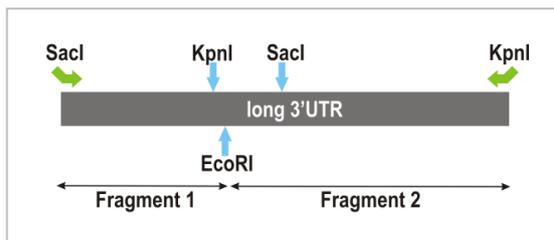
Figure 2.1. Schematic illustration of restriction sites in the long 3'UTR of mOR37A. Two primers flanking the long 3'UTR were designed with SacI and KpnI restriction sites built in. The fragment contains otherwise internal SacI and KpnI sites.

Table 2.8. Information on synthetic RNA used in ddPCR

	Length	Direction	Copy number / ng
Short 3'UTR	1438nt	Sense	1.27×10^9
Long 3'UTR	4727nt	Sense	3.86×10^8

Proteinase K treatment (section 3.3.3)

Proteinase K treatment was performed on non-purified OB cDNA. To ease the subsequent pipetting steps, 80µl H_2O was added to the 20µl synthesized cDNA for a final volume of 100µl. The whole reaction mix was incubated with 0.5µl Proteinase K (stock: 20mg/ml, Roche) and 1µl SDS (stock: 10%) for 30min at 55ºC. Subsequently, phenol/chloroform extraction was performed to isolate the cDNA by applying one volume of phenol/chloroform/isoamyl alcohol (Carl Roth) to the sample and vortexing shortly. After centrifugation for 2min at 13.000g, the upper phase was transferred to a new tube and was reextracted by adding one volume of chloroform and vortexing briefly. It was then centrifuged for another 2min at 13.000g and the upper phase was transferred to a new tube. The reextraction was repeated one more time. To further concentrate the cDNA, ethanol precipitation was conducted with 0.1 volume of 3M NaAc (pH 5.2) and 2 volumes of ice-cold 100% ethanol for over 1h at -80ºC. Proteinase K treated and EtOH purified cDNA could be either directly applied as template in ddPCR or subject to RNase treatment and spin column purification. ddPCR was carried out as described.

2.2.4 Rapid amplification of cDNA ends (RACE) and Southern blot (SB)

To obtain the 5' and 3' end sequence of OR transcripts, RACE-ready cDNA was generated with SMARTer RACE 5'/3' Kit according to manufacturer's protocol (Clonetech). The subsequent RACE-PCR was performed with Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific), applying a touchdown PCR program with different cycle numbers (20x, 25x, 30x) (Tab.2.9). When combined with Southern blot (see detailed Southern blot protocol in Appendix), the RACE-PCR products were loaded in a 1x TBE agarose gel. The electrophoresis-separated DNA fragments in the gel were first depurinated in 0.25M HCl for 5min and denatured in 0.5M NaOH/1.5M NaCl for 30min on a shaker platform before incubation in 1M NH_4Ac transfer solution for 30min. The transfer procedure was done via downwards capillary transfer with 1M NH_4Ac to a positively charged nylon membrane (Roche) for 2h. The gel size and slot positions were marked with a syringe needle. The DNA fragments were then immobilized to the membrane by UV crosslinking with 70000µJ and the membrane was incubated in the pre-hybridization

solution for 1h before overnight hybridization with the DIG-RNA probe at the corresponding hybridization temperature. In the following day, after strigency washes with pre-hybridization solution for 2x 10min and 0.1x SSC/0.1% SDS for 2x 15min at the hybridization temperature, the membrane was rinsed with wash buffer, blocked with 1% blocking reagent and subject to Anti-Digoxigenin-Ab antibody detection for 30min on a shaker platform. Subsequently, the membrane was rinsed with wash buffer and equilibrated in reaction buffer shortly. By incubating in 1ml CDP-Star (1:20 dilution in reaction buffer) for 5min, the membrane was ready to be imaged in the chemiluminescent mode in Imager 600 (GE Healthcare). The size of the detected bands was calculated by calibration to the original agarose gel. When not combined with Southern blot, the RACE-PCR products were separated in 1x TAE agarose gel.

As the next step, nested PCRs were performed on either the entire RACE-PCR reaction mix or the excised bands (QIAquick Gel Extraction Kit, Qiagen) identified by Southern blot, with a gene specific primer and the nested universal primer (NUP). The nested PCR products were purified (Zymo Research) and submitted to DNA sequencing (LGC Genomics). The resulting sequences were mapped to the mouse genome in the Ensembl database (Release 97).

Table 2.9. Touchdown program for Phusion RACE-PCR

Step	Temperature	Time	Cycles
Initialization	98ºC	30s	1x
Denaturation	98ºC	10s	
Annealing	[1]Tm+3ºC	30s (-0.5ºC/cycle)	6x
Extension	72ºC	30s/kb	
Denaturation	98ºC	10s	
Annealing	[1]Tm	30s	20x/25x/30x
Extension	72ºC	30s/kb	
Final extension	72ºC	10min	1x
Final hold	4ºC	hold	

[1]Annealing temperature is calculated from ThermoFisher Tm calculator (www.thermofisher.com /tmcalculator).

2.2.5 Fluorescence *in situ* hybridization (FISH)

Production of RNA antisense probes

In this work, DNA templates for *in vitro* transcription of labeled RNA probes were generated by cloning and plasmid preparation (plasmid backbone: pBluescript KS). RNA probes were of a typical length from 500nt to 1000nt.

A pair of PCR primers was designed flanking the target sequence, with SacI and KpnI restriction sites built in at the 5' ends, respectively. The target sequence should therefore be devoid of SacI and KpnI restriction sites. DNA template was generated by PCR with Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific) and purified with spin columns (Zymo Research). Both the PCR product and the plasmid backbone were double digested with SacI and KpnI at manufacturer's suggestions (Thermo Fisher Scientific), followed by dephosphorylation of the vector with Shrimp Alkaline Phosphatase (New England Biolabs) and purification of both with

spin columns (Zymo Research). The target DNA was then ligated into the transcription vector, downstream of an RNA polymerase (T7/T3) promoter site, with T4 DNA ligase (Thermo Fisher Scientific) applying a 5:1 molar ratio of insert DNA to vector. The ligation was incubated either at RT for 1h or at 16ºC overnight. Competent *E. coli* cells (XL1-Blue strain) were transformed with 1ng ligation product by electroporation at 2500V. After incubating for 30min in SOC medium, 100µl of the transformation was plated onto prewarmed LB agar plates with three dilution series and incubated overnight. In the following day, the correct colonies were selected via colony PCR with the HotStart Taq DNA Polymerase with Antibody (Genaxxon) using T3/T7 primers, which flanked the multiple cloning sites. The colony PCR was done by picking individual colony with pipette tips and suspending it in reaction tubes. The tips were then streaked onto another replicate LB agar plate. After incubating at 99ºC for 10min, suspended *E. coli* cells were used template. The colonies containing the correct insert were amplified in an overnight culture and the plasmids were isolated with QIAprep Spin Miniprep Kit (Qiagen). An analytical digest with suitable restriction enzymes was usually performed to further confirm the correct insertion. The template for *in vitro* transcription was generated in two ways, either by PCR amplification using the forward primer and M13F/M13R primer from the plasmid, or by linearizing the plasmid with either SacI or KpnI, depending on the orientation of the target and then blunt ending with T4 DNA Polymerase (Thermo Fisher Scientific). Both were then purified with QIAquick PCR Purification Kit (Qiagen).

The transcriptional labeling of RNA probes was achieved using the corresponding RNA promoter and T3/T7 RNA Polymerase (Roche) with DIG/Fluorescein RNA Labeling Mix (Roche), referring to manufacturer's protocol. Subsequently, ethanol precipitation was done with 0.1 volume of 4M LiCl and 3 volumes of ice-cold 100% ethanol for over 1h at -80ºC or overnight at -20ºC. It was then centrifuged, washed with 70% ethanol, air dried and dissolved in 30µl RNase-free H_2O (Qiagen). The probe yield and quality was determined by Qubit 3 Fluorometer (Thermo Fisher Scientific) and 1x TAE agarose gel electrophoresis, and the probe was snap frozen in liquid nitrogen and stored at -80ºC.

Fluorescence *in situ* hybridization

To maintain an RNase-free working environment, all glass jars, graduated cylinders and metal tweezers were baked at 200ºC for 3h prior to hybridization and all solutions were either directly treated with 0.1% DEPC (Carl Roth) and autoclaved or dissolved in DEPC-treated H_2O. To prevent nucleic acids from binding to the silicate surface of the coverslips, RNase-free coverslips were immersed in Repel-silane ES shortly (GE Healthcare) followed by rinsing in acetone and ethanol, and were air dried before further use.

Fluorescence *in situ* hybridizations were performed on coronal OE sections of the P7 mice (see detailed FISH protocol in Appendix). In day one, the sections were fixed in 4% PFA for 10min

with subsequent permeabilization in 0.2M HCl for 10min. To inactivate the endogenous peroxidase activity, the sections were treated with 3% H_2O_2 for 25min. Afterwards, an acetylation step with 5mM acetic anhydride in 0.1M TEA for 10min was aimed to block the binding site of the RNA probes on amine groups in the tissue. Acetic anhydride should be freshly mixed to 0.1M TEA. Between each functional step there was the wash step with PBS for 5min. The hybridization solution for each slide was 150µl and it contained 50% Formamide, 5x SSC (stock: 20x), 0.4µg/µl Torula RNA (stock: 20mg/ml), 0.1µg/µl tRNA (stock: 10mg/ml) with 1-3ng/µl of digoxigenin- or fluorescein-labeled RNA probe. It was denatured at 99ºC for 3min, chilled on ice before the addition of 5x Denhardt's (stock: 50x). Subsequently, the hybridization solution was pipetted onto each slide, which was covered with silanized coverslips to prevent evaporation. The sections were incubated in a humid chamber at hybridization temperature overnight.

The hybridization tempature was calculated with the following formula:

$Tm(ºC) = 79.8 + 18.5*(logNa^+) + 58.5*(\%GC) + 11.8*(\%GC)^2 - 820/L - 0.35*(\%formamide) - 25$
Na^+: concentration of Na^+ ions in moles per liter in the hybridization buffer
%GC: GC content of the probe (number between 0 and 1)
L: length of the probe
%formamide: percentage of formamide in the hybridization buffer

In day two, the slides were first washed in 5x SSC to remove the coverslips. In order to eliminate non-specific hybridization, several steps of stringency washes were performed, namely the slides were washed at hybridization temperature in the water bath with 50% Formamide, 2x SSC for 30min followed by 0.2x SSC for 1h, and then at room temperature with 0.2x SSC for 15min. Afterwards, a wash step with PBST for 5min was carried out. Before antibody detection, the sections were blocked in 2% Blocking reagent for 1h. The fluorescein-labeled RNA probe was first detected with Anti-Fluorescein-POD antibody (Roche) (1:1500 dilution in 2% Blocking reagent) by pipetting 150µl onto each slide. The sections were covered with Parafilm (Bemis) and the antibody incubation was overnight in a humid chamber at 4ºC.

In day three, all steps were performed at room temperature. The slides were first washed in PBST for 3x 10min to remove the unbound antibody. Before the signal detection by methyldopamine deposition, sections were treated with 1mM BHR in PBST (stock: 20mM) by pipetting 150µl onto each slide and incubated for 10min. BHR can conjugate tyrosine-like residues to the amines in tissue, which provided extra anchoring positions for the methyldopamine in the following color reaction. As the BHR was highly sensitive to humidity, it was freshly mixed to PBST before application. The slides were washed for 3x 5min in PBST, and the third wash step was supplied with 15µM DETAPAC (stock: 10mM), a metal ion chelating agent to remove metal ions. The subsequent color reaction was incubated for 1h in the dark in the humid chamber with 300mM mutarotated glucose (stock: 1.5M), 15µM DETAPAC (stock: 10mM), 5µM Alexa488-

methyldopamine to a final volume of 149μl with Borate buffer, and 1μl Glucose oxidase (1:1000 dilution from stock, Sigma Aldrich). By oxidizing the mutarotated glucose, the reaction needed hydrogen peroxide was generated and the peroxidase could thus anchor the Alexa488-methyldopamine to tyrosine residues in the tissue. The diluted Glucose oxidase should be freshly mixed to the color reaction mix, so that no excessive amount of hydrogen peroxide was produced. The slides were washed for 3x 5min in PBST after the color reaction. The first Anti-Fluorescein-POD (Roche) was then inactivated by incubating in 0.1M glycin (pH 2.0) for 10min and subsequently in 6% H_2O_2 for 30min. After washing the slides in PBS for 5min, they were blocked in 2% Blocking reagent for 1h. The Anti-Digoxigenin-POD antibody was diluted 1:1500 in 2% Blocking reagent and likewise 150μl was pipetted onto each slide, which was covered with Parafilm (Bermis) and incubated overnight in a humid chamber at 4ºC.

Similar to the detection procedure in day three, the Anti-Digoxigenin-POD antibody (Roche) was utilized to catalyze the color reaction in day four. Likewise the sections were sensitized with 1mM BHR and the color reaction was realized with Cy3-methyldopamine. After washing for 3x 5min in PBST, the sections were counterstained with DAPI (1:1000 dilution in PBS) for 25min in the humid chamber by pipetting 150μl onto each slide. After final wash with PBS for 5min, the sections were embedded in Mowiol and dried up overnight before imaging.

If single color RNA-FISH was performed, depending on the labeling of the RNA probe, Anti-Digoxigenin-POD or Anti-Fluorescein-POD can be freely combined with Alexa488-, Cy3-, or Cy3B-methyldopamine. Instead of a four-day protocol, only one color reaction was performed and no inactivation of the peroxidase was necessary.

After RNA-FISH, fluorescent images were obtained using Zeiss ApoTome (Carl Zeiss) and Zeiss LSM 800 with Airyscan (Carl Zeiss). Image processing was performed with the ZEN software (Blue edition, Carl Zeiss). For quantifications, the "relative height" of OSNs in the epithelium or the fluorescent intensity of individual cell was measured in Fiji (Fiji is just ImageJ). Cell area segmentation was achieved by applying the algorithm "RenyiEntropy" as the thresholding method (select *Image>Adjust>Threshold*). Subsequently a mask was created (select *Edit>Selection>Create Mask*) and individual cell was first outlined with the wand tool in the mask and then transferred to the original channel (select *Edit>Selection>Restore selection*). Different parameters can be measured (select *Analyze>Set Measurements*) with the Measure command (select *Analyze>Measure*). Either the median value or the integrated density was plotted. Data analysis and graphing was done with Origin 2016 (OriginLab).

2.2.6 Mathematical derivation of the distal/proximal ratio in immature neurons

To calculate the distal/proximal ratio in the immature mOR37A-expressing neurons, the following mathematical formula was used. In brief, the distal/proximal ratio measured from

ddPCR in the entire epithelium is, in other words, the ratio of the absolute copy of the distal transcripts and the proximal transcripts. OSNs can be subdivided into mature and immature populations based on their differentiation state. Assuming that the isoform expression profile of each OSN in the immature or mature population is the same, the proximal or the distal transcript copies can then be reflected as the cell number times the transcripts that each OSN makes. As mOR37A demonstrates a maturation-dependent on/off switch of the long isoform, it is assumed that the mature mOR37A-expressing neurons do not contain the long isoform. Eventually, to derive the distal/proximal ratio in the immature population, three variables are needed, namely, the ratio between the distal and proximal transcripts in the entire epithelium (measured by ddPCR), the ratio between the mature and immature neurons (quantified by RNA-FISH cell counting, Falk, 2015), and the ratio of the proximal transcripts in the mature and immature neurons (quantified by RNA-FISH intensity measurements).

$$\left(\frac{d}{p}\right)^{(all)} = \frac{\mathcal{N}_{(imm.)} \times n_{dist.}^{(imm.)}}{\mathcal{N}_{(m.)} \times n_{prox.}^{(m.)} + \mathcal{N}_{(imm.)} \times n_{prox.}^{(imm.)}}$$

$$= \frac{\mathcal{N} \times r_{(imm.)} \times n_{dist.}^{(imm.)}}{\mathcal{N} \times \left(r_{(m.)} \times n_{prox.}^{(m.)} + r_{(imm.)} \times n_{prox.}^{(imm.)} \right)}$$

$$= \frac{r_{(imm.)} \times n_{dist.}^{(imm.)} / n_{prox.}^{(imm.)}}{r_{(m.)} \times n_{prox.}^{(m.)} / n_{prox.}^{(imm.)} + r_{(imm.)}}$$

$$\frac{n_{dist.}^{(imm.)}}{n_{prox.}^{(imm.)}} = \left(\frac{d}{p}\right)^{(all)} \times \frac{1}{r_{(imm.)}} \times \left(r_{(m.)} \times \frac{n_{prox.}^{(m.)}}{n_{prox.}^{(imm.)}} + r_{(imm.)} \right)$$

$$\left(\frac{d}{p}\right)^{(imm.)} = \left(\frac{d}{p}\right)^{(all)} \times \left(\frac{r_{(m.)}}{r_{(imm.)}} \times \frac{n_{prox.}^{(m.)}}{n_{prox.}^{(imm.)}} + 1 \right)$$

d/p^{all}: distal/proximal ratio in the OE measured by ddPCR
$d/p^{imm.}$: distal/proximal ratio of the immature population in the OE
$N_{imm.}$, N_m: number of mature/immature cells
N: total number of mOR37A-expressing cells
$n_{prox}^{(m)}$, $n_{prox}^{(imm.)}$: number of proximal transcripts in the mature/immature cells
$r_{(m.)}$, $r_{(imm.)}$: ratio of the mature/immature cells

3. Results

We are particularly interested in the olfactory axon guidance mechanism of how 1000 different types of neurons in the olfactory epithelium form highly stereotyped connections in the olfactory bulb. In our previous search of receptor-specific guidance-related molecules for the exemplary receptor mOR37A, we found an overrepresentation of a non-coding transcript by differentially screening single-cell cDNA libraries of outgrowing and mature neurons. This transcript was localized next to the mOR37A gene and was expressed exclusively in mOR37A immature neurons (Haag, 2009). Further experimental data demonstrated that it might be part of the mOR37A mRNA, potentially generated by alternative polyadenylation, in parallel to a much shorter isoform (Helisch, 2014; Falk, 2015).

Based on these findings, this thesis mainly explored the following questions using the transgenic mOR37A-IRES-tauEGFP strain. First, the transcript structure, especially the 3' ends of mOR37A was clarified with 3' RACE in combination with Southern blots. Two major transcript isoforms, generated by alternative polyadenylation, were identified with their corresponding 3'UTRs being 1438nt and 4727nt in length. The spatial distribution of different isoforms was characterized by RNA fluorescence *in situ* hybridization on coronal epithelial sections, where the long isoform exhibited a predominant expression by the immature olfactory sensory neurons. In addition to this, the subcellular details were highlighted with Airyscan imaging, revealing potential "transcriptional hotspots" inside the nucleus. Furthermore, the relative expression level of different isoforms was quantified by ddPCR at different developmental stages. In accordance with previous observations, the long isoform was downregulated during development. Secondly, as the interplay between splicing and alternative polyadenylation could occur, the link between these two events was examined in the mOR37A transcripts. No correlation between the 5' intron splicing and the usage of the polyA sites was, however, discovered. Thirdly, to address the potential functional relevance of different isoforms in the axon guidance process, the ratio between both isoforms in the axon termini was placed under scrutiny and compared to that in the cell soma. No differential localization was found in different subcellular compartments, but a potential RNA-RNA interaction partner of the short isoform emerged in the axon termini. Interestingly, the observed RNA-RNA interaction was only present in the juvenile but not the adult stage. The identity of the interaction partner remains, however, unknown at this stage. Finally, alternative polyadenylation was analyzed in other receptors by 3' RACE, RNA fluorescence *in situ* hybridization and qPCR, corroborating the basic findings from mOR37A and rendering it a more general phenomenon.

3.1 Alternative transcripts of receptor mOR37A in the mOR37A-IRES-tauEGFP strain

3.1.1 The structure of the mOR37A-IRES-tauEGFP locus

In this work, if not stated otherwise, the transgenic mouse strain mOR37A-IRES-tauEGFP was used, in which an *IRES-tauEGFP* sequence was inserted 3nts downstream of the stop codon of the mOR37A coding region (Strotmann et al., 2000).

To exclude any inclusion of unwanted nucleotides by the insertion, which might lead to a premature termination of the transcript, the EGFP/ mOR37A 3'UTR fusion site was amplified with two primers flanking it [GFP_539F; 37A_3'UTR_416R], and was sequenced. The sequencing result showed an insertion of 67 nucleotides between the EGFP stop codon and the mOR37A 3'UTR (Fig.3.1). This insertion originates from a Cre-loxP recombination to remove the neomycin-selectable marker in generating the targeted mutagenesis of the *mOR37A* gene (Strotmann et al., 2000). Although two canonical polyadenylation signals *"ATTAAA"* were contained in the insertion site, they were not recognized as potential signals for transcription termination, possibly due to the lack of upstream U-rich and downstream GU-rich elements. Thus, no deviations from the WT C57BL/6 strain are expected with regard to the polyadenylation process.
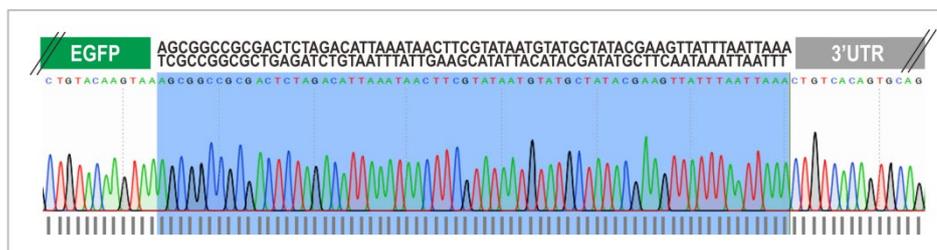

Figure 3.1. Insertion site between EGFP and mOR37A-3'UTR in the mOR37A-IRES-tauEGFP mouse. Sequencing data indicate an insertion of 67 extra nucleotides between the EGFP sequence and the mOR37A 3'UTR, which do not contain any recognizable polyadenylation signal for a shortened transcript. The amplicon was amplified with P7 mouse gDNA with primers GFP_539F and 37A_3'UTR_416R.

3.1.2 Alternative polyadenylation of mOR37A transcripts

To get a comprehensive image of the 3' ends of mOR37A transcripts, 3' RACE in combination with Southern blotting was performed with both P7 and adult OE RNA. The 3' RACE gene specific primer [37A_3'Race_GSP(GFP)] was chosen in the EGFP sequence 5' upstream of mOR37A 3'UTR, and should provide high specificity due to its non-mouse origin (Fig.3.2B). Based on previous findings of the 3'-extended transcript, three Southern blot probes were designed to target the proximal part (probe 1/2) and the distal part (probe 3) of potential mOR37A transcripts respectively (Fig.3.2B). RACE-PCR products after different cycle numbers (20x, 25x, 30x) were first loaded on an agarose gel, where multiple bands could be seen, before

subsequent Southern blotting (Fig.3.2A). In total, five products were identified by Southern blotting, with their corresponding sizes (from 3' RACE GSP to the potential 3' end) calculated by calibration to the original gel (Fig.3.2C). Notably, both proximal probes 1 and 2 generated comparable patterns between P7 and adult OE, while the distal probe 3 could only hybridize to RACE-PCR products from P7 OE (Fig.3.2C). Another observation was that lower cycle numbers promoted the amplification of larger products and *vice versa* (Fig.3.2C). As the next step, nested PCRs were carried out on either the entire RACE-PCR reaction mix or the excised bands identified by Southern blots (Fig.3.2C). All products resulting from the nested PCR were sent out for sequencing.

Sequencing data revealed three potential transcript isoforms, which corresponded to the 1.1kb, 1.7kb and 5kb band in the Southern blot (Fig.3.2C.D). The remaining 2kb and 3kb fragments mapped either to genomic A rich regions in the mOR37A 3'UTR, indicating mispriming of the UPM, or somehow other genomic locations (data not shown), and were therefore not considered as real ends. Among the three identified ends, end 1 (green asterisk) possessed a rare polyadenylation signal (*"AATAAT"*) and an unusually short distance between this signal and its polyA tail (5nt apart) (Fig.3.2D). Furthermore, the mapped location was much more 5' upstream than what the Southern blot hinted at. Taken together, this indicates that end 1 might not be the true termination position of the mOR37A transcripts. The other two ends (blue and red asterisks), on the contrary, had the canonical polyA signals of *"ATTAAA"* and *"AATAAA"*, as well as a recognizable spacing of 16nt and 21nt to their polyA tail, respectively (Fig.3.2D). Moreover, they matched to the two 3' ends identified in the WT C57BL/6 strain (Helisch, 2014). This observation confirmed that the genomic composition of IRES-tauEGFP knock-in did not alter the alternative polyadenylation process of mOR37A mRNA. Noteworthy, end 3 (red asterisk) could only be detected in the P7 OE, while end 2 (blue asterisk) in both the P7 and adult OE (Fig.3.2C), implying potential developmental regulation of the long isoform.

Up to this point, two major transcript isoforms of mOR37A have been identified in mOR37A-IRES-tauEGFP transgenic mouse, with the short 3'UTR being 1438nt and the long 3'UTR being 4727nt. They are dubbed "short isoform" and "long isoform" for future references.
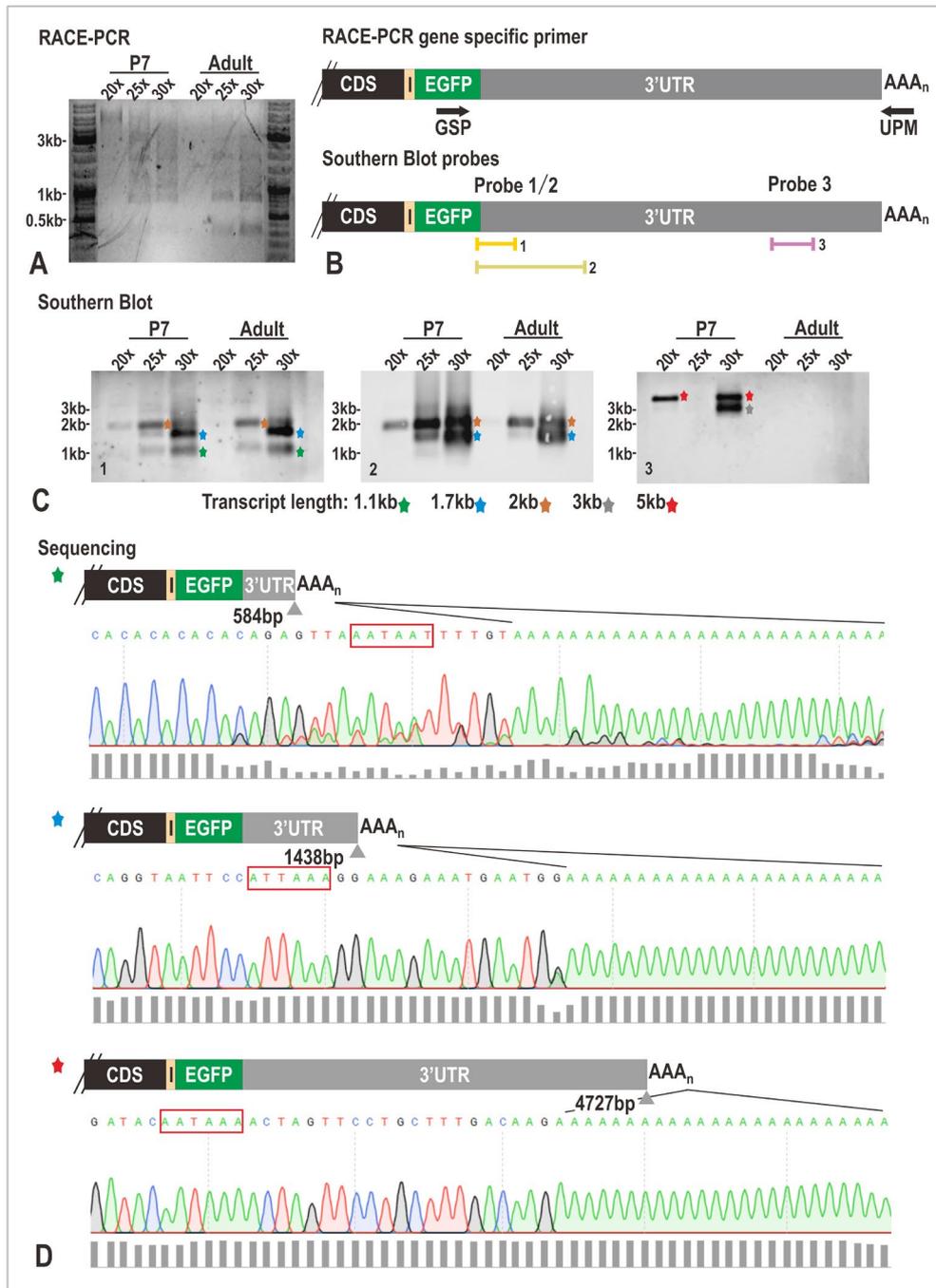
Figure 3.2. 3'UTR alternative polyadenylation of mOR37A transcripts in the mOR37A-IRES-tauEGFP strain analyzed by 3' RACE combined with Southern blotting.

A. 3' RACE-PCR with P7 and adult OE RACE-cDNA. RACE-PCR was performed with different cycle numbers (20x, 25x, 30x), and 10μl of the reaction was loaded on an agarose gel. Multiple bands can be seen.

B. Schematic illustration of the RACE-PCR gene specific primer and the Southern blot probes. As the 3' RACE gene specific primer, 37A_3'Race_GSP(GFP) in the EGFP sequence was used in combination with 10x UPM. For Southern blotting, three probes were designed for hybridization.

Probe 1: derived from PCR template with Si_mOR37A_prox_fwd; Ki_mOR37A_prox_rev. 341bp
Probe 2: derived from PCR template with Si_mOR37A_prox_fwd; Ki_qPCR_rev. 1327bp
Probe 3: derived from PCR template with Si_mOR37A_dist_fwd; Ki_mOR37A_dist_rev. 617bp

C. Southern blot on 3' RACE-PCR. Five potential transcript isoforms were identified by Southern blotting (marked by asterisks), with their corresponding sizes (from GSP to potential 3' end) calculated by calibration to the original agarose gel. The loss of the 5kb band from P7 OE 25x cycles with probe 3 was likely due to a gel loading error.

D. 3' ends of mOR37A transcripts by sequencing. Three potential 3' ends were revealed by sequencing. Among them, end 2 and 3 match to the ends in the WT C57BL/6 mouse, while end 1 might be a novel one.

However, the lack of canonical polyadenylation signal and the unusually short distance between its polyA signal and polyA tail render it less likely to be a true end. Numbers correspond to the 3'UTR length. Red rectangles mark the polyadenylation signals.
Nested primer for end 1: 37A_3'UTR_329F; NUP.
Nested primer for end 2: 37A_s4-s3_fwd; NUP.
Nested primer for end 3: 37A_3'UTR_4374F; NUP.

3.1.3 Expression pattern of mOR37A transcript isoforms in the olfactory epithelium

To investigate the expression pattern of different isoforms, RNA-FISH was performed on coronal OE sections of the P7 mice. Two antisense RNA probes were designed to target the proximal and distal regions of mOR37A transcripts, respectively (Fig.3.3A). The proximal probe could label the entire transcript repertoire, not distinguishing short from long isoforms, whereas the distal probe should exclusively label the long isoform.

Both probes labeled single OSNs across the OE (Fig.3.3B.C). The subcellular details were highlighted with Airyscan imaging. First to notice was the unique nuclear heterochromatin architecture of the OSNs (Fig.3.3D). In accordance with previous publication (Armelin-Correa et al., 2014), OSNs were characterized by one or a few large centrally localized heterochromatin blocks together with several smaller surrounding heterochromatin clusters (Fig.3.3D). This pattern was not seen in the apical supporting sustentacular cells, in which only smaller but densely packed heterochromatin blocks were present (Fig.3.3D). Concerning the subcellular distribution, both the proximal and distal probe labeled transcripts showed cytosolic and nuclear localizations (Fig.3.3B'.C', white lines delineate the nuclei). Noteworthy, the bright spot inside the nucleus was always exactly in the heterochromatin-free regions directly neighboring a heterochromatin block (Fig.3.3B'.C'), representing potential "transcriptional hotspots". This is as expected from the current model of OR selection that an OR gene escapes the heterochromatin blocks and gets addressed by multiple intergenic olfactory receptor enhancers (Monahan et al., 2019).

To quantify the epithelial localization of the cells labeled by different probes, the "relative height" of each cell in the OE was measured by dividing the "total OE thickness" by the "distance to basal side" (Fig.3.3E). Since OSNs migrate from the basal side to the apical side upon maturation, the "relative height" could be an indicator of the maturation level of a cell. Quantification suggested that cells labeled by the proximal probe were found across the whole OE with a peak accumulation in the mature OSNs (0.6-0.8), whereas cells labeled by the distal probe were found clearly at a more basal position (0.4-0.6) (Fig.3.3E). This quantified our previous qualitative observation that the long isoform expression was restricted to immature mOR37A OSNs (Falk, 2015).

Taken together, RNA-FISH demonstrates a potential maturation-dependent regulation with the long isoform being transcribed to a lesser extent in the mature mOR37A-expressing population.

Meanwhile, the cytosolic localization of the long isoform promotes the assumption of its involvement in the cellular processes instead of only regulatory functions in the nucleus.
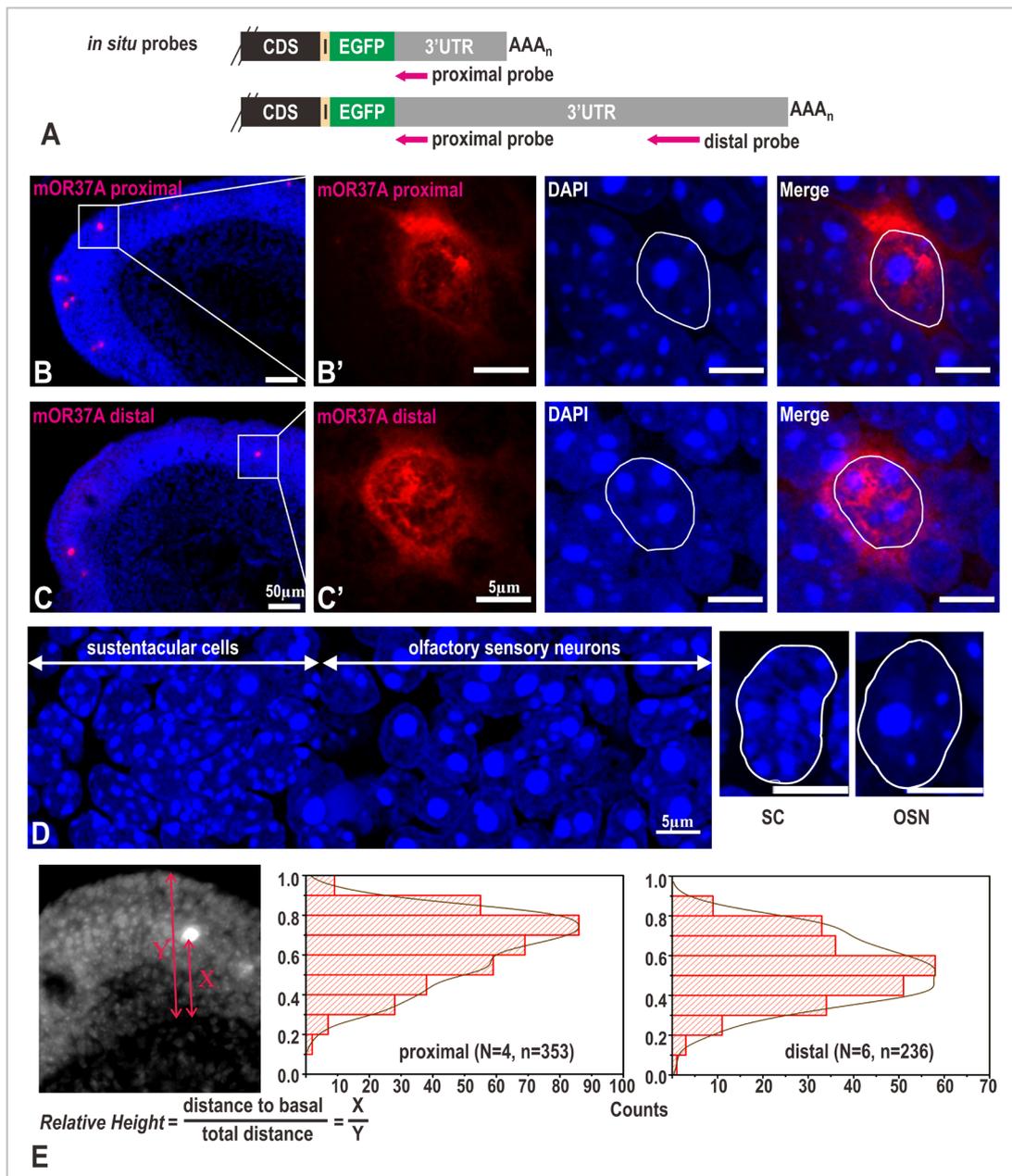

Figure 3.3. Distribution pattern of mOR37A transcript isoforms in the P7 OE analyzed by RNA-FISH.
Single color RNA-FISH was performed on the P7 OE sections with DIG-labeled antisense RNA probes and DIG-POD catalyzed Cy3B-methyldopamine color reaction.
A. Schematic illustration of the *in situ* antisense RNA probes. The proximal probe detects both isoforms, while the distal probe detects only the long isoform.
Proximal probe: derived from PCR template with Si_mOR37A_prox_fwd; Ki_mOR37A_prox_rev. 341bp
Distal probe: derived from PCR template with Si_mOR37A_dist_fwd; Ki_mOR37A_dist_rev. 617bp
B.C. Representative images of the proximal and distal probe labeled cells. Single OSNs across the OE are labeled by both the proximal and distal probes, with DAPI counterstaining. Image acquisition with Zeiss ApoTome, 20x/0.5 EC Plan-Neofluar objective. Scale bars, 50µm.
B'.C'. Subcellular distribution of the proximal and distal probes labeled transcripts. Transcripts labeled by both the proximal and distal probes show cytosolic and nuclear localizations. The nuclear localization is found in the heterochromatin-free regions with DAPI counterstaining. White lines delineate the nuclei. Image acquisition with Zeiss LSM 800 with Airyscan, 40x/water immersion objective. Z-stack with orthogonal projection. Scale bars, 5µm.

D. Nuclear organization of the OSNs. DAPI counterstaining was done on P7 OE sections. OSNs have typically one or a few centrally localized large heterochromatin blocks with smaller clusters surrounding it, whereas sustentacular cells have smaller but much more densely packed heterochromatin clusters. White lines delineate the nuclei. Image acquisition with Zeiss LSM 800 with Airyscan, 40x/water immersion objective. Z-stack with orthogonal projection. Scale bars, 5μm.

E. Quantification of the "relative height" of the proximal and distal labeled cells in the OE. The "relative height" is calculated by dividing the "total OE thickness" by the "distance to basal side". A value towards 1 means apical localization (mature OSNs) and a value towards 0 means basal localization (immature OSNs). The entire mOR37A-expressing population is distributed across the entire OE with a peak in the mature OSNs (0.6-0.8). Cells labeled by the distal probe localize, however, within the immature OSNs (0.4-0.6). N: independent experiments; n: numbers of analyzed cells.

3.1.4 Expression level of mOR37A transcript isoforms in the olfactory epithelium

To assess the relative expression level of the long isoform, ddPCR was performed with OE RNA from P1, P7 and adult mice. Similar to the situation in the RNA-FISH experiments, two primer pairs were chosen targeting the proximal and distal region of mOR37A transcripts (Fig.3.4A). Correspondingly, two TaqMan probes were designed in each amplicon to enable duplex assays with enhanced sensitivity and specificity (Fig.3.4A). The final quantification is presented as the ratio between the distal targets and the proximal targets (total transcripts), *i.e.*, distal/proximal. ddPCR input was 20ng-100ng of random primed, non-purified cDNA. The concentration here referred to the amount of total RNA input in cDNA synthesis, assuming the conversion efficiency being 100%.
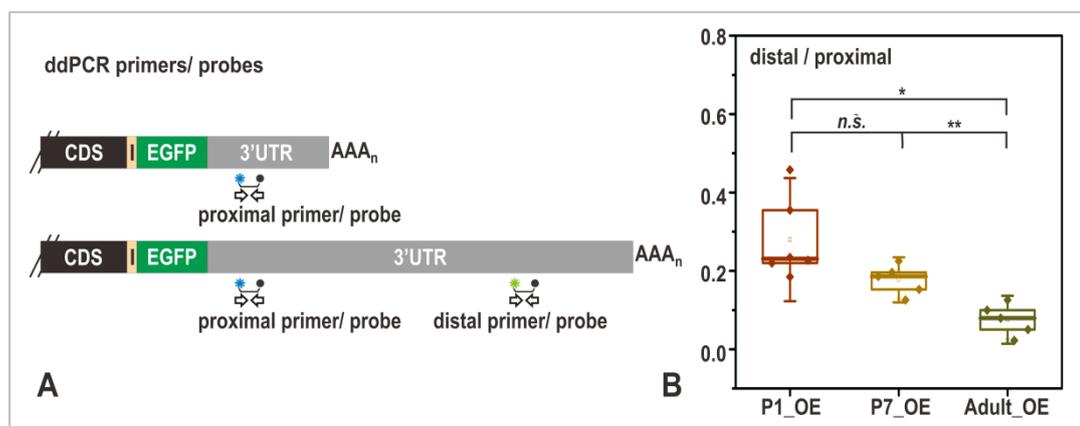

Figure 3.4. Expression level of mOR37A transcript isoforms in the OE measured by ddPCR.
ddPCR was performed with OE cDNA from different developmental stages. Random primed, non-purified cDNA was used as template, with 20ng-100ng input per reaction.
A. Schematic illustration of the ddPCR primers and the TaqMan probes. Two primer pairs together with two TaqMan probes were designed. The proximal probe detects both isoforms, while the distal probe detects only the long isoform. The proximal probe is labeled with FAM/BHQ1and the distal probe with HEX/BHQ1.
Proximal primers: 37A_3'UTR_175F; 37A_3'UTR_416R. 262bp
Distal primers: 37A_3'UTR_3429F; 37A_3'UTR_3676R. 267bp
B. Quantification of the distal/proximal ratio during development. The relative expression level of the long isoform is downregulated during development. Each dot corresponds to one animal and the whiskers represent standard deviation; T-test with n.s.: $\alpha \geq 0.05$, *: $\alpha < 0.05$, **: $\alpha < 0.01$.
P1_OE: N=6, distal/proximal ratio 0.231.
P7_OE: N=5, distal/proximal ratio 0.186.
Adult_OE: N=5, distal/proximal ratio 0.0793.

The distal/proximal ratio was significantly downregulated between the juvenile (P1: 0.231/P7: 0.186) and adult stages (0.0793) (Fig.3.4B). This underrepresentation of the long isoform during development might be due to a loss of the immature OSNs in the adult stage.

3.1.5 Expression level of mOR37A transcript isoforms in the immature neurons

The relative expression level of the long isoform has been examined at different developmental stages in the entire OE (Fig.3.4B), which consists of a mixed population of both mature and immature OSNs. To further pin down the expression level of different transcript isoforms in the immature population, single-cell based quantitative RNA-FISH was applied to quantify the fluorescent intensity of each labeled cell. The theoretical underpinning relied on the linear relationship between the transcript copy numbers and their detectable intensity, or in other words, more transcripts, higher intensity.

As described in section 3.1.3, two antisense RNA probes were designed to target the proximal and distal region of mOR37A transcripts, respectively (Fig.3.3A). The median intensity of the proximal probe labeling was measured with ImageJ (Fiji), applying the algorithm "RenyiEntropy" as thresholding method for cell area selection (Fig.3.5A, see Materials and Methods for details). As the "relative height" of individual cell in the epithelium contains the information about their maturation state, the immature population residing in the basal compartment of the epithelium was assumed to have a "relative height" of lower than 0.5 and the mature population residing in the apical compartment of higher than 0.6. The cells localized in between were not taken into consideration due to the potential overlap of the two populations. Calculated from three independent experiments (Fig.3.5B), the average intensity ratio of the proximal probe labeling between the mature and the immature population was 1.27 (Tab.3.1). According to the following formula (see Materials and Methods for details) and our existing data, the distal/proximal ratio in the immature neurons can be extrapolated with a value of approximately 0.515. This corresponds to a ratio of the short and long isoform of 1:1 in the immature population, which differs substantially from the distal/proximal ratio of the entire epithelium (Fig.3.4B).

$$\left(\frac{d}{p}\right)^{(imm.)} = \left(\frac{d}{p}\right)^{(all)} \times \left(\frac{r_{(m.)}}{r_{(imm.)}} \times \frac{n_{prox.}^{(m.)}}{n_{prox.}^{(imm.)}} + 1\right)$$

d/p^{all}*: distal/proximal ratio in the entire epithelium*
$r_{(m.)}$*,* $r_{(imm.)}$*: ratio of the mature/immature cells*
$n_{prox}^{(m)}$*,* $n_{prox}^{(imm.)}$*: number of proximal transcripts in the mature/immature cells*

d/p^{all} (P7 OE) = 0.186 (Fig.3.4B)

$r_{(m.)}/r_{(imm.)}$ = 0.582/0.418 (Falk, 2015)

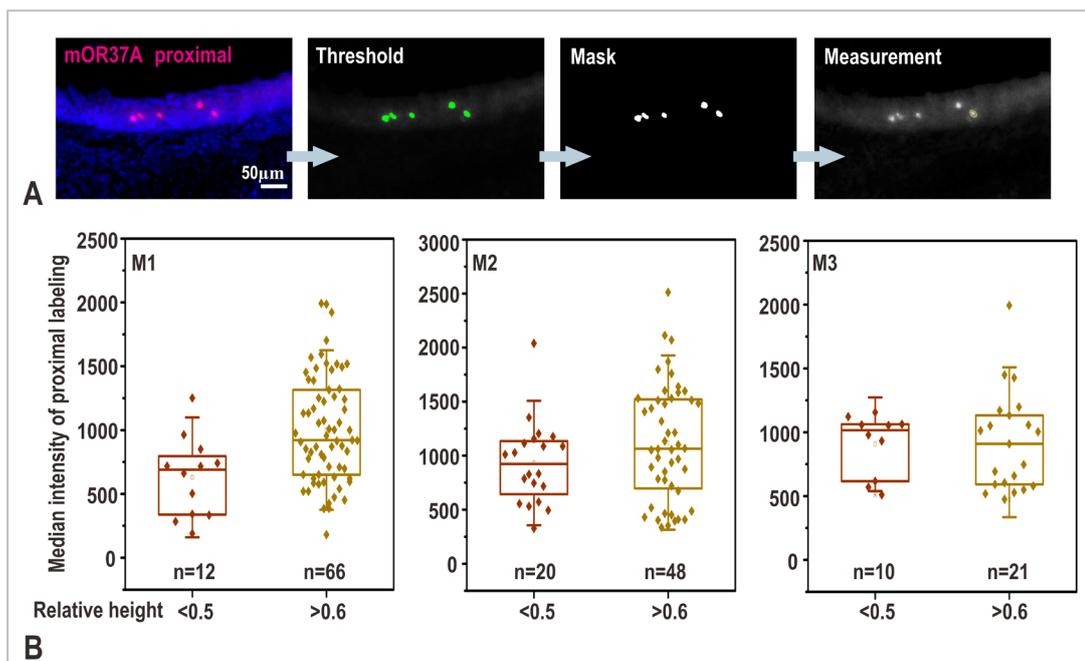$n_{prox}^{(m.)}/n_{prox}^{(imm.)}$ = 1.27 (Tab.3.1)

Figure 3.5. Fluorescent intensity measurements of the proximal probe labeling in the mature and immature OSNs analyzed by RNA-FISH.

Single or double color RNA-FISH was performed on the P7 OE sections with DIG-labeled proximal antisense RNA probe and DIG-POD catalyzed Cy3B-methyldopamine color reaction. Only the proximal probe labeling was quantified by the intensity measurement.

A. Schematic illustration of the quantification procedure. The quantification was done with ImageJ (Fiji). To select individual cell, the thresholding method "RenyiEntropy" was applied. After creating a mask, the individual cell was outlined and reselected back in the original channel. The median fluorescent intensity was measured.

B. Quantification of the median intensity of the proximal labeling in the immature and mature OSNs. The "relative height" of the immature population was set to be smaller than 0.5 and of the mature population larger than 0.6. The median intensity of both populations is comparable. Three independent experiments were quantified and the exposure time was normalized to 8ms. Each dot corresponds to one cell and the whiskers represent standard deviation.

Table 3.1. Mean value of the median fluorescent intensity measured from the proximal probe labeling in the mature and immature neurons.

	M1	M2	M3	Average
Immature cells (relative height <0.5)	629.4	932.5	906.4	/
Mature cells (relative height >0.6)	1000.7	1120.3	921.3	/
Intensity$_{(mat.)}$/Intensity$_{(imm.)}$	1.59	1.20	1.02	1.27

Up to this point, the transcript structure of mOR37A has been explored. Most likely, alternative polyadenylation shapes its transcripts by generating two isoforms with distinct 3'UTRs. Importantly, the distal/proximal ratio exhibits an age-dependent developmental regulation at tissue level measured by ddPCR, which correlates with the observation seen at single cell level analyzed by RNA-FISH that the expression of the long isoform in immature neurons disappears later during maturation.

3.2 Interplay between splicing and alternative polyadenylation

3.2.1 Alternative splicing of mOR37A transcripts in the WT C57BL/6 strain

It is widely appreciated that mammalian OR genes are devoid of introns within their coding region (Buck & Axel, 1991), but an intron is typically present between the non-coding leader exon and the coding exon (Glusman et al., 1996; Asai et al., 1996). In the mOR37A transcript, an intron of 2251bp is annotated directly upstream of mOR37A coding region in the Ensembl database (Release 97). Since the activity of the alternative polyadenylation and alternative splicing machineries could be intertwined and co-regulated, I explored the possible link between these two events in the OSNs. Given the well-known difficulty in amplifying the IRES site, all PCR-based experiments addressing alternative splicing were performed in the WT C57BL/6 background. Our previous work has already revealed two alternatively polyadenylated isoforms in the WT C57BL/6 mouse (Helisch, 2014), with the exact same ends seen in the transgenic mouse, validating this approach.

To first check the splicing of the 5' intron, RT-PCR was performed using two sets of primers, with the same forward primer in the intron, and a reverse primer either before or after the 1st polyA site (Fig.3.6A). RT-PCR was done with random primed OE cDNA from adult WT C57BL/6 mouse. Both primer sets generated one single band of the expected size, which was confirmed by the gDNA control (Fig.3.6B). This hinted at nuclear unprocessed pre-mRNA or at intron retention in cytosolic mature transcripts. It cannot be judged at this stage from which population the intron-containing transcripts originates, as TRIzol® (phenol/chloroform) based total RNA extraction isolates both cytosolic and nuclear RNA. Notably, the intron was clearly contained in some fraction of the long isoform. It remained, however, unclear whether the intron was also present in the short isoform, as primer set 1 detected both isoforms.

To get a comprehensive image of the exact 5' end of mOR37A transcripts containing the intron, 5' RACE was performed with OE RNA from the P7 WT C57BL/6 mouse. The 5' RACE gene specific primer [37A_cds_5'Race_GSP] was located at around 200bp downstream of mOR37A coding region. Different cycle numbers were applied (20x, 25x, 30x), and the RACE-PCR products were first loaded on an agarose gel. For nested PCRs, using the RACE-PCR reaction mix as template, a reverse primer in the intron sequence was combined with NUP, and the resulting PCR bands were sequenced. In brief, sequencing data mapped to the same genomic location as seen previously in the intron-depleted transcripts with one nucleotide difference (green, Fig3.6C) (Helisch, 2014). A short extension (red) was identified to the 5'UTR (magenta) annotated in the Ensembl database (Release 97) (Fig.3.6C). Thus, combining the previous 5' RACE data, there are two transcript populations concerning the 5' intron splicing, namely one with the intron contained and the other with the intron removed. It is, however, not clear whether these intron-containing transcripts are mature or unprocessed transcripts due to the

aforementioned reasons. It indicates that splicing does occur in mOR37A-expressing OSNs, and alternative splicing and alternative polyadenylation were therefore examined by RNA-FISH in more detail, in which the cytosolic and the nuclear transcripts can be visualized separately.
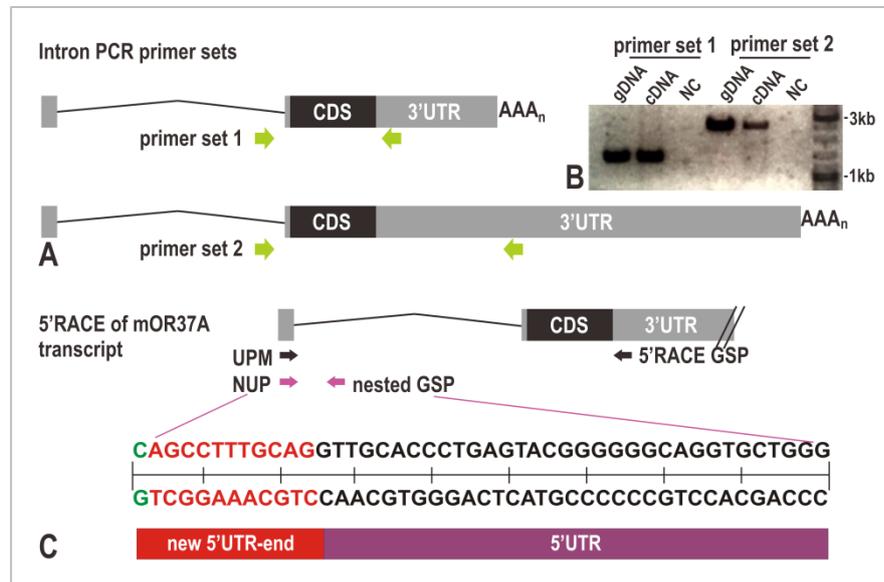

Figure 3.6. mOR37A transcript isoforms in the WT C57BL/6 mouse.
A. Schematic illustration of the localization of the 5' intron PCR primers. Both sets of primers contain the same forward primer located in the intron, and the reverse primers before or after the 1st polyA signal.
Primer set 1: 37A_intron_-47F; 37A_3'UTR_416R. 1462bp
Primer set 2: 37A_intron_-47F; 37A_dd prox 1stpA rev2. 2528bp
B. 5' intron alternative splicing analyzed by RT-PCR. PCR was done with random primed OE cDNA from the adult WT C57BL/6 mice. 5' intron is retained at least in some transcripts, especially in the long isoform.
C. 5' ends of mOR37A transcripts containing the 5' intron analyzed by 5' RACE. 5' RACE was performed with OE RNA from the P7 WT C57BL/6 mice. As the 5' RACE gene specific primer, 37A_cds_5'Race_GSP downstream of mOR37A coding region was used in combination with 10x UPM. For nested PCRs, the nested primer 37A_5'UTR_-2291R in the intron was chosen. Sequencing result of the intron-containing transcripts matches to previously identified 5' end of the intron-depleted transcripts (Helisch, 2014) with one nucleotide difference (green). A short extension (red) is added to the annotated 5'UTR (magenta) in the Ensembl database (Release 97).

3.2.2 Expression pattern of mOR37A intron in the olfactory epithelium

Though it remains unknown whether the 5' end modification seen in the WT C57BL/6 mouse also applies to the transgenic mouse, the consistent 3' alternative polyadenylation events in both strains promotes me to further probe the link between these two machineries in the transgenic mouse by RNA-FISH. For this purpose, one antisense RNA probe was designed in the intron sequence (Fig.3.7A) and the localization of the intron-containing cells in the epithelium as well as the subcellular distribution of the intron-containing transcripts was investigated by RNA-FISH on OE sections of the P7 transgenic mouse.
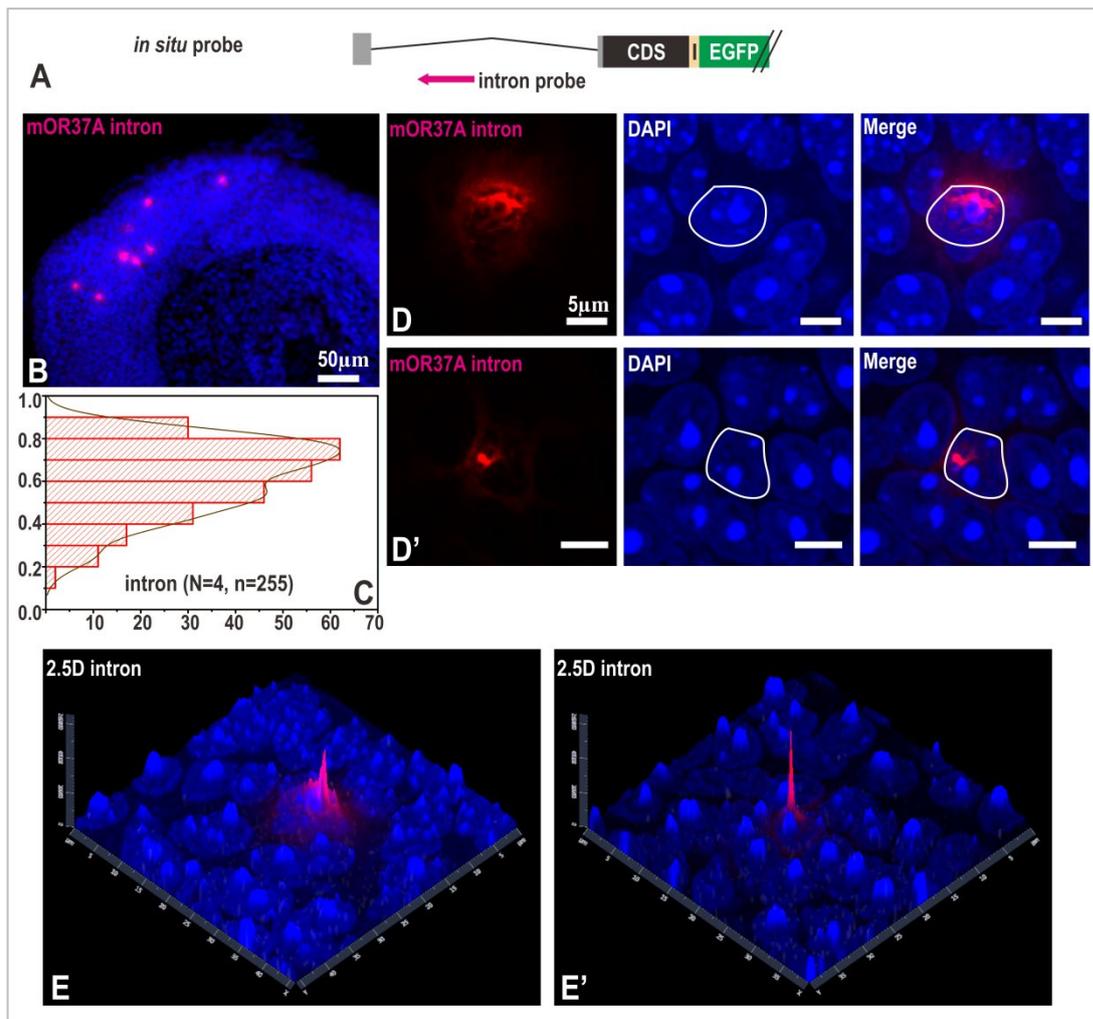
Figure 3.7. Distribution pattern of mOR37A intron-containing transcripts in the P7 OE analyzed by RNA-FISH.
Single color RNA-FISH was performed on the P7 OE sections with a DIG-labeled antisense RNA probe and DIG-POD catalyzed Cy3B-methyldopamine color reaction.
A. Schematic illustration of the *in situ* antisense RNA probe. The intron probe is in the intron sequence and detects intron-containing transcripts.
Intron probe: derived from PCR template with Si_mOR37A_5'UTR_-1987F; Ki_mOR37A_5'UTR_-1371R. 627bp
B. Representative image of the intron probe labeled cells. Single OSNs across the OE are labeled by the intron probe. Image acquisition with Zeiss ApoTome, 20x/0.5 EC Plan-Neofluar objective. Scale bar, 50μm.
C. Quantification of the "relative height" of intron-containing cells in the OE. The "relative height" is calculated by dividing the "total OE thickness" by the "distance to basal side" (see Fig.3.3E). A value towards 1 means apical localization (mature OSNs) and a value towards 0 means basal localization (immature OSNs). The localization of the intron-containing population resembles that of the entire mOR37A-expressing population with a peak in the mature OSNs (0.6-0.8). N: independent experiments; n: number of analyzed cells.
D.D'. Subcellular distribution of the intron probe labeled transcripts. Intron-containing transcripts mainly show nuclear localizations in heterochromatin-free regions. Image acquisition with Zeiss LSM 800 with Airyscan, 40x/water immersion objective. Z-stack with orthogonal projection. Scale bars, 5μm.
E.E'. 2.5D graphical illustration of the intron probe labeled transcripts. Intron probe labeled transcripts show one major fluorescent peak inside the nucleus. Image acquisition with Zeiss LSM 800 with Airyscan, 40x/ water immersion objective. Z-stack with orthogonal projection.

Notably, a substantial amount of cells were labeled by the intron probe (Fig.3.7B). When couting the number of cells labeled by the intron probe and comparing it to the cells labeled by the proximal probe in alternating section series, the intron-containing cells made up more than 70%

of the entire mOR37A-expressing population. The missing 30% was likely due to the sensivity difference between the intron and proximal probes. Furthermore, after quantifying the "relative height" of these intron-positive cells in the epithelium, a highly comparable localization pattern to the proximal probe labeled cells was observed with the peak in mature OSNs (0.6-0.8) (Fig.3.3E, Fig.3.7C). Despite the highly resemblance of the localization pattern, the subcellular organization of the intron-containing transcripts exhibited slightly different features as compared to the proximal or distal probe labeled transcripts, namely that they were highly concentrated in the nucleus (Fig.3.7D.D'). This was also evident in the 2.5D graphs of the intron-containing transcripts with one major fluorescent peak inside the nucleus (Fig.3.7E.E'). This observation might address the previous question of the origin of those intron-containing transcripts, and most of them would possibly be unprocessed nuclear pre-mRNAs. When inspecting the nuclear localization closely, a similar organization of the transcripts being localized between the heterochromatin clusters was observed, representing potential transcriptional active site or more specifically in this case the site for transcript processing by intron splicing.

In sum, the distribution pattern of the intron-containing cells in the OE contrasts with that of the long isoform-containing cells, rendering a strict co-regulatory mechanism of alternative polyadenylation of the long isoform and alternative splicing of the intron unlikely. Despite the similarity of the height distributions, the intron cannot be exclusively associated with the short isoform, since the primer set 2 generates a band in the intron RT-PCR (Fig.3.6B). Thus, not solely associated with either the short or the long isoform, alternative splicing and alternative polyadenylation might not share common regulatory mechanism and a certain percentage of both isoforms might carry the intron.

3.2.3 Expression level of intron-containing transcripts in the olfactory epithelium

To quantify the relative expression level of the intron-containing transcripts, ddPCR was performed with OE RNA from P1, P7 and adult mice. One pair of primers as well as one TaqMan probe was designed in the intron sequence and was combined with the proximal primer/probe, which was otherwise used for the quantification of the alternative isoforms (Fig.3.8A, Fig.3.4A). Similarly, the final quantification is presented as the ratio between the intron-containing transcripts and the total transcripts, *i.e.*, intron/proximal. ddPCR input was 20ng-100ng of random primed, non-purified OE cDNA.

Surprisingly, despite a high resemblance of the localization patterns in the epithelium, intron-containing transcripts made up only approximately 20% of the total transcripts (Fig.3.8B). Though a significant difference was detected between the P1 (0.302) and adult (0.148) stages, which might correspond to a developmental regulation or result from data fluctuation, there was no change between the P7 (0.185) and adult (0.148) stages (Fig.3.8B), which was notable

because the developmental regulation of the long isoform was clearly observed between these two stages (Fig.3.4B). Thus, corroborating the intron RNA-FISH conclusion, a direct association between alternative polyadenylation and alternative splicing could be ruled out. There are in total four mOR37A isoforms in the transgenic strain, namely the short and the long isoforms with or without the intron.
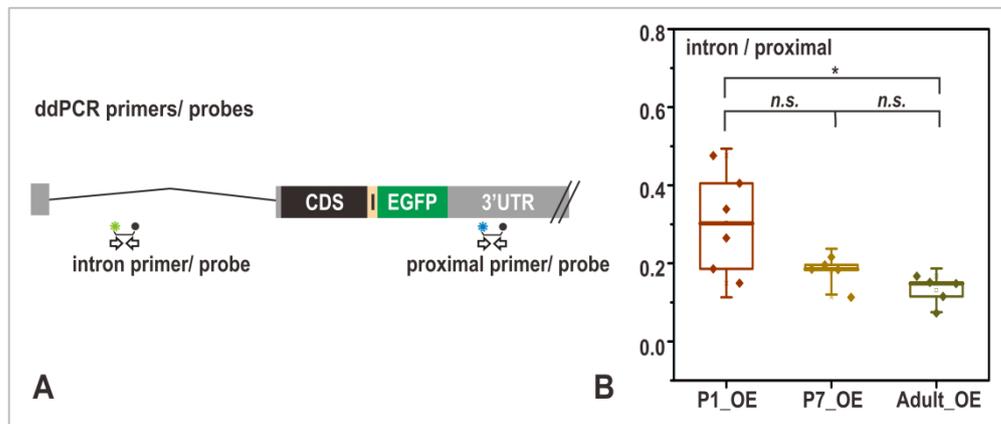

Figure 3.8. Expression level of mOR37A intron-containing transcripts in the OE measured by ddPCR.
ddPCR was performed with OE cDNA from different developmental stages. Random primed, non-purified cDNA was used as template, with 20ng-100ng input per reaction.
A. Schematic illustration of the ddPCR primers and the TaqMan probes. One pair of primers and one TaqMan probe were designed in the intron sequence. The intron probe is labeled with HEX/BHQ1, so that a combination with the existing proximal primer/probe is possible in ddPCR.
Intron primer: 37A_intron_-1995F; 37A_intron_-1822R. 174bp
B. Quantification of the intron/proximal ratio during development. The relative expression level of the intron-containing transcripts is not altered between P1 and P7 stages or P7 and adult stages. A significant reduction is, however, seen between P1 and adult stages, which might be due to flunctuations. Each dot corresponds to one animal and the whiskers represent standard deviation; T-test with n.s.: $\alpha \geq 0.05$, *: $\alpha < 0.05$.
P1_OE: N=6, intron/proximal ratio 0.302.
P7_OE: N=5, intron/proximal ratio 0.185.
Adult_OE: N=5, intron/proximal ratio 0.148.

In summary, the experiments conducted on the level of the olfactory sensory cell soma (*i.e.*, in the OE) so far corroborate the idea of a maturation-dependent regulation of mOR37A transcript isoforms, which promotes me to further probe the possibility of different functional roles of the different isoforms in the axon termini. In the next section, the research focus will therefore be switched from the OE to the OB.

3.3 A potential interaction partner of mOR37A transcripts in the axon termini

3.3.1 The effect of RNase treatment on the distal/proximal ratio in the P7 axon termini

Given neurons being highly polarized cells, the first working hypothesis regarding different isoforms in the cell soma is that there might be a selective axonal transport of either of them. In fact, receptor mRNA is known to be localized in the distal axonal compartment of OSNs based on ultrasensitive radioactive *in situ* hybridizations (Ressler et al., 1994; Vassar et al., 1994), however no previous publications have ever reported any isoform-dependent localization in the axon termini.

To address this possibility, the ratio between both isoforms was checked in axon termini in the OB and compared to that in the OE. Since axonal transport might bring very few copies of the target transcripts to the axon termini, the ddPCR protocol was modified with extra purification steps to increase sensitivity by removing possible contaminants or inhibitors. Owing to the anatomical separation of the OE (containing OSN cell bodies) and the OB (containing OSN axons), RNA isolation could be performed with ease. After cDNA synthesis with random primers, cDNA was purified with spin columns (Fig.3.9A). Another step in between was the RNase treatment, including both RNase H and RNase A (Fig.3.9A). The RNase treatment could remove both single-stranded RNA (RNase A) and RNA in the RNA-cDNA hybrid (RNase H), leaving pure cDNA for ddPCR (Fig.3.9A). In this section, typical input was 200ng-1000ng of purified cDNA for OB ddPCR and 20ng-100ng for OE ddPCR. The concentration here referred to the UV-spectrometry of ssDNA. The same primer/probe sets were used as in the OE (Fig.3.4A) and the ratio of distal/proximal was evaluated.

Surprisingly, the distal/proximal ratio in the OB showed a significant reduction upon RNase treatment (Fig.3.9B). Such difference, however, was only present in the OB, but not in the OE under the same experimental condition (Fig.3.9B). The distal/proximal ratio was 0.20 in the cell soma and axon termini with RNase treatment, and was 0.76 in the axon termini without RNase treatment. Thus, as demonstrated by the RNase treated samples, both isoforms occur in the axons in the same ratio as in the soma and no isoform is preferentially transported. However, a new and unexpected question emerged, that is why the OB cDNA exhibits sensitivity towards RNase treatment.
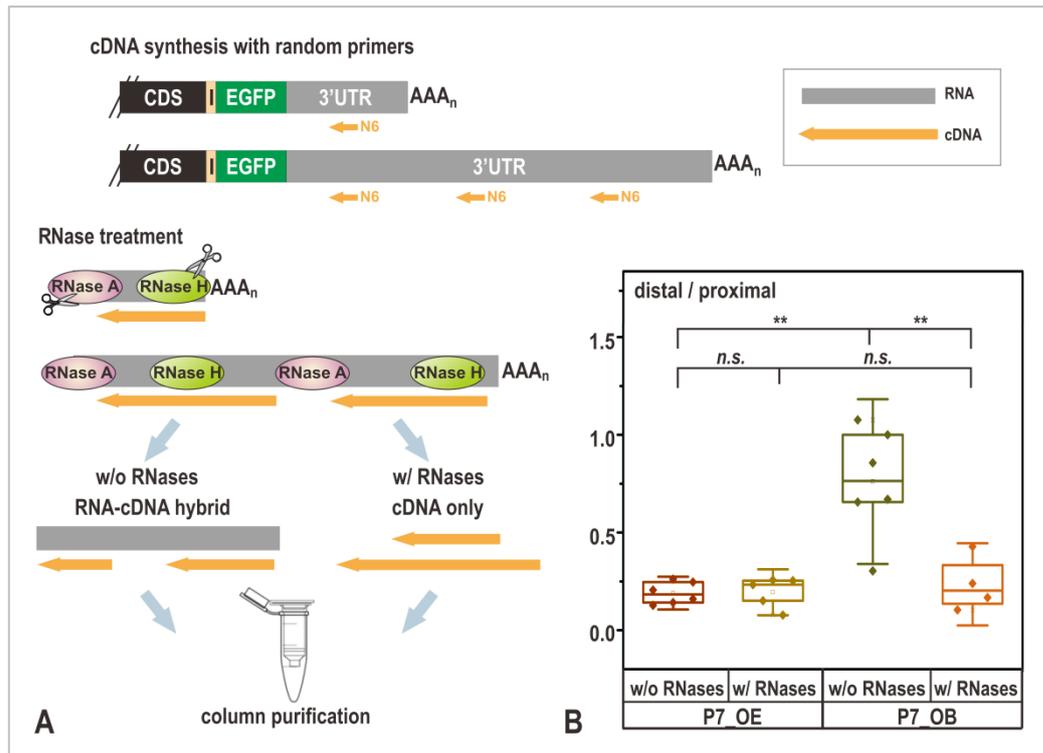
Figure 3.9. Effect of RNase treatment on the distal/proximal ratio in P7 OE and OB measured by ddPCR.
A. Schematic illustration of RNase treatment and column purification of OE/OB cDNA. In general, random primed cDNA was either directly purified with spin columns or was subject to RNase treatment prior to column purification. RNase treatment included RNase H and RNase A, which aimed to remove the RNA in RNA-cDNA hybrids (RNase H) and the single-stranded RNA (RNase A).
B. Effect of RNase treatment on the distal/proximal ratio. ddPCR was performed with P7 OE/OB cDNA. OE cDNA input was 40ng per reaction and OB cDNA input was 200ng-1000ng per reaction. The distal/proximal ratio is significantly reduced upon RNase treatment in the P7 OB, but not in the P7 OE. Each dot corresponds to one animal and the whiskers represent standard deviation; T-test with n.s.: $\alpha \geq 0.05$, *: $\alpha < 0.05$, **: $\alpha < 0.01$.
P7_OE, w/o RNases: N=6, distal/proximal ratio 0.183.
P7_OE, w/ RNases: N=5, distal/proximal ratio 0.233.
P7_OB, w/o RNases: N=6, distal/proximal ratio 0.764.
P7_OB, w/ RNases: N=4, distal/proximal ratio 0.203.

3.3.2 Effect of RNase treatment on control cDNA templates

To systematically test whether the aforementioned RNase effect on the axonal cDNA measurement originated from the true cellular context or from any artifacts, a series of control experiments were conducted by applying RNase treatment on cDNA samples from other resources.

To start with, cDNA from both P7 and adult OE was used for ddPCR under the same condition of RNase treatment and spin column purification. cDNA input was 20ng-40ng per reaction. In both cases, the distal/proximal ratio stayed unaltered, regardless of the treatment (Fig. 3.10A).

Furthermore, a synthetic system was developed after cohesive-end cloning of the short and long 3'UTR into a transcription vector, respectively (Fig.3.10B, see Materials and Methods for details). *In vitro* transcription of both RNAs was driven by the T7 promoter and the resulting products were analyzed on a RNA formaldehyde denaturing gel (Fig.3.10B). The short RNA gave one

band at the correct size (1438nt, red arrow), whereas the long RNA had, apart from an expected band (4727nt, red arrow), an extra band at around 3000nt (Fig.3.10B). This 3000nt product persisted, regardless of the transcription conditions (data not shown), hinting at potential secondary structures at this position and thus a premature fall-off of the RNA polymerase. Nevertheless, both synthetic RNAs containing the corresponding proximal and distal targets were available and could be mixed at any desirable ratios. cDNA (2fg input) transcribed from the long RNA alone had a distal/proximal ratio of approximately 0.43, implying that the undesired 3000nt product was at around 1.3:1 in molar ratio mixed with the long RNA (Fig.3.10B). When mixing cDNA from the short and the long RNA at equal mass ratio (0.1fg or 2fg), the distal/proximal ratio was expected to have an average of approximately 0.14. This was confirmed experimentally (Fig.3.10B). Importantly, cDNA from the long RNA alone or the short and long RNA mixed at equal mass ratio also did not react to the RNase treatment (Fig.3.10B).

To this point, one might still argue about the simplicity of the chosen synthetic system that could not resemble the situation of RNA isolated from tissue. To address this, adult zebrafish total RNA extracted from the brain or the abdomen was mixed with the synthetic short/long RNA to create a situation mimicking the complexity of P7 OB RNA, namely a large amount of non-target RNA (4µg zebrafish total RNA) and a small amount of target RNA (2fg short RNA and 1.3fg long RNA) (Fig.3.10C). cDNA input was 100ng-400ng per reaction. With both brain and abdomen RNA additives, the distal/proximal ratio stayed at a comparable level, independent of the RNase treatment (Fig.3.10C). Though a significant increase of the distal/proximal ratio upon RNase treatment was seen between samples with the brain RNA additives, it could not be replicated in samples with the abdomen RNA additives (Fig.3.10C). Thus, it could more likely be attributed to statistical fluctuations than to an effect of the presence of large amount of non-target RNA or RNase treatment *per se*.

Taken together, RNA from both mouse origin (OE) and an artificial system with *in vitro* transcription could not, with the same experimental procedure, replicate the "RNase sensitivity" of cDNA measurement observed with cDNA from the P7 OB. This rules out the possibility of the experimental procedure *per se* being the cause of it. It raises the assumption that there might be some P7 OB unique factor that reacts to the RNase treatment and gives rise to a change in the distal/proximal ratio. Since it deals with RNase treatment, it is speculated that some unknown RNA molecule might contribute to this unique factor.
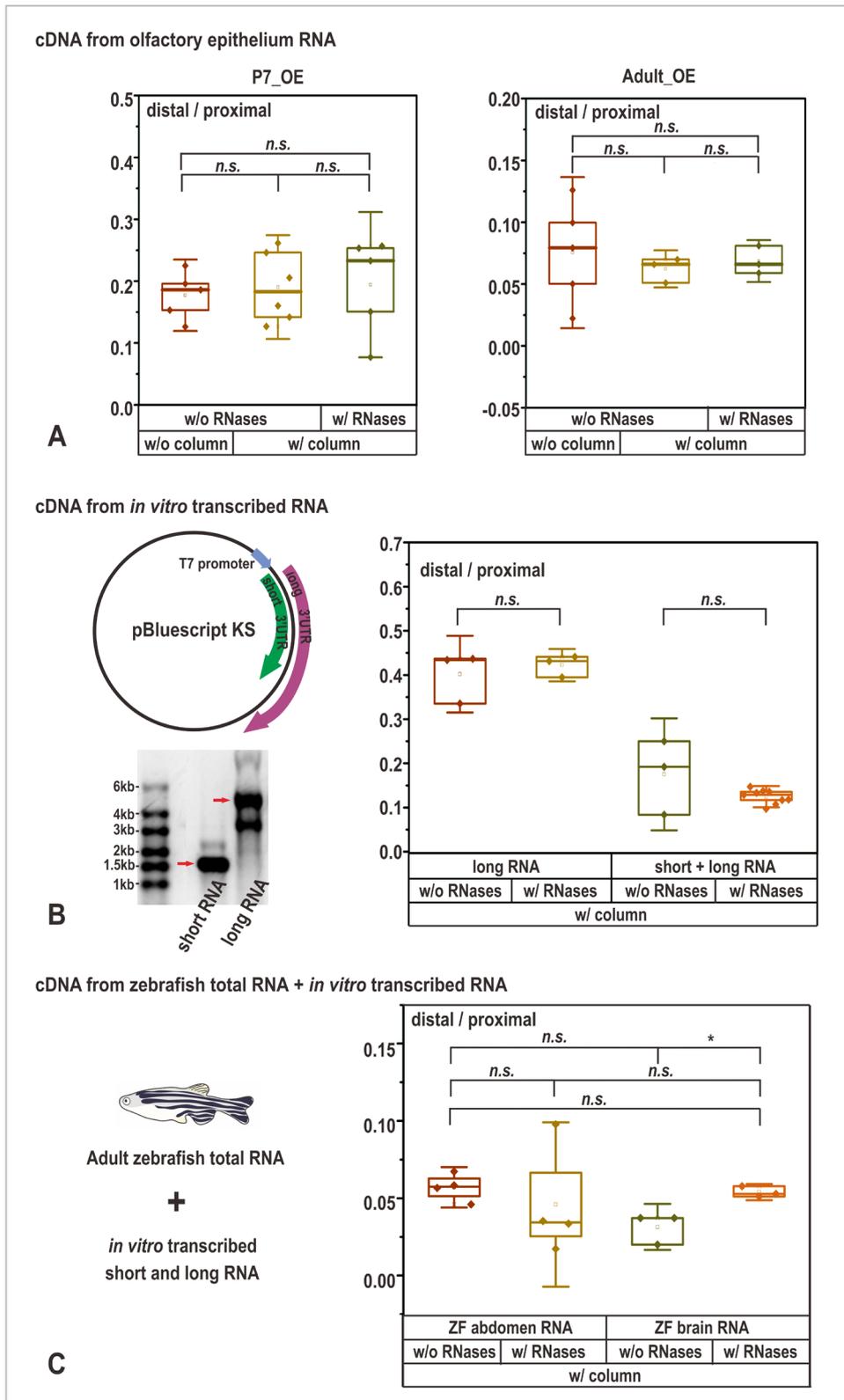
Figure 3.10. Control experiments with cDNA from various sources.
A. ddPCR with OE cDNA from different developmental stages. ddPCR was performed with OE cDNA from the P7 and adult stages. Random primed cDNA was used as template, combining different treatments. cDNA input was 20ng-40ng per reaction. RNase treatment with subsequent spin column purification does not affect the distal/proximal ratio, neither in P7 nor in adult OE. Each dot corresponds to one animal and the whiskers represent standard deviation; T-test with n.s.: $\alpha \geq 0.05$.
P7_OE, w/o RNases, w/o column: N=5, distal/proximal ratio 0.186.
P7_OE, w/o RNases, w/ column: N=6, distal/proximal ratio 0.183.

P7_OE, w/ RNases, w/ column: N=5, distal/proximal ratio 0.233.
Adult_OE, w/o RNases, w/o column: N=5, distal/proximal ratio 0.0793.
Adult_OE, w/o RNases, w/ column: N=3, distal/proximal ratio 0.066.
Adult_OE, w/ RNases, w/ column: N=3, distal/proximal ratio 0.066.
B. ddPCR with cDNA from *in vitro* synthetic RNA. ddPCR was performed with cDNA from the synthetic short/long RNA. Random primed, spin column purified cDNA was used as template combined with RNase treatment. cDNA input from the long RNA alone was 2fg per reaction and cDNA input from the short/long RNA mixed was 0.1fg or 2fg each per reaction. **Left.** Both the DNA for the short and long 3'UTR was cloned into a transcription vector. They were transcribed *in vitro* and checked by a RNA formaldehyde denaturing gel. The short 3'UTR shows one band with the expected size of 1438nt (red arrow). The long 3'UTR shows, apart from the correct size of 4727nt (red arrow), an extra band at 3000nt according to the RNA size standard. **Right.** RNase treatment with the subsequent spin column purification does not affect the distal/proximal ratio in the synthetic system, with cDNA from either the long RNA only or the short and long RNA mixed at equal mass ratio. Each dot corresponds to one replicate and the whiskers represent standard deviation; T-test with n.s.: α≥0.05.
long RNA, w/o RNases, w/ column: n=3, distal/proximal ratio 0.434.
long RNA, w/ RNases, w/ column: n=3, distal/proximal ratio 0.432.
short+long RNA, w/o RNases, w/ column: n=3, distal/proximal ratio 0.192.
short+long RNA, w/ RNases, w/ column: n=9, distal/proximal ratio 0.129.
C. ddPCR with cDNA from *in vitro* synthetic RNA mixed with zebrafish total RNA. Left. ddPCR was performed with cDNA from the synthetic short/long RNA mixed with adult zebrafish brain/abdomen total RNA to mimic the complexity of P7 OB RNA. Random primed, spin column purified cDNA was used as template combined with RNase treatment. cDNA input was 100ng-400ng per reaction, which was reversely transcribed with 4μg zebrafish total RNA mixed with 1.3fg long RNA and 2fg short RNA. **Right.** RNase treatment with the subsequent spin column purification mostly does not affect the distal/proximal ratio in the combined system. However, a weakly significant increase is seen upon RNase treatment with brain RNA additive. Each dot corresponds to one and the whiskers represent standard deviation; T-test with n.s.: α≥0.05, *: α<0.05.
ZF abdomen RNA, w/o RNases, w/ column: n=4, distal/proximal ratio 0.0575.
ZF abdomen RNA, w/ RNases, w/ column: n=4, distal/proximal ratio 0.0344.
ZF brain RNA, w/o RNases, w/ column: n=3, distal/proximal ratio 0.0371.
ZF brain RNA, w/ RNases, w/ column: n=3, distal/proximal ratio 0.0529.

3.3.3 Combination with Proteinase K treatment

Since the previous observation points towards the existence of some RNA molecule that might interfere with the measurement of the distal/proximal ratio and could be possibly removed by the RNase treatment, I further explored the possibility of whether it was purely a RNA-specific effect, by combining with other treatments such as the Proteinase K digest. Though the majority of the proteins were probably already removed in the phase separation step during RNA isolation by phenol/chloroform, tightly bound proteins might still be present. Proteinase K treatment was combined with RNase treatment on the P7 OB cDNA and in some cases EtOH precipitation was carried out instead of column purification to concentrate the cDNA (Fig.3.11A, see Materials and Methods for details).

In accordance with previous results, the distal/proximal ratios upon different treatments reacted mainly to the RNase treatment (Fig.3.11B). Speaking in detail, the distal/proximal ratio significantly decreased upon RNase treatment and a similar effect could barely be triggered by the other treatments such as the Proteinase K or different purification methods (Fig.3.11B). To ease the visualization, samples with or without RNase treatment (the same for Proteinase K treatment) were pooled and plotted (Fig.3.11C). This clearly demonstrated that RNases might affect the distal/proximal ratio, whereas Proteinase K might not (Fig.3.11C).
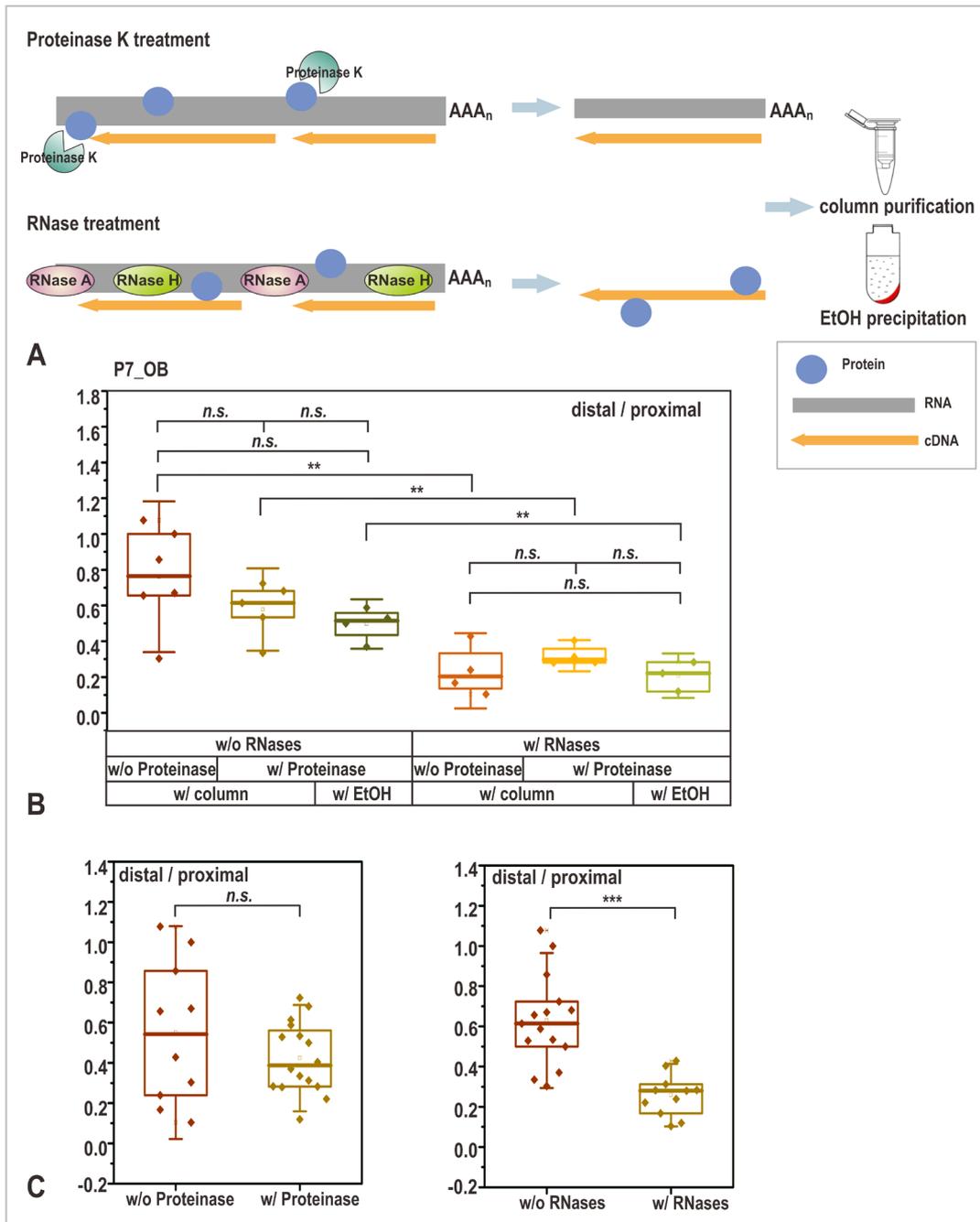
Figure 3.11. Effect of Proteinase K treatment on the distal/proximal ratio in P7 OB measured by ddPCR. ddPCR was performed with the P7 OB cDNA. Random primed, purified cDNA were used as template in combination with various treatments. cDNA input was 200ng-1000ng per reaction.

A. Schematic illustration of RNase/ Proteinase K treatment and the subsequent purification of OB cDNA. Proteinase K removed the proteins that bound to the RNA. RNase treatment included RNase H and RNase A, which aimed to remove the RNA in RNA-cDNA hybrids (RNase H) and the single-stranded RNA (RNase A). cDNA was purified either with spin column or EtOH precipitation.

B. Effect of Proteinase K treatment and different purification methods on the distal/proximal ratio. The distal/proximal ratios mainly differ in response to the RNase treatment, not in response to the Proteinase K treatment or the purification methods. Each dot corresponds to one animal and the whiskers represent standard deviation; T-test with n.s.: $\alpha \geq 0.05$, *: $\alpha < 0.05$, **: $\alpha < 0.01$.

w/o RNases, w/o Proteinase, w/ column: N=6, distal/proximal ratio 0.764.
w/o RNases, w/ Proteinase, w/ column: N=5, distal/proximal ratio 0.614.
w/o RNases, w/ Proteinase, w/ EtOH: N=4, distal/proximal ratio 0.515.
w/ RNases, w/o Proteinase, w/ column: N=4, distal/proximal ratio 0.203.
w/ RNases, w/ Proteinase, w/ column: N=4, distal/proximal ratio 0.297.
w/ RNases, w/ Proteinase, w/ EtOH: N=3, distal/proximal ratio 0.221.

C. Comparison between RNase and Proteinase K treatment. Data from B were pooled with regard to the RNase treatment or Proteinase K treatment. RNase treatment causes a significant decrease of the distal/proximal ratio, whereas Proteinase K treatment not. Each dot corresponds to one animal and the whiskers represent standard deviation; T-test with n.s.: α≥0.05, *: α<0.05, **: α<0.01, ***: α<0.001.
w/o Proteinase: N=10, distal/proximal ratio 0.542.
w/ Proteinase: N=16, distal/proximal ratio 0.387.
w/o RNases: N=15, distal/proximal ratio 0.614.
w/ RNases: N=11, distal/proximal ratio 0.280.

Up to this point, it has been consistently observed that the ddPCR P7 OB cDNA measurements exhibit some kind of sensitivity towards the RNase treatment. It occurrs only under highly specific biological conditions with specific treatment, *i.e.*, in P7 axon termini with RNase treatment, as series of control experiments with cDNA from other sources or Proteinase K treatment simply cannot replicate the effect. This strengthens the hypothesis that there could indeed be some RNA molecule interacting somewhere within the amplicons and that this interaction somehow affects the PCR amplification.

3.3.4 Gain of proximal targets with RNase treatment

To address the question of where the potential interaction might take place, the distal/proximal ratio was dismantled into the absolute copy numbers of the proximal and distal targets. The target copy numbers from P7 OB were calculated from the pooled samples from Fig.3.11C (with or without RNase treatment) and normalized to 1ng input cDNA. Notably, the amount of apparent proximal targets increased significantly upon RNase treatment, whereas that of the distal targets remained at a comparable level (Fig.3.12A). In comparison to this, both the proximal and distal copy numbers from P7 OE were not influenced by different treatments (Fig.3.12A, samples from Fig.3.10A). Thus, the proximal target was somehow affected and a loss of proximal amplification occured without RNase treatment (Median values in Tab.3.2).

Noteworthy, even without RNase treatment the loss of the proximal amplification was not complete, namely there was always a portion of the proximal targets (0.128) left that was comparable to the distal targets (0.104). This numerical coincidence promoted the assumption that the loss of the proximal amplification might originate from the short isoform. To test this assumption, a distal gene specific primer located around 600bp 3' downstream of the distal amplicon was used for cDNA synthesis instead of random primers, so that the short isoform would be excluded from the scheme (Fig.3.12B). With this set of cDNA, one would expect, if the proximal target of the long isoform was affected as well, the distal/proximal ratio to go above 1 and it could only be rescued to 1, when the RNase treatment was done. In the other scenario with the proximal target of the long isoform being spared from inhibition, the ratio would stay at 1. As a positive control, the distal/proximal ratio from P7 OE cDNA primed with the distal GSP was found close to 1 (Fig.3.12B). The experimental data suggested that the distal/proximal ratio in the P7 OB was not altered by the RNase treatment, indicating that the proximal target on the long

isoform might be spared from the binding of the PCR inhibitor. It was, however, above 1 under both conditions, speaking for a loss of proximal targets independent of the RNase treatment, when compared to the OE cDNA (Fig.3.12B). Potentially, other factors such as reverse transcription efficiency or RNA secondary structures in the OB might also contribute to the suppression of the proximal targets due to its larger distance to the primer, and would thus lead to a final ratio of greater than 1. These factors must be bulb specific.

Taken together, the aforementioned change of the distal/proximal ratio is likely to be caused by the loss of proximal target amplification without RNase treatment, and the loss equals numerically the fraction from the short isoform.
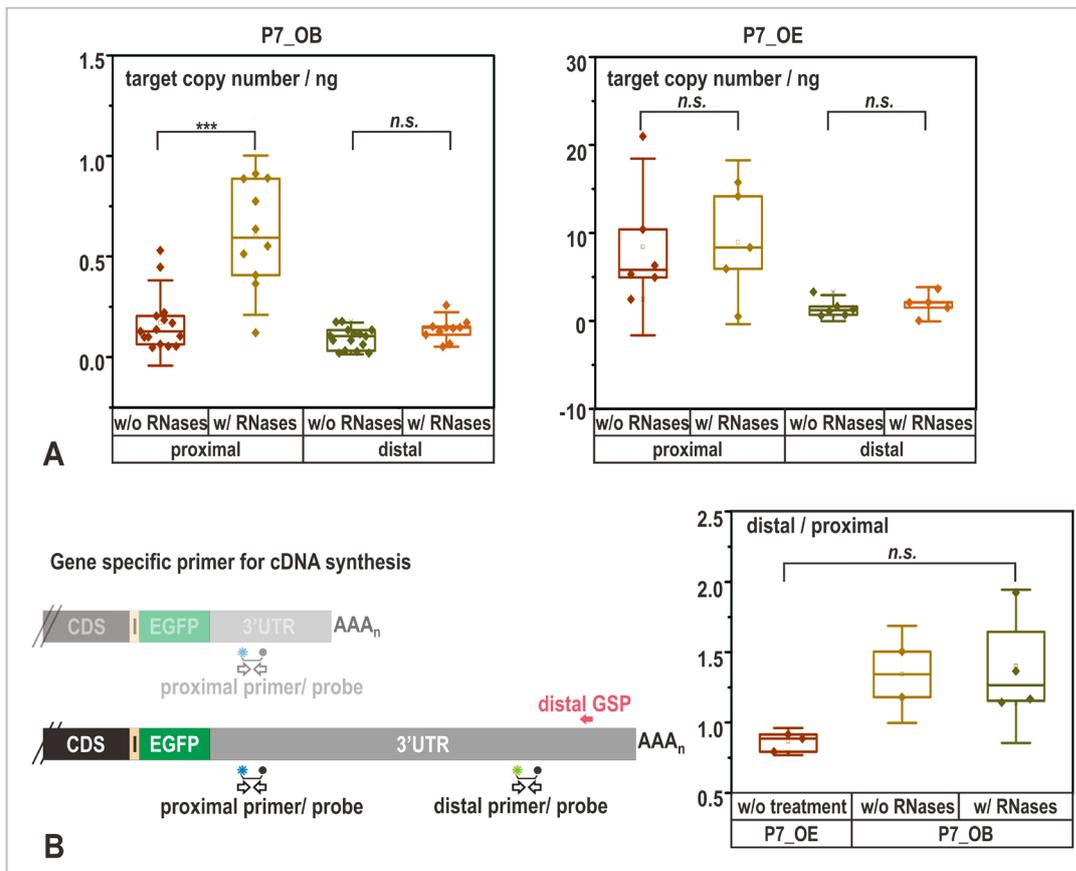

Figure 3.12. Effect of RNase treatment on the proximal targets.
A. Apparent absolute copy number of the proximal and distal targets in the P7 OE and OB under different treatments. P7 OB and OE target copy numbers are derived from Fig.3.11C and Fig.3.10A respectively, by normalizing the input to 1ng cDNA. In the P7 OB, the apparent amount of the proximal targets increases significantly upon RNase treatment, while that of the distal targets is not changed. Contrasting this, both the proximal and distal targets in the P7 OE stay at a comparable level, regardless of the RNase treatment. Each dot corresponds to one animal (P7_OB: w/o RNases N=15; w/ RNases N=10. P7_OE: w/o RNases N=6; w/ RNases N=5) and the whiskers represent standard deviation; T-test with n.s.: $\alpha \geq 0.05$, *: $\alpha < 0.05$, **: $\alpha < 0.01$, ***: $\alpha < 0.001$.
B. ddPCR with cDNA primed with the distal gene specific primer. ddPCR was performed with the P7 OE/OB cDNA. GSP primed cDNA was used as template in combination with various treatments. Input of OE cDNA and OB cDNA was 40ng and 200ng-400ng per reaction, respectively. **Left.** The distal GSP (37A_3'UTR_4124R) is located ca. 600bp downstream of the distal amplicon and can only recognize the long isoform. **Right.** cDNA from the long isoform in the P7 OB shows no response towards the RNase treatment, but the distal/proximal ratio is greater than 1 under both conditions. cDNA of the long isoform in the P7 OE works as control and the distal/proximal ratio is close to 1. Each dot corresponds to one animal and the whiskers represent standard deviation; T-test with n.s.: $\alpha \geq 0.05$.

P7_OE, w/o treatment: N=3, distal/proximal ratio 0.885.
P7_OB, w/o RNases: N=2, distal/proximal ratio 1.34.
P7_OB, w/ RNases: N=4, distal/proximal ratio 1.27.

Table 3.2. Median value of the apparent absolute copy number of the proximal and distal targets upon RNase treatment in the P7 OB, normalized to 1ng cDNA input.

	proximal target	distal target
w/ RNases	0.594	0.146
w/o RNases	0.128	0.104

3.3.5 Testing the model of a hypothetical inhibitory RNA

To explain the effect of the RNase treatment, I suggest the following working hypothesis (Fig.3.13). Since the proximal amplicon was seen responding to the RNase treatment, the interaction with an unknown RNA molecule is assumed to take place within the proximal target, due to full or partial complementarity. The binding then hinders the DNA polymerase extension during PCR. DNA polymerases can have DNA- but not RNA-exonuclease activity. When treated with RNases, this unknown RNA molecule could be removed, resulting in normal amplification of the proximal target. Otherwise a loss of the proximal amplification and an increase of the distal/proximal ratio would be expected. The binding is assumed to occur on the short isoform.
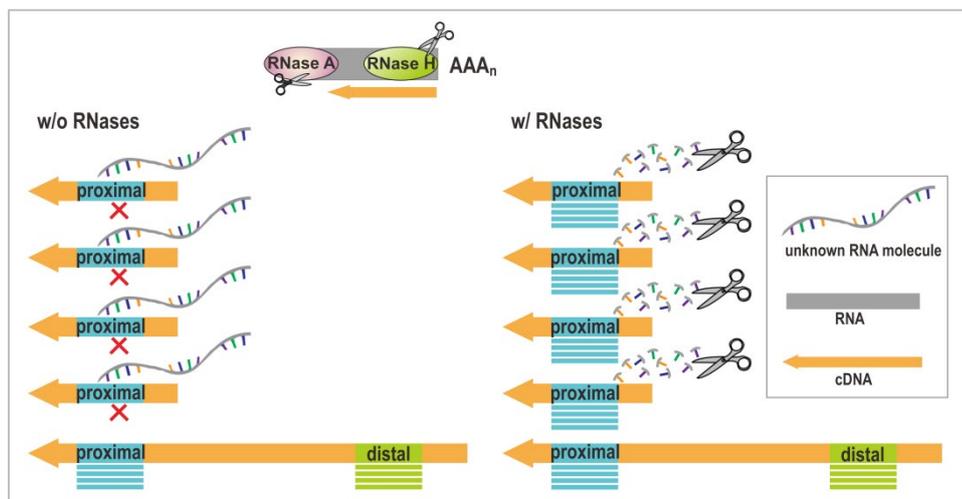

Figure 3.13. Working hypothesis on the effect of RNase treatment on the proximal target of the short isoform.
Given that the proximal target is affected by the RNase treatment, it is assumed that there is some unknown RNA molecule interacting within the proximal amplicon by full or partial complementarity. **Left.** This interaction inhibits DNA polymerase extension during PCR, leading to the loss of proximal amplification and an increased distal/proximal ratio. **Right.** Once treated with RNases, this RNA molecule can be removed, resulting in normal PCR amplification.

There are different RNA species *in vivo* that fulfill regulatory functions. To test the working hypothesis, a "gain-of-function" experiment was performed with the addition of natural small RNA molecules. Natural small RNA molecules (<200nt) were isolated from the P7 OB total RNA using the miRNeasy Mini Kit (Qiagen) and were added to the cDNA synthesized from *in vitro* transcribed RNA (Fig.3.14A), which did not show any "RNase sensitivity" previously

(Fig.3.10B). One drawback in the design here was that the isolated small RNA molecules could neither be quantified nor qualified due to the lack of proper setup. Therefore, the "small RNA concentration" given here corresponded to the amount of total RNA used for its isolation and the cDNA concentration here referred to the purified cDNA from the short/long RNA at equal mass ratio.

In accordance with previous results, equal mass ratio of cDNA from the short/long RNA without small RNA addition faithfully reproduced the distal/proximal ratio of approximately 0.14 (Fig.3.14B). Surprisingly, with increasing amount of small RNA molecules (from 0.4µg, 1.8µg or 4.8µg total RNA), the distal/proximal ratio increased drastically, mimicking the original effect seen in the P7 OB without RNase treatment (Fig.3.9B, Fig.3.14B).
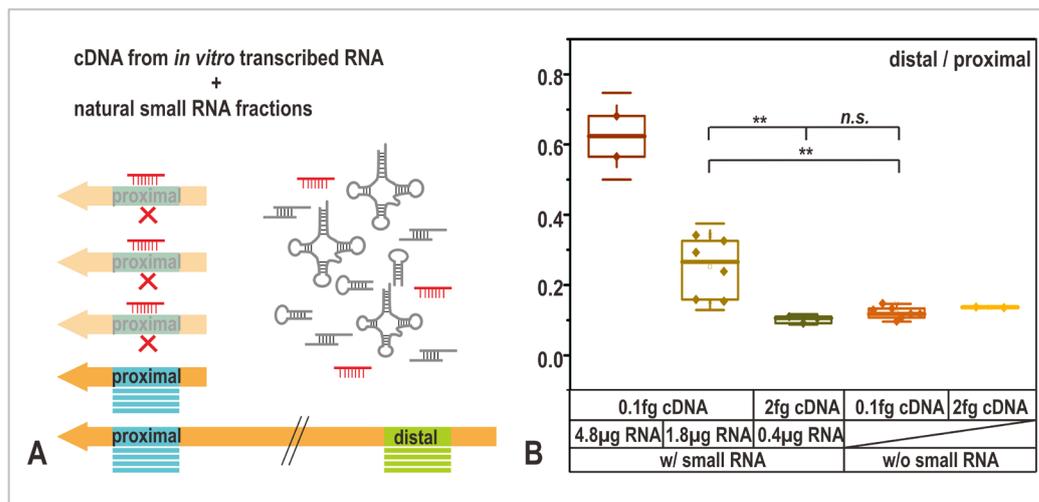

Figure 3.14. Experimental setup for testing the inhibitory effect of small RNA molecules.
ddPCR was performed with cDNA from *in vitro* synthetic short and long RNA, with small RNA additives. Random primed, RNase treated and purified cDNA from short/long RNA at equal mass ratio (0.1fg or 2fg of each) was used as template. Input of small RNA refers to the amount of total RNA used for its isolation.
A. Schematic illustration of the experimental setup. Small RNA was isolated from the P7 OB total RNA. It was added to the cDNA from synthetic short and long RNA to test its inhibitory potential. If such interaction would occur, an increase of the distal/proximal ratio is expected, since some of the proximal target could be inhibited by components from the isolated small RNA fractions.
B. Effect of small RNA additives on the distal/proximal ratio in the synthetic system. With increasing amount of small RNAs, the distal/proximal ratio increases significantly, mimicking the P7 OB situation without RNases treatment. Each dot is one replicate and the whiskers represent standard deviation; T-test with n.s.: $\alpha \geq 0.05$, *: $\alpha < 0.05$, **: $\alpha < 0.01$.
w/ small RNA, 0.1fg cDNA, 4.8µg RNA: n=2, distal/proximal ratio 0.623.
w/ small RNA, 0.1fg cDNA, 1.8µg RNA: n=6, distal/proximal ratio 0.266.
w/ small RNA, 2fg cDNA, 0.4µg RNA: n=3, distal/proximal ratio 0.105.
w/o small RNA, 0.1fg cDNA: n=7, distal/proximal ratio 0.118.
w/o small RNA, 2fg cDNA: n=2, distal/proximal ratio 0.136.

3.3.6 "RNase sensitivity" in a developmental context

Since mOR37A transcript isoforms are regulated between the juvenile and adult stages in the OE, the observed "RNase sensitivity" in P7 OB was also investigated under a different developmental perspective, namely in the adult stage. Following the identical experimental procedure, the distal/proximal ratio in the adult OB was surprisingly not influenced heavily by the RNase

treatment, neither by RNase A, RNase H combined nor separated (Fig.3.15). Despite a weakly significant increase in the distal/proximal ratio between the OE and OB when no RNase treatment was performed, the observed difference was strongly reduced in comparison to P7 OB, namely here the distal/proximal ratio was 0.065-0.075 in the cell soma and the axon termini with RNase treatment, and 0.179 in the axon termini without RNase treatment (Fig.3.15), speaking for potentially a minor loss of the proximal amplification. At P7, the loss of proximal amplification, however, corresponds almost to the entire short isoform population (Fig.3.12A).
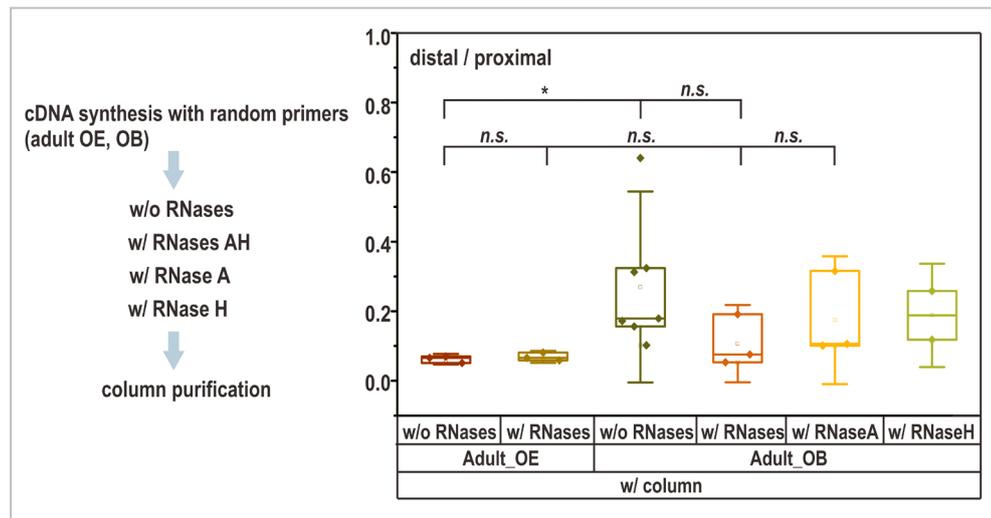

Figure 3.15. Effect of RNase treatment on the distal/proximal ratio in the adult OE and OB.
ddPCR was performed with the adult OE and OB cDNA, with the same experimental procedure as in the P7 OB. Random primed cDNA was either directly purified with spin columns or was subject to RNase treatment prior to column purification. RNase treatment included RNase H and RNase A, either combined or separated, which aimed to remove the RNA in RNA-cDNA hybrid or the single-stranded RNA. OB cDNA input was 400ng-1400ng per reaction and OE cDNA input was 20ng-40ng per reaction. The distal/proximal ratio in both adult OE and OB is not affected by the RNase treatment. Each dot corresponds to one animal and the whiskers represent standard deviation; T-test with n.s.: $\alpha \geq 0.05$, *: $\alpha < 0.05$.
Adult_OE, w/o RNases: N=3, distal/proximal ratio 0.066.
Adult_OE, w/ RNases: N=3, distal/proximal ratio 0.066.
Adult_OB, w/o RNases: N=7, distal/proximal ratio 0.179.
Adult_OB, w/ RNases: N=3, distal/proximal ratio 0.0755.
Adult_OB, w/ RNase A: N=3, distal/proximal ratio 0.106.
Adult_OB, w/ RNase H: N=2, distal/proximal ratio 0.188.

In sum, the experimental data from RNA of the olfactory axon termini strongly suggests the existence of some RNA interaction partner that interferes with the proximal target of the short mOR37A transcripts. Such interference is prominent at P7 and gradually gets lost during development. Though the identity of this RNA interaction partner remains unclear, it is a step forward towards finding any receptor-associated molecules that might finally be involved in the axon guidance.

3.4 Comparative analysis of other odorant receptors

3.4.1 Potential alternative polyadenylation in other odorant receptors

So far, alternative mOR37A transcripts have displayed a differential regulation, as well as potential interaction partners in the axon termini during development. To investigate the generality of this mechanism, 3'UTR APA was checked in other OR transcripts by 3' RACE. In total, three other receptors, mOR37C, olfr701 and olfr702, were chosen.

In the mouse, the *mOR37* subfamily is localized on chromosome 4, comprising five highly related genes (*mOR37A-mOR37E*), with *mOR37D* being a pseudogene (Strotmann et al., 1999). *mOR37C* is in close vicinity to *mOR37A*, with 18.7kb intergenic space, and the transcription orientation of both is opposite (Fig.3.16A). mOR37C 3' RACE was performed with the P7 OE RNA with 35x cycles, using a gene specific primer [37C_GSP_1018F] located 60bp downstream of its coding region. The RACE-PCR was first loaded on an agarose gel. For the nested PCRs, different forward primers were chosen based on the band size obtained from the RACE-PCR. All nested PCR products were sent for sequencing. Among them, four mapped to genomic A rich regions (data not shown) and the rest revealed two potential 3' ends (Fig.3.16B). As sequencing data suggested, the 3'UTR of the two identified isoforms was 761nt and 5934nt in length, with the polyadenylation signals *"AAGAAA"* and *"ATTAAA"*, respectively (red rectangles, Fig.3.16B).

Figure 3.16. Potential 3'UTR alternative polyadenylation of mOR37C transcripts in the mOR37A-IRES-tauEGFP strain analyzed by 3' RACE.
A. Genomic organization of the receptor genes *mOR37A* and *mOR37C*. *mOR37A* and *mOR37C* belong to the *mOR37* subfamily located on mouse chromosome 4. They are in close proximity with 18.7kb between their coding regions.
B. 3' ends of mOR37C transcripts by sequencing of 3' RACE nested PCR products. Two potential 3' ends of the mOR37C transcripts were identified by sequencing, with their 3'UTRs being 761nt and 5934nt in length. Red rectangles mark the polyadenylation signals.
Nested primer for end 1: 37C_old_primer_F; NUP.
Nested primer for end 2: 37C_3'UTR_6428F; NUP.

The other two receptors olfr701 and olfr702 were picked randomly. They belong to the *mOR283* subfamily on chromosome 7. Their transcription orientation is opposite and their intergenic distance is merely 4.5kb (Fig.3.17A). 3' RACE was performed with P7 OE RNA with 35x cycles. For olfr701, two gene specific primers were tested, with one [Olfr701_GSP761F] in the coding region and the other [Olfr701_GSP1265F] 313bp downstream of the coding region. The RACE-PCR was first loaded on an agarose gel. For the nested PCRs, different forward primers were chosen based on the band size obtained from the RACE-PCR. The nested PCRs revealed three potential 3' ends of the receptor olfr701, with their 3'UTRs being 1678nt, 2499nt and 4361nt in length, respectively (Fig.3.17B). All of them possessed the canonical polyadenylation signals of *"AATAAA"* or *"ATTAAA"* (red rectangles, Fig.3.17B). For olfr702, likewise two gene specific primers were chosen, with one [Olfr702_GSP812F] in the coding region and the other [Olfr702_3'UTR_1014F] 57bp downstream of the coding region. The nested PCRs identified two transcript isoforms with their 3'UTR length being 1009nt and 1680nt, and they both had the canonical polyadenylation signals of *"ATTAAA"* or *"AATAAA"* (red rectangles, Fig.3.17C). Thus, with all likelihood, the 3'UTR of the longest transcript isoform of olfr701 overlaps with both transcript isoforms of olfr702.

In summary, all three receptors seem to have multiple transcript isoforms potentially generated by alternative polyadenylation, rendering it a more general phenomenon in the OR transcripts.

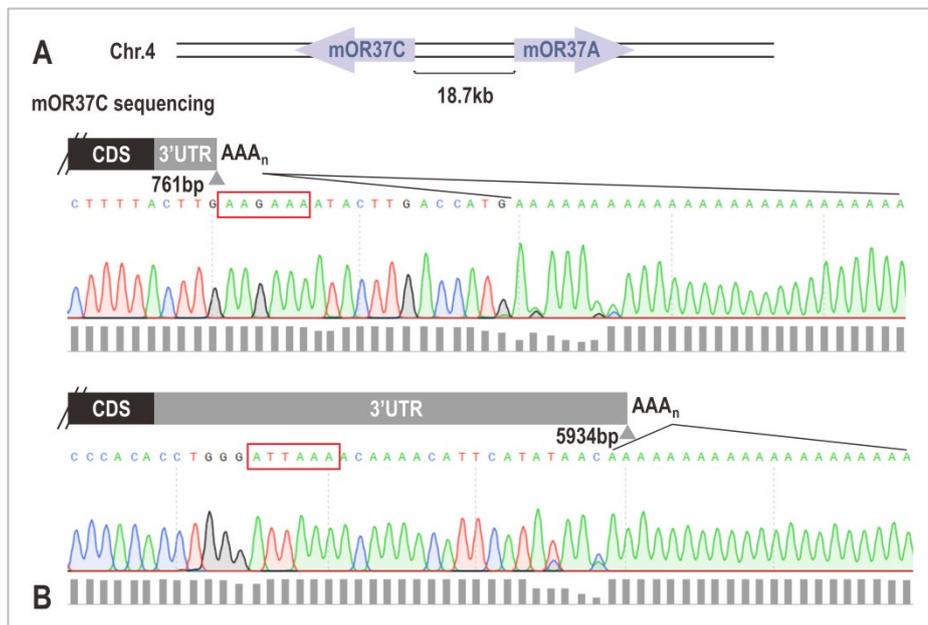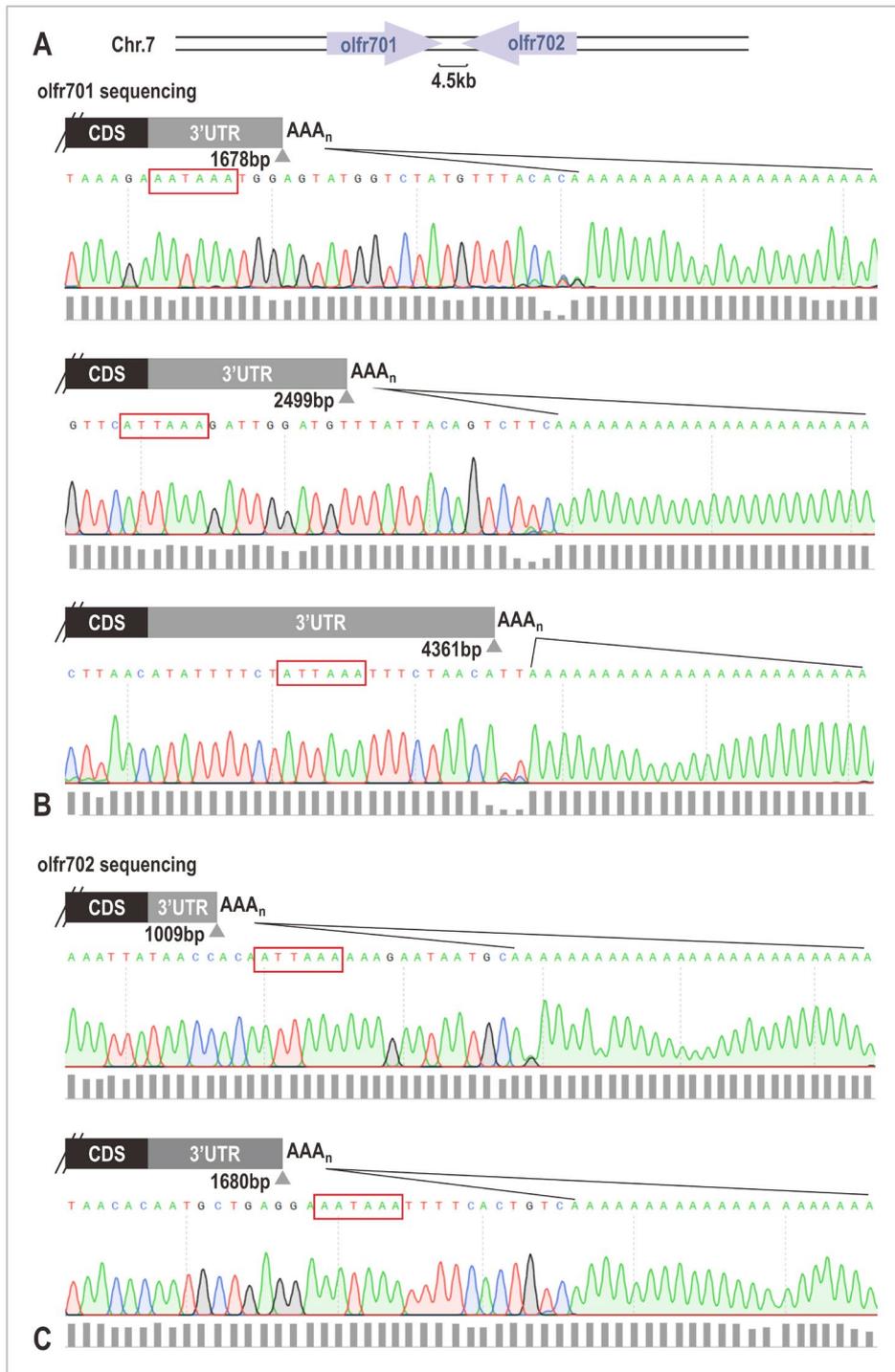**Figure 3.17. Potential 3'UTR alternative polyadenylation of olfr701 and olfr702 transcripts in the mOR37A-IRES-tauEGFP strain analyzed by 3' RACE.**

A. Genomic organization of the receptors genes *olfr701* and *olfr702*. *olfr701* and *olfr702* belong to the *mOR283* subfamily located on mouse chromosome 7. Their transcription orientation is opposite and the distance between their coding regions is merely 4.5kb.

B. 3' ends of olfr701 transcripts by sequencing of 3' RACE nested PCR products. Three potential 3' ends of the olfr701 transcripts were identified by sequencing, with their 3'UTR being 1678nt, 2499nt and 4361nt in length, respectively. Red rectangles mark the polyadenylation signals.

Nested primer for end 1: Olfr701_3'UTR_2094F; NUP.

Nested primer for end 2: Olfr701_3'UTR_3153F; NUP.

Nested primer for end 3: Olfr701_3'UTR_4770F; NUP.

C. 3' ends of olfr702 transcripts by sequencing of 3' RACE nested PCR products. Two potential 3' ends of the olfr702 transcripts were identified by sequencing, with their 3'UTR being 1009nt and 1680nt in length. Red rectangles mark the polyadenylation signals.

Nested primer for end 1: Olfr701_3'UTR_3903R; NUP.
Nested primer for end 2: Olfr701_dist_qPCR_R; NUP.

3.4.2 Expression pattern of mOR37C and olfr701 transcript isoforms in the olfactory epithelium

Further characterization with regard to the expression pattern of different transcript isoforms was achieved by RNA-FISH on OE sections of the P7 mice. Two antisense RNA probes were designed in the proximal and distal region of mOR37C and olfr701 transcripts respectively, with the proximal probe labeling all possible transcript isoforms and the distal probe only the long isoform (Fig.3.18A).

Unlike mOR37A, mOR37C and olfr701 transcripts labeled by both the proximal and distal probes showed a high degree of co-localization, in other words, almost no cells were detected only with the proximal probe labeling (Fig.3.18B.C). Interestingly, the mOR37C-expressing OSNs are known to be restricted in a small patch in the epithelium and were in fact densely packed only in some regions (Fig.3.18C), while the olfr701-expressing OSNs follow the zonal expression rule and were more sparsely distributed across the epithelium (Fig.3.18B). The subcellular details of the receptor mOR37C transcripts were further probed with Airyscan imaging. Remarkably, the subcellular localization of the proximal and distal probe labeled transcripts were also eminently comparable (Fig.3.18C'.C''), regardless of the maturation state of the individual cell (apical or basal). Similar to mOR37A, the nuclear "transcriptional hotspot" next to the heterochromatin clusters was detected in mOR37C cells as well (Fig.3.18C'.C''). For quantifications, the "relative height" of each labeled cell across the OE was measured by dividing the "total OE thickness" by the "distance to basal side" (Fig.3.18D). By taking the mOR37C-expressing OSNs as an example, cells labeled by both probes were found in the entire OE with a peak accumulation in the mature OSNs (0.6-0.8) (Fig.3.18D). Thus, contrasting mOR37A, both its closely related family member mOR37C or receptor olfr701 from other OR subfamily, do not exhibit an on/off maturation-dependent regulation of their alternatively polyadenylated isoforms with the long isoforms being restricted to immature cells.

Up to this point, there seems to be no on/off developmental regulation based on the maturation level of a given OSN for mOR37C and olfr701. In fact, a simple co-localization is an indicator of a lack of qualitative switching on/off of the transcript isoforms. A quantitative regulation should, however, be addressed with quantitative *in situ* hybridization or quantitative PCR.

Figure 3.18. Distribution pattern of olfr701 and mOR37C transcript isoforms in the P7 OE analyzed by RNA-FISH.

Double color RNA-FISH was performed on the P7 OE sections with Fluorescein- and DIG-labeled antisense RNA probes, and Fluorescein-POD, DIG-POD catalyzed Alexa488-, Cy3-methyldopamine color reactions, respectively.

A. Schematic illustration of the *in situ* antisense RNA probes. For both receptors, the proximal probe detects all isoforms, while the distal probe detects only the long isoform.

Olfr701 proximal probe: derived from PCR template with Si_olfr701_prox_F; Ki_olfr701_prox_R. 476bp

Olfr701 distal probe: derived from PCR template with Si_olfr701_dist_F; Ki_olfr701_dist_R. 455bp

mOR37C proximal probe: derived from PCR template with 37C_old_primer_F; 37C_old_primer_R. 470bp

mOR37C distal probe: derived from PCR template with Si_37C_3'UTR_4760F; Ki_37C_3'UTR_5381R. 621bp

B. Representative image of olfr701 proximal and distal probes labeled cells. Cells labeled by both probes show a high degree of co-localization. Image acquisition with Zeiss ApoTome, 20x/0.5 EC Plan-Neofluar objective. Scale bars, 50µm.

C. Representative image of mOR37C proximal and distal probes labeled cells. Cells labeled by both probes show a high degree of co-localization. Image acquisition with Zeiss ApoTome, 20x/0.5 EC Plan-Neofluar objective. Scale bars, 50µm.

C'.C''. Subcellular distribution of mOR37C proximal and distal probes labeled transcripts. At subcellular level, both the proximal and distal probes labeled transcripts also display a high degree of co-localization. This co-localization is independent of the cell maturation level. Image acquisition with Zeiss LSM 800 with Airyscan, 40x/water immersion objective. Z-stack with orthogonal projection. Scale bars, 5µm.

D. Quantification of the "relative height" of mOR37C proximal and distal labeled cells in the OE. The "relative height" is calculated by dividing the "total OE thickness" by the "distance to basal side". A value towards 1 means apical localization (mature OSNs) and a value towards 0 means basal localization (immature OSNs). mOR37C long isoform expressing cells exhibit a similar distribution pattern as the entire population, both with a peak in the mature OSNs (0.6-0.8). N: independent experiments; n: number of analyzed cells.

3.4.3 Expression level of olfr701 transcript isoforms in the immature and mature neurons

To investigate the possibility of a quantitative maturation-dependent regulation of different transcript isoforms, the distal/proximal ratio measured from RNA-FISH was calculated for both the immature and mature populations and compared.

Olfr701 was taken as an example. As described in section 3.4.2, two antisense RNA probes were designed to target the proximal and distal regions of the olfr701 transcripts, respectively (Fig.3.18A). Given the permanent co-localization of the proximal and distal probe labeled cells (Fig.3.18B), the integrated density of each individual cell from both labeling was measured with ImageJ (Fiji), applying the algorithm "RenyiEntropy" as thresholding method for the cell area selection (see Fig.3.5A and Materials and Methods for details). Since the "relative height" of individual cell in the epithelium contains the information about their maturation state, the basal 50% epithelium ("relative height" < 0.5) was considered to be the immature compartment and the apical 30% epithelium the mature compartment ("relative height" > 0.7). The cells localized in between were not taken into consideration due to the potential overlap of the two populations. A direct comparison of the distal/proximal intensity ratio between different populations can cancel out the efficiency difference caused by the detection system (*e.g.*, DIG- or Fluorescein-POD, Alexa488- or Cy3-methyldopamine) and can therefore provide information about a maturation-dependent regulation.

The labeling of the proximal and distal probes was swapped in two sets of independent experiments (Fig.3.19A), so that the detection system coupled to each probe could also be exchanged. For image acquisition, the exposure time of both channels was set identical in each independent experiment for all the cells analyzed. For different individuals, the measured fluorescent intensity was normalized to an equivalent of 90ms exposure time and cells from different individuals were pooled together in the quantification. In both cases, the distal/proximal intensity ratio, regardless of the detection system (distal_Fluo_Alexa488/proximal_Dig_Cy3 or distal_Dig_Cy3/proximal_Fluo_Alex488) was not significantly different between the immature

and mature populations, ruling out a potential maturation-dependent regulation of the olfr701 transcript isoforms (Fig.3.19B).
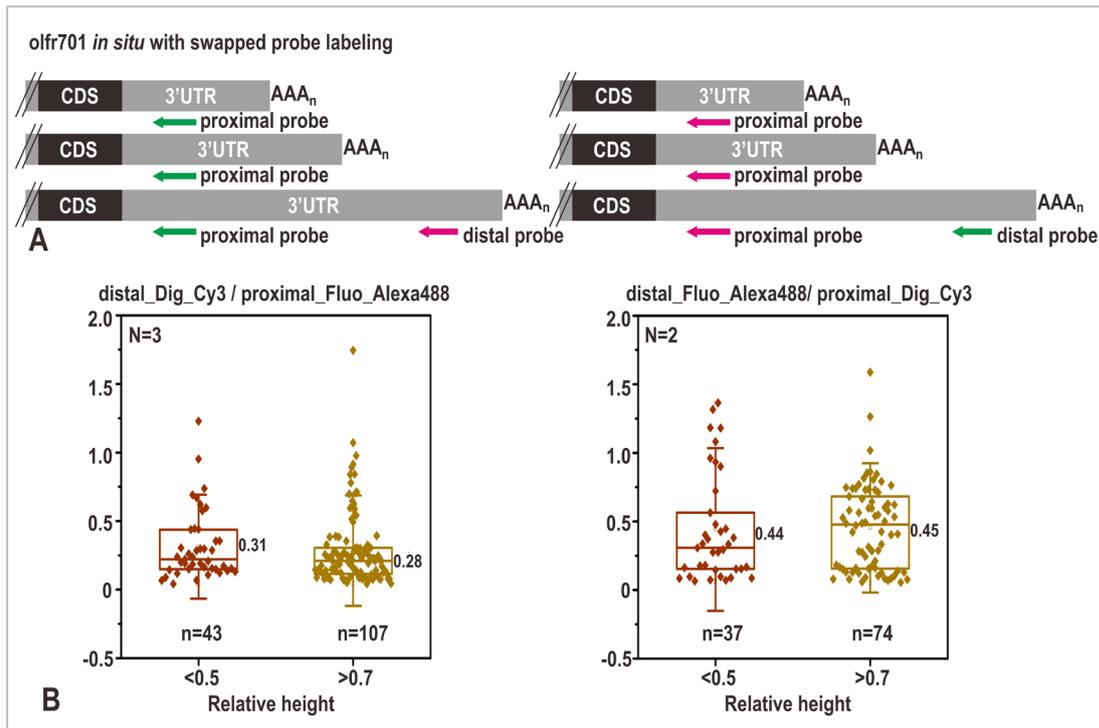

Figure 3.19. Expression level of olfr701 transcript isoforms in the immature and mature neurons in the P7 OE by intensity measurement of RNA-FISH.
Double color RNA-FISH was performed on the P7 OE sections with Fluorescein- or DIG-labeled antisense RNA probes and Fluorescein-POD or DIG-POD catalyzed Alexa488- or Cy3-methyldopamine color reactions. Fluorescent intensity of both channels was measured in individual cell with ImageJ (Fiji), using "RenyiEntropy" as the thresholding method. The distal/proximal intensity ratio was calculated and compared between the immature ("relative height" < 0.5) and the mature ("relative height" > 0.7) populations.
A. Schematic illustration of *in situ* antisense RNA probes. The proximal probe detects all isoforms, while the distal probe detects only the longest isoform. The labeling of the proximal and the distal probes was swapped in two sets of experiments, namely proximal_Fluo_Alexa488 with distal_Dig_Cy3, or proximal_Dig_Cy3 with distal_Fluo_Alexa488.
Proximal probe: derived from PCR template with Si_olfr701_prox_F; Ki_olfr701_prox_R. 476bp
Distal probe: derived from PCR template with Si_olfr701_dist_F; Ki_olfr701_dist_R. 455bp
B. Quantification of the distal/proximal intensity ratio in the immature and mature populations. Independent of the detection system, the distal/proximal intensity ratio between the immature and mature population does not differ significantly, indicating no maturation-dependent regulation of the olfr701 transcript isoforms. The exposure time was normalized to 90ms for each individual. Each dot corresponds to one cell and the whiskers represent standard deviation
distal_Dig_Cy3/proximal_Fluo_Alexa488: N=3, <0.5 n=43, >0.7 n=107.
distal_Fluo_Alexa488/proximal_Dig_Cy3: N=2, <0.5 n=37, >0.7 n=74.

3.4.4 Expression level of mOR37C and olfr701 transcript isoforms in the olfactory epithelium

Despite no maturation-dependent regulation of the olfr701 isoforms, the possibility of an age-dependent regulation of the transcript isoforms from the receptors mOR37C and olfr701 was probed by qPCR with SYBR Green labeling using OE RNA from P1, P7 and adult mice. Absolute quantification of both the proximal and distal targets was achieved by calibration to the standard curve obtained from the serial dilutions of templates of known concentrations. Similar to

mOR37A, the final quantification was presented as the ratio between the distal and proximal targets (total transcripts) from the same cDNA sample, *i.e.*, distal/proximal. As only the ratio was evaluated, no housekeeping genes were needed in this case. Typical input for OE qPCR was 30ng of random primed, non-purified cDNA. The concentration here referred to the amount of total RNA in the cDNA synthesis, assuming a conversion efficiency of 100%.

For mOR37C, two sets of primers were designed, with one in the proximal region targeting both isoforms and the other exclusively the long isoform (Fig.3.20A). Contrasting a maturation-independent co-localization of the proximal-positive with the distal-positive cells in RNA-FISH, the distal/proximal ratio showed a significant reduction between juvenile (P1/P7) and adult stages (Fig.3.20C). Notably, the ratio here was very similar to the distal/proximal ratio in receptor mOR37A (Fig.3.4B).

For olfr701, due to the overlapping 3'UTRs of the receptors olfr701 and olfr702, the choice of primers was limited. Based on the findings from 3' RACE, two sets of primers were designed with one in the proximal region of all three isoforms of olfr701 and the other in the distal region of the longest isoform (Fig.3.20B). The distal primers flanked the 3' end of the olfr702 long isoform, so that a co-amplification from olfr702 should be avoided (Fig.3.20B). To first evaluate the relative abundance of olfr701 and olfr702 transcripts, one set of proximal primers was designed after the coding region of olfr702, flanking the 3' end of the longest isoform of olfr701 (Fig.3.20B). When combining the proximal primers of these two receptors, the relative abundance of both populations was revealed to be constant between the P7 and adult stages, with olfr701 being slightly higher expressed than olfr702 (mean value olfr701/olfr702=1.7) (Fig.3.20C). Similar to mOR37C, an age-dependent regulation of the longest isoform was also seen in receptor olfr701, in which the distal/proximal ratio was significantly reduced between juvenile (P1/P7) and adult stages (Fig.3.20C). Notably, the distal/proximal ratio of the receptor olfr701 in different developmental stages was much higher than that of mOR37A and mOR37C (Fig.3.20C, Fig.3.4B).
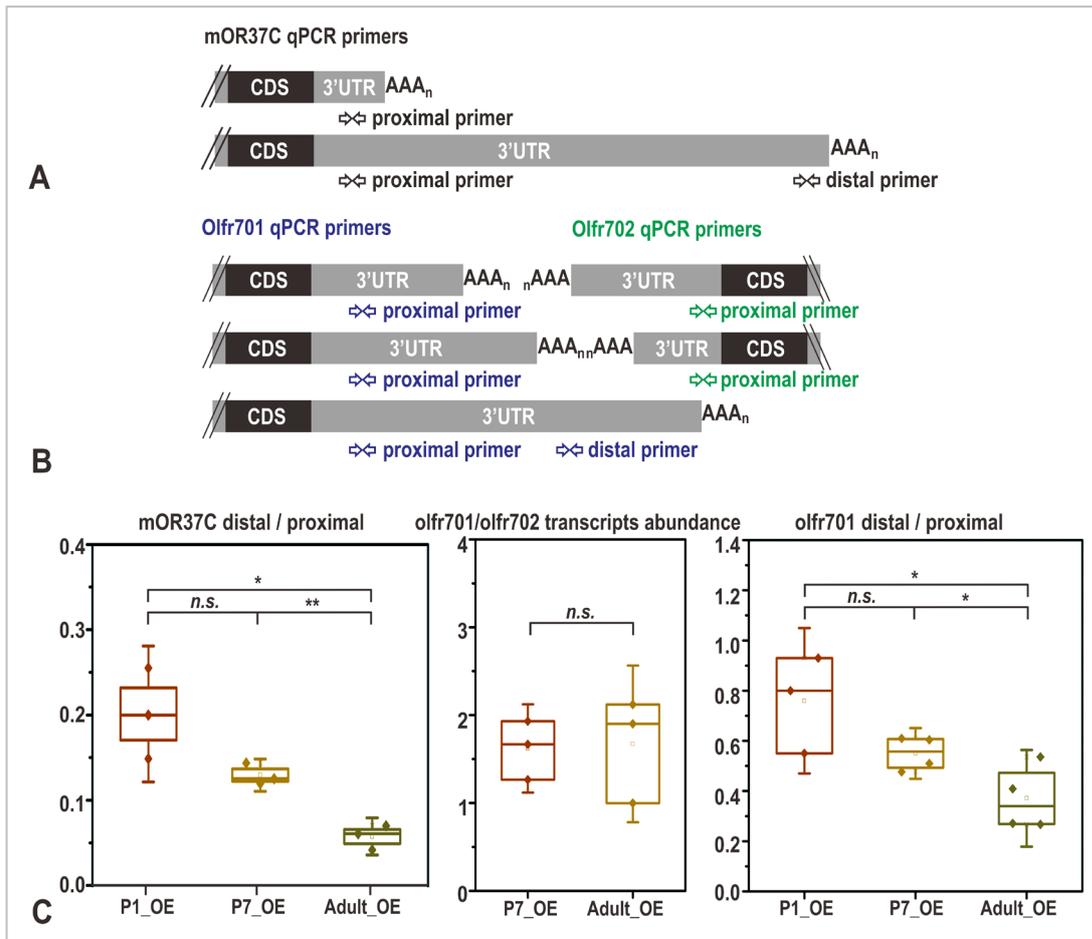
Figure 3.20. Expression levels of mOR37C and olfr701 transcript isoforms in the OE measured by qPCR (SYBR Green).

qPCR was performed with OE cDNA from different developmental stages. Random primed, non-purified cDNA was used as template, with 30ng input per reaction.

A. Schematic illustration of mOR37C qPCR primers. Two primer pairs were chosen based on the 3' ends identified by the 3' RACE, with the proximal primer detecting both isoforms and the distal primer only the long isoform.

Proximal primers: 37C_3'UTR_1233F; 37C_3'UTR_1392R. 159bp

Distal primers: 37C_3'UTR_6429F; 37C_3'UTR_6592R. 164bp

B. Schematic illustration of olfr701/olfr702 qPCR primers. For olfr701, two primer pairs were chosen based on the 3' ends identified by the 3' RACE, with the proximal primer detecting all three isoforms and the distal primer only the longest isoform. To assess the relative abundance of olfr701 and olfr702 transcripts, one primer pair for olfr702 was selected in its proximal region.

Olfr701 proximal primers: olfr701_prox_qPCR_F; olfr701_prox_qPCR_R. 190bp

Olfr701 distal primers: olfr701_dist_qPCR_F; olfr701_dist_qPCR_R. 144bp

Olfr702 proximal primers: olfr702_qprox _1071F; olfr702_qprox _1244R. 173bp

C. Quantification of mOR37C and olfr701 distal/proximal ratio during development. The relative expression level of the mOR37C long isoform is downregulated between juvenile (P1/P7) and adult stages. The relative abundance of olfr701 and olfr702 transcripts is maintained at constant level between P7 and adult stages, with olfr701 displaying a slightly higher expression level than its counterpart (olfr701/olfr702=1.7). Similar to mOR37C, the expression level of olfr701 long isoform is significantly downregulated between juvenile (P1/P7) and adult stage. Each dot corresponds to one animal and the whiskers represent standard deviation. T-test with n.s.: $\alpha \geq 0.05$; *: $\alpha < 0.05$, **: $\alpha < 0.01$.

mOR37C, P1_OE: N=3, distal/proximal ratio 0.200.

mOR37C, P7_OE: N=3, distal/proximal ratio 0.125.

mOR37C, Adult_OE: N=3, distal/proximal ratio 0.0607.

olfr701/olfr702, P7_OE: N=3, olfr701/olfr702 ratio 1.67.

olfr701/olfr702, Adult_OE: N=3, olfr701/olfr702 ratio 1.90.

olfr701, P1_OE: N=3, distal/proximal ratio 0.800.

olfr701, P7_OE: N=4, distal/proximal ratio 0.557.

olfr701, Adult_OE: N=4, distal/proximal ratio 0.340.

Taken all together, the results presented in this work suggest that alternative polyadenylation, which gives rise to multiple mRNA isoforms, might be a general modification of the odorant receptor transcripts. There are so far two types of developmental regulation observed. The first one, represented by mOR37A, is a maturation-dependent regulation of switching off the expression of mOR37A long isoform in the mature neurons, which could, however, not be validated in the other receptors. The second one, also seen in the other receptors, is an age-dependent regulation by downregulating the relative expression level of the long isoform during development. Furthermore, in the case of mOR37A, different isoforms might take on different functional roles during the axon guidance phase by interacting with other molecule(s). So far what is known about these interaction partners are their RNA nature and their potential interaction site within the mOR37A proximal amplicon of the short isoform. Interestingly, the observed RNA-RNA interaction was only present in the juvenile but not the adult stage. The exact identity of these interaction partners remains to be revealed, which will potentially benefit the understanding of the olfactory axon guidance mechanism.

4. Discussion

This work aimed to characterize the exact structure of mOR37A transcripts and to investigate any potential functional roles of different isoforms during the phase of olfactory axon guidance. Applying 3' RACE combined with Southern blotting, two major transcript isoforms, likely generated by alternative polyadenylation, were identified. The short 3'UTR ends 1438nt after the mOR37A coding region and the long 3'UTR 4727nt. The epithelial localization of the long isoform expressing cells, analyzed by RNA-FISH, demonstrated a predominant expression by the immature OSNs, which corroborated our previous findings of a maturation-dependent regulation of the long isoform. In accordance with these observations, the relative expression level of the long isoform quantified by ddPCR was seen downregulated during development. Furthermore, highlighted by Airyscan imaging, potential "transcriptional hotspots" of the receptor mRNA were revealed in the nucleus next to the heterochromatin blocks due to their robust and dominant expression in the OSNs. Importantly, the subcellular localization of the long isoform was observed in both the nucleus and cytoplasm, rendering it likely for the long isoform to be involved in cellular processes instead of pure regulatory functions in the nucleus. The link between alternative splicing and alternative polyadenylation was examined in the context of 5' intron splicing and 3' polyA site choice using mOR37A transcript as an example. However, no direct correlation was discovered. Finally, in the axon terminal, the short isoform was seen regulated in a developmental stage-dependent manner by an unknown RNA interaction partner. Through literature search and bioinformatic analysis, there was a perfect seed match in the short isoform with two olfactory-enriched miRNAs from the miR-200 family. In the other receptors, mOR37C, olfr701 and olfr702, alternative polyadenylation was likely to shape the transcript structure by generating mRNA isoforms with distinct 3'UTRs. Despite no maturation-dependent regulation, a similar age-dependent downregulation of the relative expression level of the long isoform was confirmed in receptors mOR37C and olfr701.

4.1 Alternative transcript structures of mouse odorant receptors

4.1.1 Alternative polyadenylation of mouse odorant receptor mRNA

The mOR37A-IRES-tauEGFP strain by enabling us to identify one species of OSNs, has greatly facilitated our previous single-cell cDNA library screen, in which the differential regulation of the receptor mOR37A mRNA between outgrowing and mature mOR37A-expressing OSNs was observed for the first time (Haag, 2009). In this transgenic strain, an IRES-tauEGFP sequence is inserted between the mOR37A coding region and its 3'UTR (Strotmann et al., 2000). Through PCR amplification, the insertion site downstream of the EFGP sequence was clarified here, with an inclusion of 67 extra nucleotides (Fig.3.1), which are the remnants of the excision of a floxed

neomycin cassette. The presence of two canonical polyA signals *"ATTAAA"* in the insertion sequence were recognized by web tools DNAFSMiner (Liu et al., 2005) and Dragon PolyA Spotter (Kalkatawi et al., 2013) as potential polyadenylation sites, whose influence on the transcript structure was illustrated later by comparing the transcript isoforms in the WT C57BL/6 and transgenic strains.

Transcript isoforms of the receptor mOR37A generated by 3' alternative polyadenylation were investigated in the transgenic strain and compared to the WT strain. By combining 3' RACE and Southern blotting, three possible ends were identified (Fig.3.2D). Among them, two match to the 3' ends in the WT C57BL/6 strain (Helisch, 2014) and one is novel. The novel end possesses the shortest 3'UTR of 584nt and a non-canonical polyadenylation signal *"AATAAT"* (Fig.3.2D). As Southern blotting greatly enhances the detection sensitivity and specificity, this end could have been overlooked in the WT C57BL/6 strain, where no blotting was done. Moreover, the mapped region is neither rich in templated genomic As nor does it show a high degree of complementarity to the 3' RACE CDS primer, which further argues against potential mispriming problems from the PCR. Thus, the authenticity of this end is technically validated. However, the usage of the *"AATAAT"* signal in the mouse has a very low frequency compared to other hexamers (Gruber et al., 2016), and in this case it is also not positioned 15-30nt upstream of the cleavage site, where the polyadenylation signal is typically located. One more piece of evidence that speaks against the existence of this isoform is that the location revealed by sequencing (584bp after CDS) is much more 5' upstream than what the Southern blot suggests (ca. 850bp after CDS). In sum, this novel transcript isoform might only contribute to a very small portion of the entire mOR37A transcript repertoire, if it exists at all. Contrary to this, the other two ends fit exactly to what has been shown in the WT C57BL/6 strain, with the short 3'UTR being 1438nt (PAS: *"ATTAAA"*) and the long 3'UTR being 4727nt (PAS: *"AATAAA"*) (Fig.3.2D). Furthermore, bioinformatic PAS predictions with various web tools (DNAFSMiner, Dragon PolyA Spotter) predict their presence but largely miss the novel one. Only with DNAFSMiner using the manually defined PAS motif (*"AATAAT"*), the novel end can be recognized. With regard to the polyA signals in the insertion site caused by the excision of a loxP site, no extra transcript isoforms were identified utilizing these signals.

In sum, the genetic manipulation of inserting IRES-EGFP does not alter the 3' end alternative polyadenylation process and there are two major transcript isoforms in the transgenic strain like in the WT C57BL/6 strain.

One well-established method to visualize RNA is the Northern blot. A sensitivity test was carried out with synthetic target RNA and probe (both 940nt) in a dilution series to determine the detection threshold. The lowest target RNA input lies at around 30fg (data not shown), which corresponds to 5.99×10^4 molecules. As the long isoform is the underrepresented fraction of the entire mOR37A transcript repertoire, the detection threshold should be adjusted based on its

abundance. When taking into account the quantitative information obtained from ddPCR, 4.1 copies of the long isoform (data from P7 OE without any treatment) are present in 1ng of random primed OE cDNA. Assuming the conversion efficiency from RNA to cDNA being 100%, 5.99×10^4 molecules should be present in 14.6µg of P7 OE total RNA. One trial with 15µg of OE total RNA and a probe of similar length (1327nt) in the proximal region of mOR37A mRNA was performed to capture all transcript isoforms, but it only cross hybridized to the 18S and 28S rRNAs according to the band sizes (1.9kb and 4.7kb), which were distinguishable from the expected length of the mOR37A isoforms (short 2.4kb and long 5.7kb) (data not shown). Although 15µg of total RNA is in a reasonable range, rRNA is very abudant in the total RNA pool, which might be the reason for such cross-hybridizations. One way to improve the specificity is to enrich mRNA from total RNA.

To extend the observations from mOR37A, alternative polyadenylation was investigated in other receptors as well. Interestingly, all other receptors, mOR37C, olfr701 and olfr702 studied so far, either closely related to mOR37A (mOR37C, 88% CDS animo acid identity) or with high sequence divergence (olfr701, 45% CDS animo acid identity), demonstrate multiple transcript isoforms generated mostly likely by alternative polyadenylation (Fig.3.16, Fig.3.17). Although the RACE-PCR here was done without Southern blot hybridization, which might raise sensitivity and specificity issues, all of the identified 3' ends were carefully examined so that they do not map to genomic A rich regions and do contain a recognizable polyA signal at the appropriate position upstream of the cleavage site (Fig.3.16, Fig.3.17). Thus, even if there were isoforms which might have been overlooked, the identified isoforms already reflect the active post-transcriptional modifications on the receptor mRNAs. Finally, the alternative polyadenylation of receptor mRNAs shall be proven by methods that allow the direct visualization of the full-length isoforms such as Northern blot or 3' RACE combined with Southern blot. To this point, data gathered in this thesis imply that alternative polyadenylation could be a general phenomenon in modifying mouse odorant receptor transcripts.

Indeed, it has been reported by other publications that OR transcripts undergo extensive post-transcriptional modifications in their UTRs, such as alternative splicing and alternative polyadenylation (Young et al., 2003; Ibarra-Soria et al., 2014; Doulazmi et al., 2019). For instance, Young et al. reported that more than half of the receptors from their cDNA collection (419 OR genes) utilize more than one polyadenylation site (Young et al., 2003), and Doulazmi and colleagues concluded that more than 77% of odorant receptors are subject to alternative polyadenylation (Doulazmi et al., 2019).

A direct comparison of the 3'UTR length of the four studied receptors with publicly available datasets is illustrated in Fig.4.1. Alternative polyadenylation of all four receptors were explored by 3' RACE in this thesis. The annotated 3'UTRs refer to the annotations in the Ensembl database (Release 97) and the RNAseq data are from Ibarra-Soria et al., 2014. The multiple 3'

ends of receptor mRNA are principally confirmed by the Ensembl annotations. Moreover, the longest isoform of each receptor by 3' RACE is also roughly detected by RNAseq. Notably, the RNAseq method used in the respective publication is not 3'-Seq but whole transcriptomic analysis, which might account for the absence of other 3'UTR isoforms (Ibarra-Soria et al., 2014). Altogether, alternative polyadenylation is likely to be a common 3' modification on mouse odrant receptor mRNAs and multiple isoforms are generated with varying 3'UTRs.
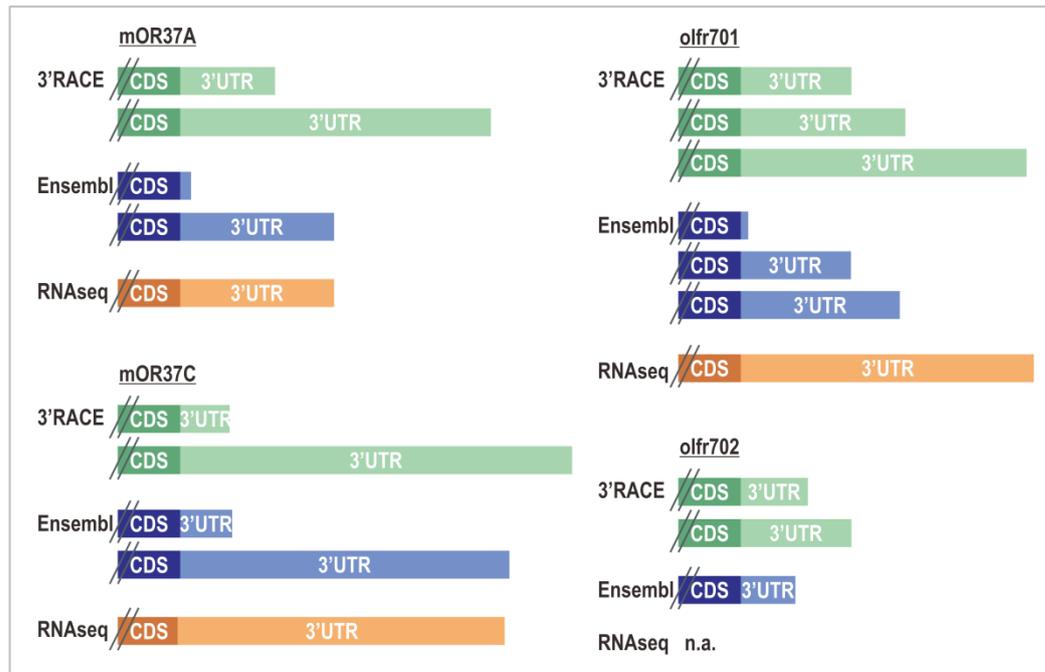

Figure 4.1. Comparison of 3'UTR lengths of the four studied receptors with publicly available datasets. The 3'UTRs of all four receptors investigated by 3' RACE, from the Ensembl database annotations (Release 97) and from RNAseq data (Ibarra-Soria et al., 2014) are listed in green, blue and orange.

4.1.2 Nuclear and cytosolic localization of the receptor transcript isoforms

The visualization of the subcellular localization of different isoforms was enabled by Airyscan imaging. In line with previous publications (Clowney et al., 2012; Armelin-Correa et al., 2014), it was noticed that OSNs possess unique heterochromatin organizations in their nuclei, by harboring one or a few large centrally localized heterochromatin clusters surrounded by smaller heterochromatin islands (Fig.3.3D). This nuclear architecture is essential for the singular receptor choice (Clowney et al., 2012). When comparing the subcellular localization of mOR37A transcripts labeled by the proximal and distal probes, a nuclear and cytosolic localization can be seen with both (Fig.3.3B'.C'). Interestingly, the nuclear spot is located exactly neighbouring the heterochromatin clusters, representing potential "transcriptional hotspots". This observation also fits to previous descriptions by the others that during the receptor choice, the active OR allele loses its heterochromatin signatures and is moved to euchromatin regions (Magklara et al., 2011). Notably, such nuclear "transcriptional hotspots" were confirmed in the receptor mOR37C as well (Fig.3.18C'.C''), but they were not detected for every transcript, as the OMP mRNA in the

mature OSNs labeled by RNA-FISH was more or less homogeneously distributed only in the cytoplasm (data not shown). The hotspot presumably indicates the very high expression of a given gene. Of higher relevance in this context is, however, the cytosolic fraction, which strengthens the involvement of the long isoform in cellular processes instead of pure regulatory functions in the nucleus.

4.1.3 Developmental downregulation of the long isoform

Of the two alternatively polyadenylated mOR37A isoforms, a maturation-dependent downregulation of the long isoform has long been indicated from our previous work, by the presence of the 3'-extended transcripts exclusively in the immature mOR37A neurons in the initial single-cell cDNA library screen (Haag, 2009), and the switching off of the long isoform in the mature (GAP43⁻,OMP⁺) mOR37A neurons by RNA-FISH (Falk, 2015).

The present thesis corroborates this interesting finding by providing yet other independent evidence. To start with, the 3' RACE-Southern blot hybridization pattern between juvenile and adult stages is highly similar when hybridized with probe 1 and 2 (proximal probes), whereas probe 3 (distal probe) only detects RACE-PCR products in the juvenile stage (Fig.3.2C). The complete absence of the long isoform in the adult stage is in line with previous observations. In other words, the 3'-extension of the mOR37A transcripts is mainly enriched in the juvenile stage, where the immature neurons are highly abundant. Apart from this, RNA-FISH with the quantification of "relative height" in the OE also clearly demonstrates a more basal localization of the long isoform-expressing OSNs in comparison to the entire mOR37A-expressing population (Fig.3.3E). Finally, quantitative ddPCR reveals a reduction of the relative expression level of the long isoform during development from P1/P7 to adult stages (Fig.3.4B). It has to be pointed out that the downregulation of the distal/proximal ratio at tissue level might have several reasons. On the one hand, there is a reduction of the immature OSNs in the adult stage. When combined with the switching off behavior of the long isoform upon maturation, its expression level thus stays further underrepresented in the adult stage. On the other hand, there could also be an age-dependent regulation of different isoforms. When combining the cell numbers analyzed by RNA-FISH and the transcript level measured by ddPCR, from P7 to adult stages the percentage of the immature cells decreases from 42% to 2% (Falk, 2015) but the long transcript level drops only from 0.186 to 0.079 (Fig.3.4B). A sharp reduction of the immature cell numbers together with a mild decrease of the long transcripts indicate that either each immature neuron at the adult stage habors more long isoforms in absolute quantity than at the P7 stage or each mature neuron at the adult stage strongly decreases its expression of the short isoform or both combined. In the two extreme situations, where the expression of the short/long isoform is expected to be significantly down/upregulated in their corresponding population, a qualitative difference in

comparison to the P7 stage should be visible in RNA-FISH. Furthermore, in the P7 animal, the distal/proximal ratio in the immature population was calculated to be approximately 0.5 according to the formula (section 3.1.5) by measuring the proximal labeling intensity in both the mature and imature populations from RNA-FISH (Fig.3.5), which differs significantly from the 0.186 distal/proximal ratio in the entire epithelium (Fig.3.4B). Similarly, the distal/proximal ratio in the immature population can also be measured in the adult animal. With that piece of information, a better understanding of the cause of the age-dependent regulation can be acquired. Thus, the developmental regulation of mOR37A isoforms might come in two flavors, namely a maturation-dependent on/off switch and an age-dependent regulation.

Contrasting this, receptors mOR37C and olfr701 do not show a qualitative maturation-dependent "on/off" switch as RNA-FISH demonstrates a persistent co-localization of the proximal and distal probe labeled cells, regardless of the maturation state (Fig.3.18B.C). Moreover, the relative fluorescent intensity of the distal/proximal labeling of the receptor olfr701 measured from RNA-FISH also does not show any significant difference between the immature and mature compartments in the OE, potentially ruling out a quantitative maturation-dependent regulation (Fig.3.19B). Given that three transcript isoforms were discovered in the receptor olfr701 (Fig.3.17B), one reason for the absence of a quantitative maturation-dependent regulation might be that not every isoform is regulated and a "wrong" isoform was checked in this case. But multiple isoforms do co-exist based on the relative transcript levels measured by qPCR (Fig.3.20C) and the distal/proximal intensity ratio quantified from RNA-FISH (Fig.3.19B). Thus, the observed co-localization at least is not due to the presence of only the long isoforms which contain apparently both targets, but from different isoforms. Despite no downregulation of the long isoform upon OSN maturation, an age-dependent regulation can be detected in the other receptors by qPCR (Fig.3.20C).

Notably, the distal/proximal ratio by qPCR seems to vary between different receptors, for instance, the ratio of mOR37A and mOR37C is comparable, whereas that of olfr701 is much higher than the other two (Fig.3.4B, Fig.3.20C). A possible reason that might account for such differences is that olfr701 and its neighboring olfr702 have overlapping 3'UTRs (Fig.3.17), which could lead to co-amplification of the other receptor in any PCR-based methods. When comparing the absolute copy numbers of different receptors quantified by qPCR and ddPCR, the expression level of the olfr701 long isoform is seen much higher than that of the other receptors (Tab.4.1). Despite a careful choice of the olfr701 qPCR distal primers to flank the distal end of olfr702 (Fig.3.20B), there might still be longer unidentified olfr702 isoforms, which could eventually be co-amplified and thus contribute to a high distal/proximal ratio.

Table 4.1. Mean value of the apparent absolute copy number of the proximal and distal targets in the P7 OE, normalized to 1ng cDNA input.

	proximal target	distal target
mOR37A (ddPCR)	26.33	4.09
mOR37C (qPCR)	47.49	6.25
Olfr701 (qPCR)	40.47	19.75

The potential functional relevance of the maturation-dependent regulation of the receptor mOR37A will be viewed in two contexts, namely the receptor choice and the axon guidance. The major argument for that lies in the temporal coincidence of these events, as OSN maturation coincides with the final stabilization of the receptor choice as well as the formation of the olfactory circuits.

Olfactory sensory neurons conform to the "one neuron-one receptor" rule, which specifies the choice of one singular receptor per neuron from a repertoire of more than 1000 receptor genes. The molecular mechanism of this monogenic and monoallelic expression depends largely on the chromatin state and the nuclear organization of OSNs. It has been demonstrated that OR genes are silenced epigenetically and condensed into a few foci close to the heterochromatin clusters prior to functional receptor choice (Magklara et al., 2011; Clowney et al., 2012). A stochastic choice is made, when a single OR gene gets addressed by multiple Greek island enhancers and escapes the foci to get positioned in the nearby heterochromatin-free regions (Clowney et al., 2012). Could the long isoform assist the escape of a receptor gene from its suppressive state? The subcellular localization of the long isoform was indeed observed between the heterochromatin clusters (Fig.3.3C'). It should, however, be viewed as a transcriptional active site instead of the event that initiates transcription. The receptor choice occurs in the DNA level, whereas the post-transcriptional modification of alternative polyadenylation can only take place when the transcription has already been initiated. Therefore, from a temporal point of view, the assumption of "the long isoform facilitates receptor choice" is not valid, given that the receptor choice happens ahead of it. Moreover, if the function of the long isoform is restricted in aiding and stabilizing the singular receptor choice, it should be expected to be highly enriched in the nucleus. Experimental evidence against this assumption can be seen from the cytosolic localization of the long isoform captured by RNA-FISH (Fig.3.3C'), and the presence of the long isoform in the distal compartment of axon termini by both RT-PCR (data not shown) and quantitative ddPCR (Tab.3.2). Therefore, it becomes less likely that the long isoform is directly associated with receptor choice.

What about the second possibility of a role in the axon guidance? Olfactory map formation is known to take place already in the prenatal phase. Soon after OSN differentiation, they extend axons navigating towards the telencephalon where the OB will later develop. At E15, OSN axons accumulating on the surface nerve layer finally penetrate deeper into the OB and the glomerulogenesis starts, establishing a coarse olfactory map (Treloar et al., 1999). As the

olfactory stem cells in the OE constantly generate OSNs, olfactory axon guidance becomes a lifelong requirement for those newly-generated OSNs to find their targets and to integrate into existing circuits. In the P7 mice, the thickness of GAP43 (immature neuron marker) labeled basal epithelium is comparable to the OMP (mature neuron marker) labeled apical epithelium (data not shown), indicating the presence of large number of immature OSNs with growth cones and thus active axonal pathfinding. Contrary to this, in the adult mice, the GAP43 stripe becomes much thinner in the entire epithelium, since most of the OSNs already form synaptic connections in the OB. Between the P7 and adult stages where the developmental downregulation of the long isoform was seen, such downregulation can be linked to a decreased axonal pathfinding activity due to the reduction of the immature OSNs in the epithelium. Therefore, the physiological time course of axon guidance and the developmental regulation of the long isoform at least coincide in a temporal manner. Notably, there is no quantitative coincidence of the immature cell number and the long isoform expression, as from P7 to adult stages the ratio of the immature cells decreases from 42% to 2% (Falk, 2015) but the long transcript level drops from 0.186 to 0.079 (Fig.3.4B). This implies that the long isoform remains overrepresented in the immature neurons at the adult stage, corroborating its role in the lifelong regeneration of OSNs. Moreover, literature suggests that the receptor mRNAs are transported to the distal compartment of the OSNs (Ressler et al., 1994; Vassar et al., 1994) and RT-PCR data from this work also confirms the presence of the long isoform in the axon termini (data not shown). For this reason, I later switched the experimental focus to the OB and tried to identify any guidance-related functions of different isoforms.

4.1.4 No co-regulation of alternative polyadenylation and alternative splicing

Both mOR37A isoforms possess the canonical polyA signals of proximal *"ATTAAA"* and distal *"AATAAA"* (Fig.3.2D). Among them, *"ATTAAA"* accounts for 12.3% of the mouse 3' sequences and *"AATAAA"* 39.5%, according to a recent publication that comprehensively analyzed 3' end sequencing data from public databases (Gruber et al., 2016). Such an arrangement of a stronger polyA signal at the distal end has been reported in HEK293 cells, when the distal site is chosen from multiple cleavage sites (Martin et al., 2012). As a general rule of thumb, in proliferating cells the proximal polyA site tends to be utilized and in differentiating cells the distal polyA site is preferred (Sandberg et al., 2008; Ji and Tian, 2009; Elkon et al., 2012). Here upon OSN differentiation the proximal polyA site is chosen, which seems to be an exception of the pre-existing examples, but the polyA site preference in various biological processes mainly depends on the cellular context which eventually influences the expression levels or the interactions of the 3' end processing factors via different regulatory mechanisms. It remains largely unknown, what is the molecular mechanism behind polyA choice in odorant receptors. The relation between

alternative splicing and alternative polyadenylation in mOR37A transcripts was examined here due to the potential crosstalk between these two machineries by common regulators (Wang et al., 2008).

Demonstrated already in the seminal work from the early 1990s (Buck and Axel, 1991), OR genes are devoid of introns in their coding regions, but they do contain 5' non-coding exons and introns (Glusman et al., 1996; Sosinsky et al., 2000; Ibarra-Soria et al., 2014). In the Ensembl database (Release 97), an intron of 2551bp is annotated 5' upstream of the mOR37A coding region. The possibility of 5' splicing was addressed here with RT-PCR in the WT C57BL/6 mouse given the difficulty in amplifying the IRES sequence in the transgenic strain. With a forward primer in the annotated intron region and two reverse primers before or after the 1st polyA site, products with the correct size were amplified, indicating the presence of the intron in at least some transcripts (Fig.3.6A). The intron retention was confirmed by 5' RACE (Fig.3.6C). Importantly, it is not possible at this stage to distinguish mature transcripts from unprocessed nuclear pre-mRNAs because TRIzol® (phenol/chloroform) based total RNA isolation can isolate both and lead to intronic sequencing reads from nascent transcripts (Sultan et al., 2014). Alternative splicing and alternative polyadenylation were then investigated by RNA-FISH, so that the cytosolic and the nuclear transcripts can be separately visualized.

Though it is unclear whether the intron retention also occurs in the transgenic strain, the identical transcript processing demonstrated for the 3' end modification by alternative polyadenylation in both strains makes it likely. Therefore, RNA-FISH was performed with an intronic RNA antisense probe on P7 sections of the transgenic strain (Fig.3.7A). Notably, a substantial amount of cells were in fact labeled by the intron probe (Fig.3.7B). By assessing the "relative height" of each labeled cell in the entire olfactory epithelium, a strikingly similar localization pattern between the intron-expressing cells and the entire mOR37A-expressing population (proximal probe labeled cells) was seen. In other words the intron-expressing cells are distributed across the entire height of the epithelium with a major peak in the mature neurons (0.6-0.8) (Fig.3.3E, Fig.3.7C). This is clearly distinguishable from the distal probe labeled cells, which show an accumulation in a more basal position (Fig.3.3E). Thus, an exclusive association between the 5' intron and the long isoform can be ruled out.

It should be noted here that the relative expression level of the intron-containing transcripts is comparable to that of the long isoform, and there is even a slight tendency of a developmental regulation (Fig.3.8B), mimicking the ddPCR quantification of the long isoform (Fig.3.4B). However, the developmental regulation of the long isoform between P7 and adult stages cannot be replicated with the intron-containing transcripts, which again excludes any systematic link between the long isoform and the intron (Fig.3.4B, Fig.3.8B). Moreover, given that the localization of the intron-expressing cells are also in the mature neruons (Fig.3.7B.C), there must be long isoforms that do not contain the 5' intron.

Can the intron then solely be present in the short isoform? Clear evidence from intron RT-PCR argues against it (Fig.3.6B). Moreover, quantitative measurements from ddPCR show that the intron-containing transcripts make up only a small portion of the entire mOR37A transcripts at all three tested developmental stages, making it mathematically impossible for every short isoform to harbor an intron (Fig.3.8B).

Furthermore, the subcellular details of the intron-containing transcripts also contrast that of the short and long isoform by a relatively concentrated nuclear and low cytoplasmic localization (Fig.3.3B'.C', Fig.3.7D.D'). Such nuclear distribution has been noticed in other receptors, such as MOR28 and M50, when hybridized with intronic probes in RNA-FISH (Clowney et al., 2012). Taken together, the observed "intron retention" by RT-PCR and RNA-FISH could be unprocessed nuclear nascent RNAs, and they are mainly spliced out or rarely retained in the mature transcripts. Combining these results, the 5' intron is believed to be assigned randomly to the short and long isoforms and shall largely be removed from the mature transcripts.

To achieve the numerical equilibrium of a comparable localization pattern to that of the entire mOR37A-expressing population and a comparable expression level to the long isoform, one could imagine, for instance, at the P7 stage, among every five mOR37A transcripts, four of the short isoform and one of the long isoform (distal/proximal=0.186, Fig.3.4B), and one intron could be assigned randomly to any of these transcripts (intron/proximal=0.185, Fig.3.8B). This fulfills the quantitative measurements from ddPCR at tissue level. At the single cell level, since the distribution of the intron is random, its relative position could therefore mimic the entire mOR37A-expressing population. In sum, alternative splicing and alternative polyadenylation seems to be independent from each other in processing the mOR37A transcripts.

It remains unknown how the alternative polyadenylation machinery is regulated. As key regulator proteins for the 3' end processing, the expression levels of, for instance, the CtsF or the CPSF protein complex could be studied and compared at different stages to further shed light on this issue.

4.2 Potential functional roles of different isoforms during axon guidance

4.2.1 Interaction partner in the axonal termini

Odorant receptor mRNA is known to be translocated to the axon termini (Ressler et al., 1994; Vassar et al., 1994) and the odorant receptor protein is established as a key instructive determinant of the olfactory mapping (Feinstein and Mombaerts, 2004). Through high throughput RNAseq (Ibarra-Soria et al., 2014) and cDNA library screening (Young et al., 2003), complex transcriptional isoforms of ORs have first been noticed, however no evidence about their biological functions has ever been in the limelight. Given the observed specific expression of the long isoform during the axon guidance phase and its downregulation thereafter, I set to investigate possible functions of receptor mRNA isoforms in the process of axon guidance using receptor mOR37A as an example. This has hinted at the existence of specific interaction partners of the short isoforms in the axon termini.

The first experimental evidence comes from ddPCR with P7 OB cDNA, where presence or absence of RNase treatment prior to ddPCR leads to significantly different distal/proximal ratios (Fig.3.9). To confirm the specificity of this effect, diverse control experiments were carried out. However, whether it is Proteinase K treatment with different purification methods (Fig.3.11) or with cDNA template from other sources (Fig.3.10), the sensitivity of distal/proximal ratio towards RNase treatment observed on P7 OB RNA cannot be replicated. It is thus assumed that the P7 OB cDNA exhibits a peculiar "RNase sensitivity" and this sensitivity occurs only under highly specified conditions (P7 not adult, OB not OE) when treated with RNase A and H. This RNase sensitivity could hint at an interaction partner of RNA nature, which happens to bind to the amplicons during PCR amplification and which can be removed by RNases.

When talking about the ratio change, one has to assume a situation where the two targets are differently affected. In order to figure out which target reacts towards the RNase treatment, the ratio was dismantled into the absolute copy numbers of the proximal and distal targets. This revealed that it is the proximal target which reacts to the RNase treatment (Fig.3.12A). Notably in the P7 OE, both targets remain at similar levels and the ratio remains unaltered upon RNase treatment, which renders the increase of P7 OB proximal target upon RNase treatment so noteworthy (Fig.3.12A). Although the absolute value of the target copy numbers normalized to 1ng cDNA input might not be accurate due to the difficulty in assessing cDNA concentrations, especially as with or without RNase treatment, the cDNA is present either as single-stranded cDNA or cDNA-RNA hybrid, the true indicator is the numerical change affecting the proximal target. It is believed that the proximal target number with RNase treatment reflects the actual copy numbers, and the readout is lower without RNase treatment due to the interaction between the proximal amplicon and some unknown RNA molecule. Thus, the interaction site must be in the proximal amplicon. In support of this interpretation, RNA secondary structures of the short

and long 3'UTRs were predicted by the RNAfold webserver from ViennaRNA based on a computational algorithm minimizing the free energy (Gruber et al., 2008). According to this prediction, both 3'UTRs show complicated secondary structures, but the region of the proximal amplicon has one of the lowest intramolecular base-pairing probabilities among the entire RNA (color code, Fig.4.2). This might be essential for the intermolecular interactions because the target region should be unstructured or at most engaged only in weak self-folding structures.
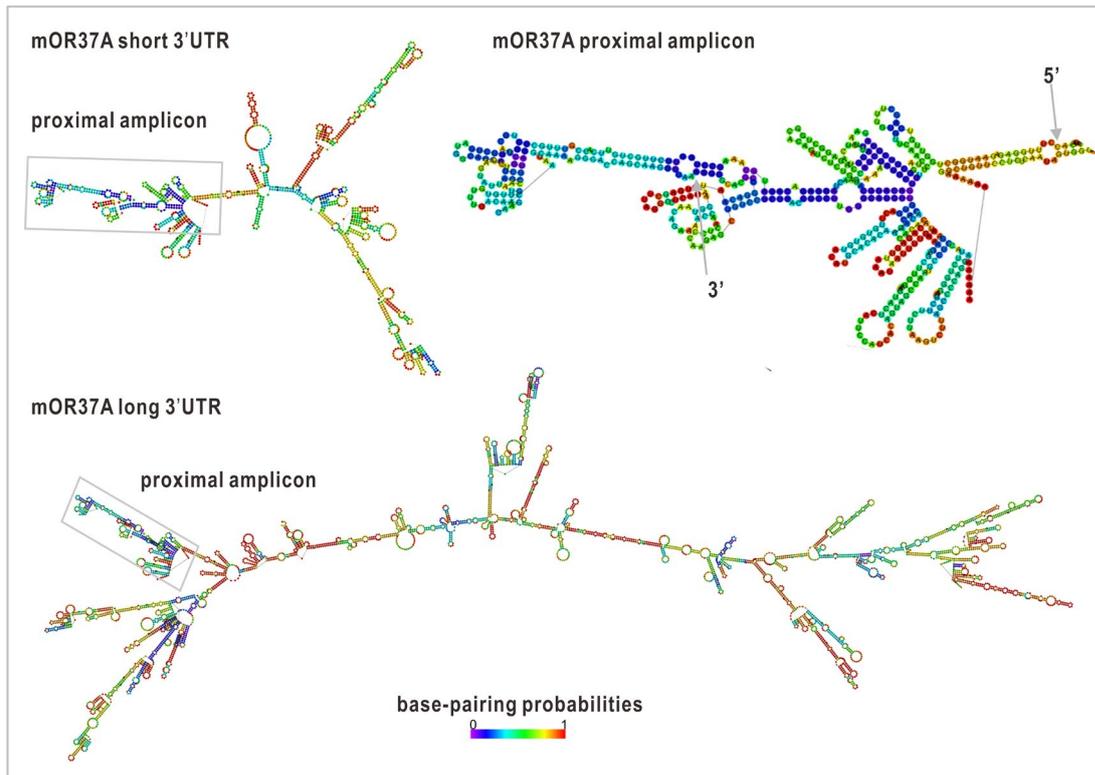

Figure 4.2. ViennaRNA's prediction of RNA secondary structures with bases color-coded for their pairing probabilities. Both mOR37A short and long 3'UTRs have complex secondary structures, but the region of the proximal amplicon stays identical (marked with the grey rectangles and enlarged on the right). In the enlarged view, the base-pairing probabilities in the proximal amplicon are seen among the lowest in the whole 3'UTR. 5' and 3' indicate the 5' and 3' end of the proximal amplicon.

To explain the observed phenomena, a working model was put forward (Fig.3.13). It is assumed that there is some unknown RNA interaction partner, which inhibits the proximal target amplification during PCR. Only when digested with RNases, this inhibitory factor can be removed and efficient PCR amplification can be restored. Though it might seem unlikely at first glance that PCR reactions can be inhibited by an RNA molecule, oligoribonucleotide interference-PCR based on the principle of RNA-DNA hybridization has been reported before (Yuen et al., 2001; Tanigawa et al., 2014; Fujita et al., 2018). In brief, a complementary RNA molecule can bind to one of the template strands and prevent the elongation of the DNA polymerase. The exact length of such complementary RNA varies in different publications. For instance, Fujita et al. suggested that 17-25 bases of oligoribonucleotide can inhibit PCR (Fujita et al., 2018), whereas Yuen et al. tested a 750nt long RNA, which also successfully blocked the

PCR amplification (Yuen et al., 2001). Such kind of PCR inhibition even has its own applications, for example when using degenerate PCR primers to identify new members of a multigene family, the known family members can be suppressed by using complementary RNAs so that the new ones can be preferentially amplified (Yuen et al., 2001). From a mechanistic point of view, there are a few aspects that should be mentioned. First of all, the DNA polymerase in the ddPCR Supermix is the Taq polymerase. It possesses the 5'-3' exonuclease activity on DNA substrates (Holland et al., 1991), which is essential to cleave the TaqMan probe into fluorescent fragments. Taq polymerase also exhibits some RNA template-dependent 5' nuclease activity according to Ma and colleagues, but it is significantly lower compared to the DNA-based nuclease activity (Ma et al., 2000). Moreover, it is not clear where the complementarity is formed between the mOR37A proximal amplicon and the unknown RNA molecule, at 3' end or 5' end, and whether the interacting RNA molecule is circular or linear. All these critical points might eventually affect whether the inhibitory RNA molecule can be digested by the limited nuclease activity of the Taq polymerase. At least as the experimental data suggest, the potential RNA molecule discussed here does persist and exerts an impact on the PCR amplification. Another aspect that might favor RNA-DNA hybridization is the ddPCR procedure itself. In order to stabilize the droplets by ensuring a uniform thermal transfer, the ramp rate of the whole reaction was set to 2°C/sec. A slow pace of heating and cooling is beneficial for RNA annealing. Remarkably, the inhibitory RNA molecule must be present in large quantities. As the cDNA template is randomly partitioned into 20.000 droplets, to achieve a global inhibition of the proximal target, there must be much more inhibitory molecules than its target.

Taken together, observations obtained from ddPCR suggest the presence of a proximal-specific RNA interaction partner of mOR37A transcripts in the axon termini at P7.

4.2.2 Interaction partner of the short isoform

As the interaction site appears to be in the proximal amplicon, a question is raised of whether the proximal amplicons of both isoforms are affected equally. When comparing the distal/proximal ratio in the OE and OB (Fig.3.9B), the ratio in the P7 OB goes down upon RNase treatment to a comparable level as seen in the P7 OE, indicating that both isoforms might be packed and translocated to the axon termini in the same ratio as they are in the cell soma. In fact, the loss of the proximal amplification was never complete, even when no RNase treatment was performed. More precisely, there is a similar amount of proximal target left as there is of the distal target without RNase treatment (Tab.3.2). This numerical coincidence promotes the idea that the interaction might only take place within the short isoform. Experimentally, one trial to exclude the short isoform from the scheme was performed by cDNA synthesis with a gene specific primer downstream of the distal target (Fig.3.12B). Notably, after singling out the long isoform, the

distal/proximal ratio loses its sensitivity towards the RNase treatment (Fig.3.12B). This experimentally supports the idea that the interaction partner might only bind to the short isoform. What could then be the reason that spares the long isoform from interaction despite the identical sequence? One possibility would be that the long isoform *in vivo* might recruit RNA modifying enzymes, which lead to sequence alterations of the proximal amplicon so that it can no longer bind the inhibitor.

Taken together, the loss of the proximal amplification numerically matches exactly the proximal targets from the short isoform, which might be the first evidence of differential interaction of different isoforms.

4.2.3 Developmental aspects of the interaction between mOR37A and its RNA partner

The existence of an interaction partner in the P7 OB exactly at the phase of axon outgrowth promotes the investigation of this phenomenon from a developmental perspective. Surprisingly, in the adult OB the distal/proximal ratio loses its sensitivity towards RNase treatment and is maintained at a low level close to that of the adult OE (Fig.3.15). This could be an indicator of the loss of the interaction partner in the adult stage. Thus, the potential interaction partner is present only in axons and only during the axon guidance phase.

As a side note, in the P7 mouse the short and the long isoforms are likely to be transported to the axon termini in the same ratio as they are in the cell soma, since the distal/proximal ratio upon RNase treatment drops to the OE level (Fig.3.9B). Observations from the adult OB corroborate this assumption by having comparable distal/proximal ratios in both compartments (Fig.3.15).

There are so far two forms of developmental regulations. The long isoform is switched off upon maturation in every OSN and the short isoform suppression in the axons is lost between P7 and adult stages. The maturation-dependent regulation of the long isoform coincides temporally with formation of the olfactory circuits and the developmental stage-dependent regulation of the axonal short isoform might correspond to the olfactory critical period. The ontogenetic and regenerative axon guidance mechanism seems to differ in the olfactory system, as a critical period of olfactory axon targeting has been discovered (Ma et al., 2014; Tsai and Barnea, 2014). Only within the critical period till the first postnatal week (P7), a disrupted glomerular map could be restored and recovered. Beyond this time window, the disruption is immutable. Thus, the observed loss of suppression on the axonal short isoform between P7 and adult stages coincides with the establishment and maintenance phases of the olfactory circuits.

Considering the release of the axonal short isoform from suppression in the adult stage, it should be noted that the odorant receptor proteins can be detected in the axon termini by immunohistochemistry from early postnatal stages to adulthood, implying the permanent presence of the OR protein in the axons (Barnea et al., 2004; Strotmann et al., 2004; Low and

Mombaerts, 2017). Thus, speculated from the two forms of regulation, the long isoform might take on the role of axon guidance in the immature neurons and be switched off after synaptogenesis in the mature neurons, while the short isoform might be needed for the maintenance of exisiting circuits in the post-critical period.

4.2.4 Bioinformatic search for the potnetial mOR37A interaction partner

In recent years, the regulatory functions of non-coding RNAs are slowly coming to light and are proven to be essential for diverse cellular processes (reviewed in Mattick and Makunin 2006; Morris and Mattick, 2014). With the help of ddPCR, evidence for an RNA interaction partner is emerging in the axon termini. First of all, to test the inhibitory effect of natural small RNA molecules from the P7 OB, one "gain-of-function" experiment was designed with isolated small RNA fractions (<200nt) added to the cDNA from *in vitro* synthesized RNA (Fig.3.14A). In this setup, the distal/proximal ratio increases drastically with increasing amount of small RNA molecules, fully mimicking the effect seen in P7 OB without RNase treatment (Fig.3.14B, Fig.3.9B). Though the concentration of the added small RNAs could not be determined due to the lack of proper setup, it highly narrows down the search for a potential interaction partner to the natural small RNA species in the OB (Fig.3.14).

Before experimental identification of the RNA interaction partner, literature was browsed for small non-coding RNAs, especially microRNAs, in the olfactory system. By a handful of publications (Choi et al., 2008; Bak et al., 2008; Sun et al., 2014; Beclin et al., 2016), several miRNAs, miR141, miR183, miR200b, miR429, are universally found in the mouse olfactory bulb with different techniques. When compared to other organisms such as the zebrafish where the miRNA expression during embryonic development has been studied in depth, the aforementioned miRNAs are confirmed to have a high olfactory association (Wienholds et al., 2005). Among them, miR-141, miR-200b and miR-429 all belong to the miR-200 family, which consists of five members in total and is among the best characterized miRNAs (reviewed in Senfter et al., 2016). In most cases, the miR-200 family is highly enriched in the epithelial cells and has prominent roles in metastatic cancers (reviewed in Humphries and Yang, 2015). There are also some detailed reports about its function in the olfactory system, such as regulating the terminal olfactory differentiation from progenitor cells to mature OSNs (Choi et al., 2007). Similar to this, the miR-183, which is in a cluster containing two other miRNAs, also has been shown to be essential for sensory neuron maturation (Fan et al., 2017).

Given their enrichment in the olfactory system, the seed regions of all identified miRNAs were aligned to the proximal amplicon where the interaction is predicted (Fig.4.3). Interestingly, the seed region of miR-200b and miR-429 has a perfect match with mOR37A mRNA in the proximal amplicon part (Fig.4.3). Apart from this, there is an "A" across in the mRNA sequence in

position 384, which would be able to base-pair with the pre-seed U, and a 3'-supplementary pairing, which might be beneficial for binding affinity and efficiency (Fig.4.3). The binding possibility of all four predicted miRNA seeds was surveyed in the mOR37A 3'UTR. Despite a few matches in other regions of the 3'UTR, no binding occurs in either the proximal or the distal amplicons except the aforementioned ones. Thus, two olfactory enriched miRNAs from the miR-200 family have notable potential target sites in the proximal amplicon.
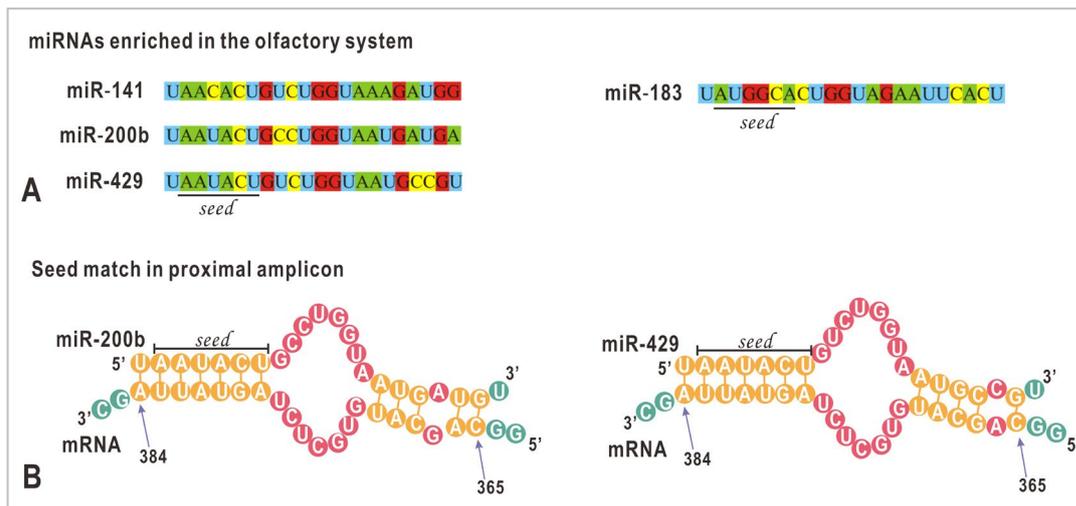
Figure 4.3. Literature suggested olfactory bulb-associated miRNAs and their seed match within the mOR37A mRNA.
A. miRNAs in the olfactory system. Four miRNAs are indicated to be enriched in the mouse olfactory bulb. They are miR-141, miR-200b and miR-429, which belong to the miR-200 family, and miR-183.
B. Seed match in the mOR37A mRNA. Both miR-200b and miR-429 have perfect seed match with the mOR37A mRNA within the proximal amplicon. Numbers indicate mRNA positions after the mOR37A coding region.

If such interaction does occur, what does it imply? One question that one has to address first is why the "RNase sensitivity" is missing in the OE, given that miRNA is also transcribed in the nucleus before any compartmentalization and functionalization. One possibility could be that the nucleus exported pre-miRNA is processed into mature miRNA only when it is translocated to the axon termini. In support of this idea, several studies have revealed that Dicer, the key enzyme responsible for the maturation of miRNAs, is present in growth cones (Hengst et al., 2006; Zhang et al., 2013; Hancock et al., 2014), making local processing of pre-miRNAs possible. Moreover, as a regulatory layer of gene expression, a number of miRNAs are reported to distribute differentially at subcellular level, such as enrichment or depletion in the axon termini (reviewed in Iyer et al., 2014). However, there are still two technical questions that cannot be neglected. One is the spin column purification prior to ddPCR, which, according to the manufacturer's instructions (Zymo Research), recovers DNA fragments ranging from 50bp to 23kb. A mature miRNA should hardly be retrieved, unless it is bound to any larger sequences (e.g., miRNA "sponges"). Another is the annealing temperature during ddPCR, which is higher than the melting temperature calculated from the predicted interaction (12 base pairs). However, the

melting temperature of mature miRNAs has been reported to have a wide range till even 92ºC (1M salt concentration) (Lee et al., 2008), which would then suffice the annealing step during ddPCR. Moreover, the exact composition of the ddPCR Supermix is not clear with regard to the salt concentration, which might influence the final stability of miRNA binding.

When the interaction partner is a miRNA, the functional consequence is likely to be miRNA-mediated target mRNA suppression. It could fit to previous discussions with regard to the developmental aspects of releasing the axonal short isoform from inhibition in the following way. The short isoform is first suppressed from translation by miRNA-mediated gene silencing during the ontogenetic axon guidance phase and afterwards in the maintenance phase, its expression is restored to serve as a maintenance factor.

Parallel to this, I also searched for other long non-coding RNAs (lncRNA) reported in literature in the mouse olfactory system. Two publications with comprehensive analysis of the mouse non-coding transcriptome in the OE and OB are used as a base of the search (Mercer et al., 2008; Camargo et al., 2019). However, as most of the non-coding transcripts are neither annotated nor functionally characterized, I was not able to pin down any candidate by manually aligning the top 25 candidates to the mouse genome. When handling big dataset and when searching for short matches, a more thorough alignment is necessary to retrieve the lncRNA data for identifying potential interaction partners of the mOR37A.

4.2.5 Attempts at the experimental identification of mOR37A interaction partner

In order to identify this unknown RNA interaction partner, several methods have been tried in this thesis, which up to this point have not yielded a result. The basic knowledge about this RNA molecule mainly relies on two facts. First, its sequence on both ends is unknown, which possibly requires ligation of linker sequences to enable PCR-based amplification. Second, it is partially complementary with the proximal target, which enables hybridization-based pulldown assays. Moreover, it is assumed to run in the antisense direction, if any regulatory functions should be exerted on the sense mRNA.

To start with, the template-switch activity of the 5' RACE technology (SMARTer RACE 5'/3' Kit, Clontech) is taken advantage of in probing the unknown 5' end of the RNA molecule. A gene specific primer [37A_3'UTR_241F] was randomly chosen in the proximal amplicon. By performing a standard 5' RACE with this primer, a band of 150bp occured in the nested PCR, however with both OE and OB RACE-ready cDNA (data not shown). This 150bp product was sequenced and it matches largely to the intron sequence of the ubiquitin-conjugating enzyme E2G 1 gene in the antisense direction, where no annotated genes are present (data not shown). It is, therefore, not very likely to be a real hit. Notably, small RNA species would hardly be amplified by this method due to their short sequences.

Following this, a second set of experiment was designed based on RNA-RNA hybridization, namely Northern blot. Difficulties were how to translate the ddPCR annealing temperature to the Northern blot hybridization temperature, which was largely influenced by the presence of formamide, a component absent in the ddPCR Supermix for Probes (Bio-Rad). By using a sense probe covering the short 3'UTR, which should bind the unknown interaction partner, and low stringency conditions of 50°C hybridization temperature, no product was, however, systematically detected (data not shown). Similar to the 5' RACE-based method, no small RNAs can possibly be detected as they might be lost upon electrophoresis due to their short sequences.

At last, a capture experiment was designed, which might bear more resemblance to the ddPCR condition. In brief, it relies on streptavidin-biotin based RNA pulldown and subsequent RT-PCR amplification and cloning. Since RNAs are highly sensitive towards degradation, the RNA pulldown experiment was designed to count on cDNA-RNA hybridization. The experimental procedure is the following. A sense single-stranded cDNA molecule ("*bait cDNA*") covering the proximal amplicon is synthesized via linear PCR. The *bait cDNA* is biotinylated at its 5' end via a biotinylated primer and anchored to the streptavidin beads (Dynabeads Streptavidin Trial Kit, M-270 Streptavidin, Thermo Fisher Scientific). By fishing in the P7 OB total RNA pool for any potential interactions, the captured RNA species (*"prey RNA"*) are eluted from the beads and subject to 3' and 5' RNA ligation of adaptor sequences for downstream PCR amplification. This ligation procedure is explicitly desgined for cloning short RNAs according to literature (Lau et al., 2001). After PCR amplification all products are cloned into a vector and sequenced. While establishing the protocol, a few improvements were seen to be necessary. First, the biotinylated primer for *bait cDNA* synthesis is modified with Biotin-TEG (Eurofins Genomics) on its very 5' nucleotide. This single biotin labeling seems to be too weak to hold the *bait cDNA* on the streptavidin beads during the pulldown elution step, as a temperature gradient ranging from 60°C to 85°C revealed the presence of the *bait cDNA* in the eluates together with the captured RNAs (data not shown). One way to eliminate the undesired elution is to enable multiple biotinylated positions in the primer, which is commercially available from the manufacturer (Eurofins Genomics). Second, the eluted *bait cDNA* cannot be excluded from the RNA ligation step, because the T4 RNA Ligase 1 (New England Biolabs) recognizes both ssRNA and ssDNA as reaction substrate. This leads to serious sequencing difficulties, as the top alignment turns out to be the *bait* itself. I have thus tried to deactivate the 3' end of the *bait cDNA* by addition of the ddTTP (Carl Roth) catalyzed by the Terminal Transferase (New England Biolabs). However, up to this stage the reaction was largely not complete and there were still traces of the *bait cDNA* that can finally be amplified by PCR.

Altogether, the identity of the unknown RNA molecule has not yet been determined experimentally, but the capture experiment carries strong promise for revealing the identity of the interaction partner and should be followed up.
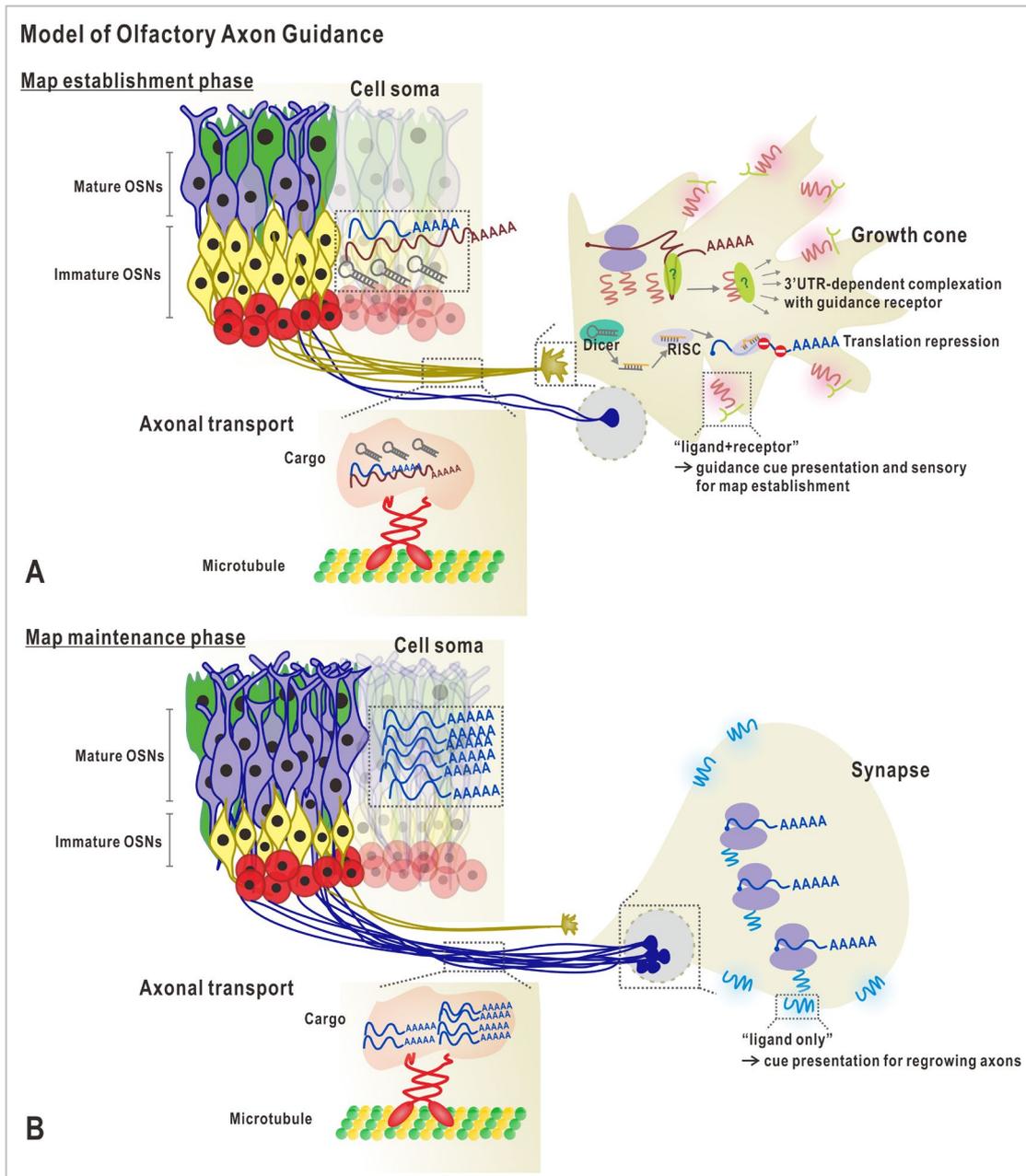
4.3   Model of olfactory axon guidance

Based on the previous knowledge and findings from this thesis, I suggest the following model (Fig.4.4). The model describes the subcellular distribution and functions of different isoforms in OSNs expressing the exemplary receptor mOR37A in two developmental stages P7, where axonal pathfinding is ongoing and the adult, where synaptic connections are largely established but where regeneration is continuing.

In the P7 animal, the olfactory epithelium consists of a large number of immature OSNs and olfactory axon guidance is vigorously taking place. It is thus termed the map establishment phase. The ratio of the short to long isoform is 1:1 in the immature population in the OE. Both isoforms are enclosed in ribonucleoprotein particles and transported to the axon termini, where the short isoform is repressed from translation potentially by small RNA-mediated suppression, and the long isoform is actively translated. The small RNA is believed to be functionally processed only in the axon termini by the endoribonuclease Dicer and subsequently loaded into the RNA-induced silencing complex (RISC). Additional receptor-specific guidance molecules are required for guidance cue sensory. They could be recruited exclusively by the long 3'UTR, by forming a functional signaling complex with the receptor protein from the long isoform. This protein complex is eventually transported to the cell membrane and is involved in the OR-mediated fiber-fiber interactions, which can bring other axons and growth cones of the same type to their stereotyped positions in the OB. Examples of the 3'UTR-dependent protein function can be found in literature. For instance, the long 3'UTR of the E3 ligase BIRC3 is required for the formation of BIRC3 protein complexes, which regulate the surface expression of the receptor CXCR4 and eventually the B cell migration (Lee and Mayr, 2019). In principle, the receptor proteins from both isoforms are identical in their sequence. The necessity of suppressing the translation from the short isoform in the map establishment phase potentially lies in its inability in interacting with or recruiting other guidance-related molecules. In the OR-dependent fiber-fiber interactions, both the ligand and receptor should be present on the OSN growth cones. One might assume that the short isoform alone could only act as a ligand, while the long isoform, by recruiting other receptors, could act as both the ligand and receptor for guidance. During the establishment phase, all axons should stay highly responsive for primary targeting or for error correction after erroneous synaptogenesis, and thus the long isoform together with its recruited guidance molecule is favoured. As a ligand only, the short isoform renders the growth cones desensitized towards the reception of guidance signaling necessary for the map formation and is thus suppressed from translation.

In the adult animal, the olfactory epithelium is largely composed of mature OSNs and the glomerular map is established. It is thus termed the map maintenance phase. Since the long isoform exhibits a maturation-dependent regulation, its expression is switched off and the short

isoform is dominant in the mature population. Similarly, the short isoform is translocated to the axon termini but the axonal repression is abolished, resulting in normal translation activity. During the maintenance phase, the framework is finished and only the regrowing axons have to be responsive towards external cues. The receptor protein from the short isoform acting as a ligand is thus primary in providing guidance cues for the small regrowing population of immature OSNs which still make the long isoform as well.
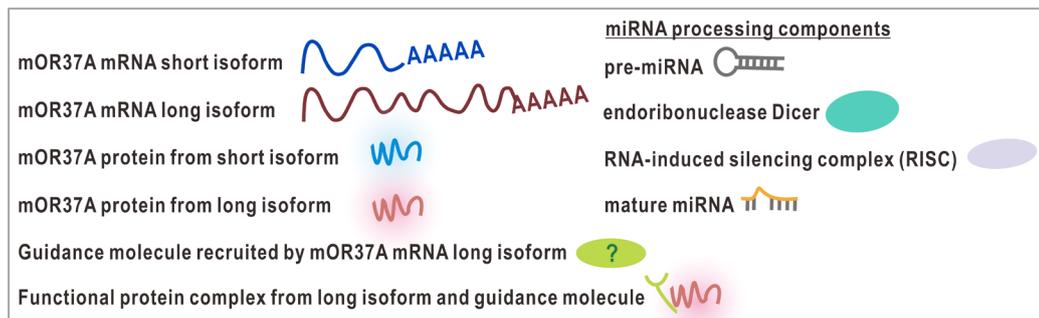
Figure 4.4. Hypothetical model of olfactory axon guidance.
A. Subcellular distribution and functions of mOR37A isoforms in the map establishment phase. In the P7 animal, the ratio of the short to long isoform is 1:1 in the immature population in the OE. Both isoforms are transported towards the axon termini, where the short isoform is inhibited from translation by small RNA-mediated repression and the long isoform is actively translated. The long isoform recruits other receptor-specific guidance molecules with its long 3'UTR and thereby forms a functional protein complex, which is localized to the membrane. This protein complex is involved in the OR-mediated fiber-fiber interactions essential for the olfactory map formation by functioning both as ligand and receptor. This enables guidance sensory during primary targeting and error correction. The short isoform is suppressed from translation during the establishment phase, as it might only function as a guidance ligand, which renders the growth cones desensitized towards signals sensing, thus impedes the map formation.
B. Subcellular distribution and functions of mOR37A isoforms in the map maintenance phase. In the adult animal, the short isoform is dominant in the mature population in the OE due to the maturation-dependent switch off of the long isoform. The short isoform is transported towards the axon termini but its repressive state is removed, resulting in normal translational activity in the axon termini. Since in the maintenance phase, the olfactory framework is largely established, the short isoform can function as a ligand for regrowing axons in order to maintain the exisiting circuits.

4.4 Outlook

To further refine the mechanism of olfactory axon guidance, several questions should be addressed in the future.

To start with, to analyze the whole olfactory transcriptome and to understand the transcript structure of odorant receptor mRNAs, RNAseq could be performed.

Furthermore, the identity of the unknown RNA interaction molecule of the proximal amplicon should be clarified. Therefore the strategy of the capture experiment would greatly aid the understanding of the dynamic regulation of different isoforms in the growth cones.

Moreover, there are a few points in the model that should be addressed. For example, the long isoform is suggested to be distinguishable from the short isoform, with regard to its immunity towards the unknown RNA molecule-mediated PCR inhibition and its ability to recruit other guidance-related molecules. Such difference is assumed to be a result of its sequence difference and its extended 3'UTR. The sequence difference, presumably caused by post-transcriptional modifications from RNA modifying enzymes, can be addressed by sequencing of the RT-PCR products with primers flanking the proximal amplicon region. To investigate the protein interaction partners that can exclusively be recruited by the long 3'UTR, both *in vitro* and *in vivo* methods can be applied (reviewed in Ramanathan, 2019). For instance, *in vitro* biotinylated-RNA pulldown assay utilizing end-labeled biotinylated RNA as a bait to pull down the interaction proteins in the cell lysate; or *in vivo* protein-RNA crosslinking based methods with covalently crosslinking of the protein-RNA complexes and extraction the interaction proteins could be performed. With both approaches, immunoprecipitation is only possible, if there are known candidate proteins; otherwise the captured proteins should ideally be characterized by mass spectrometry analysis.

Equally promising might be the establishment of an OSN culture system. In heterologous systems, the effective expression of the olfactory receptor protein is often impeded by the poor membrane transport of the receptor. With a functional *in vitro* system, genetic manipulations such as deletion of the long isoform could be performed with ease and the dynamic axonal transport could also be followed up.

5. References

Adamson, T. E., Shutt, D. C., and Price, D. H. (2005). Functional coupling of cleavage and polyadenylation with transcription of mRNA. <u>J Biol Chem</u> **280**(37): 32262-32271.

Ahn, S. H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. <u>Mol Cell</u> **13**(1): 67-76.

Albeanu, D. F., Provost, A. C., Agarwal, P., Soucy, E. R., Zak, J. D., and Murthy, V. N. (2018). Olfactory marker protein (OMP) regulates formation and refinement of the olfactory glomerular map. <u>Nat Commun</u> **9**(1): 5073.

An, J. J., Gharami, K., Liao, G. Y., Woo, N. H., Lau, A. G., Vanevski, F., Torre, E. R., Jones, K. R., Feng, Y., Lu B., and Xu, B. (2008). Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. <u>Cell</u> **134**(1): 175-187.

Armelin-Correa, L. M., Gutiyama, L. M., Brandt, D. Y., and Malnic, B. (2014). Nuclear compartmentalization of odorant receptor genes. <u>Proc Natl Acad Sci U S A</u> **111**(7): 2782-2787.

Asai, H., Kasai, H., Matsuda, Y., Yamazaki, N., Nagawa, F., Sakano, H., and Tsuboi, A. (1996). Genomic structure and transcription of a murine odorant receptor gene: differential initiation of transcription in the olfactory and testicular cells. <u>Biochem Biophys Res Commun</u> **221**(2): 240-247.

Assens, A., Dal Col, J. A., Njoku, A., Dietschi, Q., Kan, C., Feinstein, P., Carleton, A., and Rodriguez, I. (2016). Alteration of Nrp1 signaling at different stages of olfactory neuron maturation promotes glomerular shifts along distinct axes in the olfactory bulb. <u>Development</u> **143**(20): 3817-3825.

Bader, A., Klein, B., Breer, H., and Strotmann, J. (2012). Connectivity from OR37 expressing olfactory sensory neurons to distinct cell types in the hypothalamus. <u>Front Neural Circuits</u> **6**: 84.

Bak, M., Silahtaroglu, A., Moller, M., Christensen, M., Rath, M. F., Skryabin, B., Tommerup, N., and Kauppinen, S. (2008). MicroRNA expression in the adult mouse central nervous system. <u>RNA</u> **14**(3): 432-444.

Barnea, G., O'Donnell, S., Mancia, F., Sun, X., Nemes, A., Mendelsohn, M., and Axel, R. (2004). Odorant receptors on axon termini in the brain. <u>Science</u> **304**(5676): 1468.

Barreau, C., Paillard, L., and Osborne, H. B. (2005). AU-rich elements and associated factors: are there unifying principles? <u>Nucleic Acids Res</u> **33**(22): 7138-7150.

Bartel, D. P. (2018). Metazoan MicroRNAs. <u>Cell</u> **173**(1): 20-51.

Bautze, V., Bar, R., Fissler, B., Trapp, M., Schmidt, D., Beifuss, U., Bufe, B., Zufall, F., Breer, H., and Strotmann, J. (2012). Mammalian-specific OR37 receptors are differentially activated by distinct odorous fatty aldehydes. <u>Chem Senses</u> **37**(5): 479-493.

Bautze, V., Schwack, W., Breer, H., and Strotmann, J. (2014). Identification of a natural source for the OR37B ligand. <u>Chem Senses</u> **39**(1): 27-38.

REFERENCES

Beclin, C., Follert, P., Stappers, E., Barral, S., Core, N., de Chevigny, A., Magnone, V., Lebrigand, K., Bissels, U., Huylebroeck, D., Bosio, A., Barbry, P., Seuntjens, E., and Cremer, H. (2016). miR-200 family controls late steps of postnatal forebrain neurogenesis via Zeb2 inhibition. Sci Rep **6**: 35729.

Belluscio, L., Gold, G. H., Nemes, A., and Axel, R. (1998). Mice deficient in G(olf) are anosmic. Neuron **20**(1): 69-81.

Berkovits, B. D., and Mayr, C. (2015). Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. Nature **522**(7556): 363-367.

Brawerman, G. (1981). The Role of the poly(A) sequence in mammalian messenger RNA. CRC Crit Rev Biochem **10**(1): 1-38.

Brechbuhl, J., Klaey, M., and Broillet, M. C. (2008). Grueneberg ganglion cells mediate alarm pheromone detection in mice. Science **321**(5892): 1092-1095.

Breer, H., Fleischer, J., and Strotmann, J. (2006). The sense of smell: multiple olfactory subsystems. Cell Mol Life Sci **63**(13): 1465-1475.

Brown, C. E., and Sachs, A. B (1998). Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. Mol Cell Biol **18**(11): 6548-6559.

Brown, K. M., and Gilmartin, G. M. (2003). A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. Mol Cell **12**(6): 1467-1476.

Brunet, L. J., Gold, G. H., and Ngai, J (1996). General anosmia caused by a targeted disruption of the mouse olfactory cyclic nucleotide-gated cation channel. Neuron **17**(4): 681-693.

Buck, L., and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell **65**(1): 175-187.

Buck, L. B. (2000). The molecular architecture of odor and pheromone sensing in mammals. Cell **100**(6): 611-618.

Bulfone, A., Wang, F., Hevner, R., Anderson, S., Cutforth, T., Chen, S., Meneses, J., Pedersen, R., Axel R., and Rubenstein, J. L. (1998). An olfactory sensory map develops in the absence of normal projection neurons or GABAergic interneurons. Neuron **21**(6): 1273-1282.

Camargo, A. P., Nakahara, T. S., Firmino, L. E. R., Netto, P. H. M., do Nascimento, J. B. P., Donnard, E. R., Galante, P. A. F., Carazzolle, M. F., Malnic B., and Papes, F. (2019). Uncovering the mouse olfactory long non-coding transcriptome with a novel machine-learning model. DNA Res. doi:10.1093/dnares/dsz015

Carrington, J. C., and Ambros, V. (2003). Role of microRNAs in plant and animal development. Science **301**(5631): 336-338.

Chang, T. H., Huang, H. Y., Hsu, J. B., Weng, S. L., Horng, J. T., and Huang, H. D. (2013). An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. BMC Bioinformatics **14 Suppl 2**: S4.

Chess, A., Simon, I., Cedar, H., and Axel, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. Cell **78**(5): 823-834.

Choi, P. S., Zakhary, L., Choi, W. Y., Caron, S., Alvarez-Saavedra, E., Miska, E. A., McManus, M., Harfe, B., Giraldez, A. J., Horvitz, H. R., Schier A. F., and Dulac, C. (2008). Members of the miRNA-200 family regulate olfactory neurogenesis. <u>Neuron</u> **57**(1): 41-55.

Chuah, M. I., and West, A. K. (2002). Cellular and molecular biology of ensheathing cells. <u>Microsc Res Tech</u> **58**(3): 216-227.

Clowney, E. J., LeGros, M. A., Mosley, C. P., Clowney, F. G., Markenskoff-Papadimitriou, E. C., Myllys, M., Barnea, G., Larabell, C. A., and Lomvardas, S. (2012). Nuclear aggregation of olfactory receptor genes governs their monogenic expression. <u>Cell</u> **151**(4): 724-737.

Colgan, D. F., and Manley, J. L. (1997). Mechanism and regulation of mRNA polyadenylation. <u>Genes Dev</u> **11**(21): 2755-2766.

Cutforth, T., Moring, L., Mendelsohn, M., Nemes, A., Shah, N. M., Kim, M. M. Frisen, J. and Axel, R. (2003). Axonal ephrin-As and odorant receptors: coordinate determination of the olfactory sensory map. <u>Cell</u> **114**(3): 311-322.

Dalton, R. P., Lyons, D. B., and Lomvardas, S. (2013). Co-opting the unfolded protein response to elicit olfactory receptor feedback. <u>Cell</u> **155**(2): 321-332.

DeMaria, S., and Ngai, J. (2010). The cell biology of smell. <u>J Cell Biol</u> **191**(3): 443-452.

Derti, A., Garrett-Engele, P., Macisaac, K. D., Stevens, R. C., Sriram, S., Chen, R., Rohl, C. A., Johnson J. M., and Babak, T. (2012). A quantitative atlas of polyadenylation in five mammals. <u>Genome Res</u> **22**(6): 1173-1183.

Doulazmi, M., Cros, C., Dusart, I., Trembleau, A., and Dubacq, C. (2019). Alternative polyadenylation produces multiple 3' untranslated regions of odorant receptor mRNAs in mouse olfactory sensory neurons. <u>BMC Genomics,</u> **20**(1), 577. doi:10.1186/s12864-019-5927-3

Edmonds, M., and Abrams, R. (1960). Polynucleotide biosynthesis: formation of a sequence of adenylate units from adenosine triphosphate by an enzyme from thymus nuclei. <u>J Biol Chem</u> **235**: 1142-1149.

Elkon, R., Drost, J., van Haaften, G., Jenal, M., Schrier, M., Oude Vrielink, J. A., and Agami, R. (2012). E2F mediates enhanced alternative polyadenylation in proliferation. <u>Genome Biol</u> **13**(7): R59.

Elkon, R., Ugalde, A. P., and Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function. <u>Nat Rev Genet</u> **14**(7): 496-506.

Falk, W. (2015). Entwicklungsabhängige Regulation der Transkriptisoformen des olfaktorischen Rezeptors mOR37a. Dissertation, Karlsruhe Institute of Technology.

Fan, J., Jia, L., Li, Y., Ebrahim, S., May-Simera, H., Wood, A., Morell, R. J., Liu, P., Lei, J., Kachar, B., Belluscio, L., Qian, H., Li, T., Li, W., Wistow G., and Dong, L. (2017). Maturation arrest in early postnatal sensory receptors by deletion of the miR-183/96/182 cluster in mouse. <u>Proc Natl Acad Sci U S A</u> **114**(21): E4271-E4280.

Farbman, A. I. (1992). <u>Cell biology of olfaction</u>. Cambridge ; New York, N.Y., USA, Cambridge University Press.

Farbman, A. I., and Margolis, F. L. (1980). Olfactory marker protein during ontogeny: immunohistochemical localization. <u>Dev Biol</u> **74**(1): 205-215.

Feinstein, P., Bozza, T., Rodriguez, I., Vassalli, A., and Mombaerts, P. (2004). Axon guidance of mouse olfactory sensory neurons by odorant receptors and the beta2 adrenergic receptor. <u>Cell</u> **117**(6): 833-846.

Feinstein, P., and Mombaerts, P. (2004). A contextual model for axonal sorting into glomeruli in the mouse olfactory system. <u>Cell</u> **117**(6): 817-831.

Flanagan, J. G. (2006). Neural map specification by gradients. <u>Curr Opin Neurobiol</u> **16**(1): 59-66.

Fleischer, J., Hass, N., Schwarzenbacher, K., Besser, S., and Breer, H. (2006). A novel population of neuronal cells expressing the olfactory marker protein (OMP) in the anterior/dorsal region of the nasal cavity. <u>Histochem Cell Biol</u> **125**(4): 337-349.

Fletcher, R. B., Das, D., Gadye, L., Street, K. N., Baudhuin, A., Wagner, A., Cole, M. B., Flores, Q., Choi, Y. G., Yosef, N., Purdom, E., Dudoit, S., Risso D., and Ngai, J. (2017). Deconstructing Olfactory Stem Cell Trajectories at Single-Cell Resolution. <u>Cell Stem Cell</u> **20**(6): 817-830 e818.

Friedrich, R. W., and Laurent, G. (2001). Dynamic optimization of odor representations by slow temporal patterning of mitral cell activity. <u>Science</u> **291**(5505): 889-894.

Fujita, T., Yuno, M., Kitaura, F., and Fujii, H. (2018). A refined two-step oligoribonucleotide interference-PCR method for precise discrimination of nucleotide differences. <u>Sci Rep</u> **8**(1): 17195.

Fulle, H. J., Vassar, R., Foster, D. C., Yang, R. B., Axel, R., and Garbers, D. L. (1995). A receptor guanylyl cyclase expressed specifically in olfactory sensory neurons. <u>Proc Natl Acad Sci U S A</u> **92**(8): 3571-3575.

Fuss, S. H., Omura, M., and Mombaerts, P. (2005). The Grueneberg ganglion of the mouse projects axons to glomeruli in the olfactory bulb. <u>Eur J Neurosci</u> **22**(10): 2649-2654.

Garneau, N. L., Wilusz, J., and Wilusz, C. J. (2007). The highways and byways of mRNA decay. <u>Nat Rev Mol Cell Biol</u> **8**(2): 113-126.

Getchell, M. L., and Getchell, T. V. (1992). Fine structural aspects of secretion and extrinsic innervation in the olfactory mucosa. <u>Microsc Res Tech</u> **23**(2): 111-127.

Glusman, G., Clifton, S., Roe, B., and Lancet, D. (1996). Sequence analysis in the olfactory receptor gene cluster on human chromosome 17: recombinatorial events affecting receptor diversity. <u>Genomics</u> **37**(2): 147-160.

Godfrey, P. A., Malnic, B., and Buck, L. B. (2004). The mouse olfactory receptor gene family. <u>Proc Natl Acad Sci U S A</u> **101**(7): 2156-2161.

Gong, C., and Maquat, L. E. (2011). lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. <u>Nature</u> **470**(7333): 284-288.

Graziadei, P. P., and Graziadei, G. A. (1979). Neurogenesis and neuron regeneration in the olfactory system of mammals. I. Morphological aspects of differentiation and structural organization of the olfactory sensory neurons. J Neurocytol **8**(1): 1-18.

Gruber, A. J., Schmidt, R., Gruber, A. R., Martin, G., Ghosh, S., Belmadani, M., Keller, W., and Zavolan, M. (2016). A comprehensive analysis of 3' end sequencing data sets reveals novel polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation. Genome Res **26**(8): 1145-1159.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R., and Hofacker, I. L. (2008). The Vienna RNA websuite. Nucleic Acids Res **36**(Web Server issue): W70-74.

Gruneberg, H. (1973). A ganglion probably belonging to the N. terminalis system in the nasal mucosa of the mouse. Z Anat Entwicklungsgesch **140**(1): 39-52.

Haag, N. (2009). Differenzielle Einzelzell-Transkriptomanalyse zur Identifizierung geruchsrezeptorassoziierter Differenzierungsmoleküle während der Entwicklung des olfaktorischen Systems der Maus. Dissertation, Universität Jena.

Haines, D. E., and Mihailoff, G. A. (2018). Fundamental neuroscience for basic and clinical applications. Philadelphia, PA, Elsevier.

Hanchate, N. K., Kondoh, K., Lu, Z., Kuang, D., Ye, X., Qiu, X., Pachter, L., Trapnell, C., and Buck, L. B. (2015). Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. Science **350**(6265): 1251-1255.

Hancock, M. L., Preitner, N., Quan, J., and Flanagan, J. G. (2014). MicroRNA-132 is enriched in developing axons, locally regulates Rasa1 mRNA, and promotes axon extension. J Neurosci **34**(1): 66-78.

Helisch, H. (2014). Molekularbiologische Analyse eines Geruchsrezeptor-assoziierten, entwicklungsabhängig regulierten, nicht-codierenden Transkripts der Maus. Master Thesis, Karlsruhe Institute of Technology.

Hengst, U., Cox, L. J., Macosko, E. Z., and Jaffrey, S. R. (2006). Functional and selective RNA interference in developing axons and growth cones. J Neurosci **26**(21): 5727-5732.

Hilgers, V., Lemke, S. B., and Levine, M. (2012). ELAV mediates 3' UTR extension in the Drosophila nervous system. Genes Dev **26**(20): 2259-2264.

Hilgers, V., Perry, M. W., Hendrix, D., Stark, A., Levine, M., and Haley, B. (2011). Neural-specific elongation of 3' UTRs during Drosophila development. Proc Natl Acad Sci U S A **108**(38): 15864-15869.

Holbrook, E. H., Szumowski, K. E., and Schwob, J. E. (1995). An immunochemical, ultrastructural, and developmental characterization of the horizontal basal cells of rat olfactory epithelium. J Comp Neurol **363**(1): 129-146.

Holland, P. M., Abramson, R. D., Watson, R., and Gelfand, D. H. (1991). Detection of specific polymerase chain reaction product by utilizing the 5'-3' exonuclease activity of Thermus aquaticus DNA polymerase. Proc Natl Acad Sci U S A **88**(16): 7276-7280.

Hoppe, R., Breer, H., and Strotmann, J. (2003). Organization and evolutionary relatedness of OR37 olfactory receptor genes in mouse and human. <u>Genomics</u> **82**(3): 355-364.

Hoppe, R., Lambert, T. D., Samollow, P. B., Breer, H., and Strotmann, J. (2006). Evolution of the "OR37" subfamily of olfactory receptors: a cross-species comparison. <u>J Mol Evol</u> **62**(4): 460-472.

Hoque, M., Ji, Z., Zheng, D., Luo, W., Li, W., You, B., Park, J. Y., Yehia, G., and Tian, B. (2013). Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. <u>Nat Methods</u> **10**(2): 133-139.

Hu, J., Zhong, C., Ding, C., Chi, Q., Walz, A., Mombaerts, P., Matsunami, H. and Luo, M. (2007). Detection of near-atmospheric concentrations of CO_2 by an olfactory subsystem in the mouse. <u>Science</u> **317**(5840): 953-957.

Huard, J. M., and Schwob, J. E. (1995). Cell cycle of globose basal cells in rat olfactory epithelium. <u>Dev Dyn</u> **203**(1): 17-26.

Humphries, B., and Yang, C. (2015). The microRNA-200 family: small molecules with novel roles in cancer development, progression and therapy. <u>Oncotarget</u> **6**(9): 6472-6498.

Ibarra-Soria, X., Levitin, M. O., Saraiva, L. R., and Logan, D. W. (2014). The olfactory transcriptomes of mice. <u>PLoS Genet</u> **10**(9): e1004593.

Imai, T., Suzuki, M., and Sakano, H. (2006). Odorant receptor-derived cAMP signals direct axonal targeting. <u>Science</u> **314**(5799): 657-661.

Imai, T., Yamazaki, T., Kobayakawa, R., Kobayakawa, K., Abe, T., Suzuki, M., and Sakano, H. (2009). Pre-target axon sorting establishes the neural map topography. <u>Science</u> **325**(5940): 585-590.

Iwai, N., Zhou, Z., Roop, D. R., and Behringer, R. R. (2008). Horizontal basal cells are multipotent progenitors in normal and injured adult olfactory epithelium. <u>Stem Cells</u> **26**(5): 1298-1306.

Iyer, A. N., Bellon, A., and Baudet, M. L. (2014). microRNAs in axon guidance. <u>Front Cell Neurosci</u> **8**: 78.

Jensen, K. B., Dredge, B. K., Stefani, G., Zhong, R., Buckanovich, R. J., Okano, H. J., Yang, Y. Y., and Darnell, R. B. (2000). Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. <u>Neuron</u> **25**(2): 359-371.

Ji, Z., Lee, J. Y., Pan, Z., Jiang, B., and Tian, B. (2009). Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. <u>Proc Natl Acad Sci U S A</u> **106**(17): 7028-7033.

Ji, Z., Luo, W., Li, W., Hoque, M., Pan, Z., Zhao, Y., and Tian, B. (2011). Transcriptional activity regulates alternative cleavage and polyadenylation. <u>Mol Syst Biol</u> **7**: 534.

Ji, Z., and Tian, B. (2009). Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. <u>PLoS One</u> **4**(12): e8419.

Jones, D. T., and Reed, R. R. (1989). Golf: an olfactory neuron specific-G protein involved in odorant signal transduction. Science **244**(4906): 790-795.

Juilfs, D. M., Fulle, H. J., Zhao, A. Z., Houslay, M. D., Garbers, D. L., and Beavo, J. A. (1997). A subset of olfactory neurons that selectively express cGMP-stimulated phosphodiesterase (PDE2) and guanylyl cyclase-D define a unique olfactory signal transduction pathway. Proc Natl Acad Sci U S A **94**(7): 3388-3395.

Kalkatawi, M., Rangkuti, F., Schramm, M., Jankovic, B. R., Kamau, A., Chowdhary, R., Archer J. A., and Bajic, V. B. (2013). Dragon PolyA Spotter: predictor of poly(A) motifs within human genomic DNA sequences. Bioinformatics **29**(11): 1484.

Kaneko-Goto, T., Yoshihara, S., Miyazaki, H., and Yoshihara, Y. (2008). BIG-2 mediates olfactory axon convergence to target glomeruli. Neuron **57**(6): 834-846.

Kasowski, H. J., Kim, H., and Greer, C. A. (1999). Compartmental organization of the olfactory bulb glomerulus. J Comp Neurol **407**(2): 261-274.

Katada, S., Hirokawa, T., Oka, Y., Suwa, M., and Touhara, K. (2005). Structural basis for a broad but selective ligand spectrum of a mouse olfactory receptor: mapping the odorant-binding site. J Neurosci **25**(7): 1806-1815.

Kobilka, B. K., and Deupi, X. (2007). Conformational complexity of G-protein-coupled receptors. Trends Pharmacol Sci **28**(8): 397-406.

Koos, D. S., and Fraser, S. E. (2005). The Grueneberg ganglion projects to the olfactory bulb. Neuroreport **16**(17): 1929-1932.

Koushika, S. P., Soller, M., and White, K. (2000). The neuron-enriched splicing pattern of Drosophila erect wing is dependent on the presence of ELAV protein. Mol Cell Biol **20**(5): 1836-1845.

Kubick, S., Strotmann, J., Andreini, I., and Breer, H. (1997). Subfamily of olfactory receptors characterized by unique structural features and expression patterns. J Neurochem **69**(2): 465-475.

Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science **294**(5543): 858-862.

Lee, I., Ajay, S. S., Chen, H., Maruyama, A., Wang, N., McInnis, M. G., and Athey. B. D. (2008). Discriminating single-base difference miRNA expressions using microarray Probe Design Guru (ProDeG). Nucleic Acids Res **36**(5): e27.

Lee, S. H., and Mayr, C. (2019). Gain of Additional BIRC3 Protein Functions through 3'-UTR-Mediated Protein Complex Formation. Mol Cell **74**(4): 701-712 e709.

Legendre, M., Ritchie, W., Lopez, F., and Gautheret, D. (2006). Differential repression of alternative transcripts: a screen for miRNA targets. PLoS Comput Biol **2**(5): e43.

Levai, O., and Strotmann, J. (2003). Projection pattern of nerve fibers from the septal organ: DiI-tracing studies with transgenic OMP mice. Histochem Cell Biol **120**(6): 483-492.

Licatalosi, D. D., Mele, A., Fak, J. J., Ule, J., Kayikci, M., Chi, S. W., Clark, T. A., Schweitzer, A. C., Blume, J. E., Wang, X., Darnell J. C., and Darnell, R. B. (2008). HITS-

CLIP yields genome-wide insights into brain alternative RNA processing. Nature **456**(7221): 464-469.

Lin, D. M., Wang, F., Lowe, G., Gold, G. H., Axel, R., Ngai, J., and Brunet, L. (2000). Formation of precise connections in the olfactory bulb occurs in the absence of odorant-evoked neuronal activity. Neuron **26**(1): 69-80.

Liu, A. H., Zhang, X., Stolovitzky, G. A., Califano, A., and Firestein, S. J. (2003). Motif-based construction of a functional map for mammalian olfactory receptors. Genomics **81**(5): 443-456.

Liu, H., Han, H., Li, J., and Wong, L. (2005). DNAFSMiner: a web-based software toolbox to recognize two types of functional sites in DNA sequences. Bioinformatics **21**(5): 671-673.

Low, V. F., and Mombaerts, P. (2017). Odorant receptor proteins in the mouse main olfactory epithelium and olfactory bulb. Neuroscience **344**: 167-177.

Lyons, D. B., Allen, W. E., Goh, T., Tsai, L., Barnea, G., and Lomvardas, S. (2013). An epigenetic trap stabilizes singular olfactory receptor expression. Cell **154**(2): 325-336.

Ma, L., Wu, Y., Qiu, Q., Scheerer, H., Moran, A., and Yu, C. R. (2014). A developmental switch of axon targeting in the continuously regenerating mouse olfactory system. Science **344**(6180): 194-197.

Ma, M., Grosmaitre, X., Iwema, C. L., Baker, H., Greer, C. A., and Shepherd, G. M. (2003). Olfactory signal transduction in the mouse septal organ. J Neurosci **23**(1): 317-324.

Ma, W. P., Kaiser, M. W., Lyamicheva, N., Schaefer, J. J., Allawi, H. T., Takova, T., Neri, B. P., and Lyamichev, V. I. (2000). RNA template-dependent 5' nuclease activity of Thermus aquaticus and Thermus thermophilus DNA polymerases. J Biol Chem **275**(32): 24693-24700.

Magklara, A., Yen, A., Colquitt, B. M., Clowney, E. J., Allen, W., Markenscoff-Papadimitriou, E. Evans, Z. A., Kheradpour, P., Mountoufaris, G., Carey, C., Barnea, G., Kellis M., and Lomvardas, S. (2011). An epigenetic signature for monoallelic olfactory receptor expression. Cell **145**(4): 555-570.

Malnic, B., Godfrey, P. A., and Buck, L. B. (2004). The human olfactory receptor gene family. Proc Natl Acad Sci U S A **101**(8): 2584-2589.

Malnic, B., Hirono, J., Sato, T., and Buck, L. B. (1999). Combinatorial receptor codes for odors. Cell **96**(5): 713-723.

Mandel, C. R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3'-end processing. Cell Mol Life Sci **65**(7-8): 1099-1122.

Markenscoff-Papadimitriou, E., Allen, W. E., Colquitt, B. M., Goh, T., Murphy, K. K., Monahan, K., Mosley, C. P., Ahituv, N., and Lomvardas, S. (2014). Enhancer interaction networks as a means for singular olfactory receptor expression. Cell **159**(3): 543-557.

Martin, K. C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. Cell **136**(4): 719-730.

Martincic, K., Alkan, S. A., Cheatle, A., Borghesi, L., and Milcarek, C. (2009). Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. Nat Immunol **10**(10): 1102-1109.

Mattick, J. S., and Makunin, I. V. (2006). Non-coding RNA. Hum Mol Genet **15 Spec No 1**: R17-29.

Menco, B. P. (1980). Qualitative and quantitative freeze-fracture studies on olfactory and nasal respiratory epithelial surfaces of frog, ox, rat, and dog. II. Cell apices, cilia, and microvilli. Cell Tissue Res **211**(1): 5-29.

Mercer, T. R., Dinger, M. E., Sunkin, S. M., Mehler, M. F., and Mattick, J. S. (2008). Specific expression of long noncoding RNAs in the mouse brain. Proc Natl Acad Sci U S A **105**(2): 716-721.

Millar, A. A., and Waterhouse, P. M. (2005). Plant and animal microRNAs: similarities and differences. Funct Integr Genomics **5**(3): 129-135.

Miyamichi, K., Serizawa, S., Kimura, H. M., and Sakano, H. (2005). Continuous and overlapping expression domains of odorant receptor genes in the olfactory epithelium determine the dorsal/ventral positioning of glomeruli in the olfactory bulb. J Neurosci **25**(14): 3586-3592.

Mombaerts, P., Wang, F., Dulac, C., Chao, S. K., Nemes, A., Mendelsohn, M. Edmondson, J., and Axel, R. (1996). Visualizing an olfactory sensory map. Cell **87**(4): 675-686.

Monahan, K., Horta, A., and Lomvardas, S. (2019). LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. Nature **565**(7740): 448-453.

Mori, K., and Sakano, H. (2011). How is the olfactory map formed and interpreted in the mammalian brain? Annu Rev Neurosci **34**: 467-499.

Mori, K., von Campenhause, H., and Yoshihara, Y. (2000). Zonal organization of the mammalian main and accessory olfactory systems. Philos Trans R Soc Lond B Biol Sci **355**(1404): 1801-1812.

Morris, K. V., and Mattick, J. S. (2014). The rise of regulatory RNA. Nat Rev Genet **15**(6): 423-437.

Morrison, E. E., and Costanzo, R. M. (1992). Morphology of olfactory epithelium in humans and other vertebrates. Microsc Res Tech **23**(1): 49-61.

Mountoufaris, G., Chen, W. V., Hirabayashi, Y., O'Keeffe, S., Chevee, M., Nwakeze, C. L. Polleux, F., and Maniatis, T. (2017). Multicluster Pcdh diversity is required for mouse olfactory neural circuit assembly. Science **356**(6336): 411-414.

Movahedi, K., Grosmaitre, X., and Feinstein, P. (2016). Odorant receptors can mediate axonal identity and gene choice via cAMP-independent mechanisms. Open Biol **6**(7).

Nakashima, A., Ihara, N., Shigeta, M., Kiyonari, H., Ikegaya, Y., and Takeuchi, H. (2019). Structured spike series specify gene expression patterns for olfactory circuit formation. Science **365**(6488). doi:10.1126/science.aaw5030

Nakashima, A., Takeuchi, H., Imai, T., Saito, H., Kiyonari, H., Abe, T., Chen, M., Weinstein, L. S., Yu, C. R., Storm, D. R., Nishizumi, H., and Sakano, H. (2013). Agonist-independent GPCR activity regulates anterior-posterior targeting of olfactory sensory neurons. Cell **154**(6): 1314-1325.

Ngai, J., Dowling, M. M., Buck, L., Axel, R., and Chess, A. (1993). The family of genes encoding odorant receptors in the channel catfish. Cell **72**(5): 657-666.

Nusser, Z., Kay, L. M., Laurent, G., Homanics, G. E., and Mody, I. (2001). Disruption of GABA(A) receptors on GABAergic interneurons leads to increased oscillatory power in the olfactory bulb network. J Neurophysiol **86**(6): 2823-2833.

Orona, E., Rainer, E. C., and Scott, J. W. (1984). Dendritic and axonal organization of mitral and tufted cells in the rat olfactory bulb. J Comp Neurol **226**(3): 346-356.

Pifferi, S., Pascarella, G., Boccaccio, A., Mazzatenta, A., Gustincich, S., Menini, A., and Zucchelli, S. (2006). Bestrophin-2 is a candidate calcium-activated chloride channel involved in olfactory transduction. Proc Natl Acad Sci U S A **103**(34): 12929-12934.

Pinto, P. A., Henriques, T., Freitas, M. O., Martins, T., Domingues, R. G., Wyrzykowska, P. S. Coelho, P. A., Carmo, A. M., Sunkel, C. E., Proudfoot, N. J., and Moreira, A. (2011). RNA polymerase II kinetics in polo polyadenylation signal selection. EMBO J **30**(12): 2431-2444.

Price, J. L., and Powell, T. P. (1970). The mitral and short axon cells of the olfactory bulb. J Cell Sci **7**(3): 631-651.

Price, J. L., and Powell, T. P. (1970). The morphology of the granule cells of the olfactory bulb. J Cell Sci **7**(1): 91-123.

Proudfoot, N. J. (2011). Ending the message: poly(A) signals then and now. Genes Dev **25**(17): 1770-1782.

Ramanathan, M., Porter, D. F., and Khavari, P. A. (2019). Methods to study RNA-protein interactions. Nat Methods **16**(3): 225-234.

Ressler, K. J., Sullivan, S. L., and Buck, L. B. (1994). Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. Cell **79**(7): 1245-1255.

Robinow, S., and White, K. (1991). Characterization and spatial distribution of the ELAV protein during Drosophila melanogaster development. J Neurobiol **22**(5): 443-461.

Ron, D., and Walter, P. (2007). Signal integration in the endoplasmic reticulum unfolded protein response. Nat Rev Mol Cell Biol **8**(7): 519-529.

Rozenblatt-Rosen, O., Nagaike, T., Francis, J. M., Kaneko, S., Glatt, K. A., Hughes, C. M. LaFramboise, T., Manley, J. L., and Meyerson, M. (2009). The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3' mRNA processing factors. Proc Natl Acad Sci U S A **106**(3): 755-760.

Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A., and Burge, C. B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science **320**(5883): 1643-1647.

Schaefer, M. L., Young, D. A., and Restrepo, D. (2001). Olfactory fingerprints for major histocompatibility complex-determined body odors. J Neurosci **21**(7): 2481-2487.

Schwob, J. E. (2002). Neural regeneration and the peripheral olfactory system. Anat Rec **269**(1): 33-49.

Schwob, J. E., Jang, W., Holbrook, E. H., Lin, B., Herrick, D. B., Peterson, J. N., and Hewitt Coleman, J. (2017). Stem and progenitor cells of the mammalian olfactory epithelium: Taking poietic license. J Comp Neurol **525**(4): 1034-1054.

Senfter, D., Madlener, S., Krupitza, G., and Mader, R. M. (2016). The microRNA-200 family: still much to discover. Biomol Concepts **7**(5-6): 311-319.

Serizawa, S., Ishii, T., Nakatani, H., Tsuboi, A., Nagawa, F., Asano, M., Sudo, K., Sakagami, J., Sakano, H., Ijiri, T., Matsuda, Y., Suzuki, M., Yamamori, T., Iwakura Y., and Sakano, H. (2000). Mutually exclusive expression of odorant receptor transgenes. Nat Neurosci **3**(7): 687-693.

Serizawa, S., Miyamichi, K., Takeuchi, H., Yamagishi, Y., Suzuki, M., and Sakano, H. (2006). A neuronal identity code for the odorant receptor-specific and activity-dependent axon sorting. Cell **127**(5): 1057-1069.

Shepard, P. J., Choi, E. A., Lu, J., Flanagan, L. A., Hertel, K. J., and Shi, Y. (2011). Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. RNA **17**(4): 761-772.

Shi, Y., Di Giammartino, D. C., Taylor, D., Sarkeshik, A., Rice, W. J., Yates, J. R., 3rd, Frank, J., and Manley, J. L. (2009). Molecular architecture of the human pre-mRNA 3' processing complex. Mol Cell **33**(3): 365-376.

Shum, E. Y., Espinoza, J. L., Ramaiah, M., and Wilkinson, M. F. (2015). Identification of novel post-transcriptional features in olfactory receptor family mRNAs. Nucleic Acids Res **43**(19): 9314-9326.

Singer, M. S., and Shepherd, G. M. (1994). Molecular modeling of ligand-receptor interactions in the OR5 olfactory receptor. Neuroreport **5**(10): 1297-1300.

Solbu, T. T., and Holen, T. (2012). Aquaporin pathways and mucin secretion of Bowman's glands might protect the olfactory mucosa. Chem Senses **37**(1): 35-46.

Sosinsky, A., Glusman, G., and Lancet, D. (2000). The genomic structure of human olfactory receptor genes. Genomics **70**(1): 49-61.

St John, J. A., Clarris, H. J., McKeown, S., Royal, S., and Key, B. (2003). Sorting and convergence of primary olfactory axons are independent of the olfactory bulb. J Comp Neurol **464**(2): 131-140.

Strotmann, J., Conzelmann, S., Beck, A., Feinstein, P., Breer, H., and Mombaerts, P. (2000). Local permutations in the glomerular array of the mouse olfactory bulb. J Neurosci **20**(18): 6927-6938.

Strotmann, J., Hoppe, R., Conzelmann, S., Feinstein, P., Mombaerts, P., and Breer, H. (1999). Small subfamily of olfactory receptor genes: structural features, expression pattern and genomic organization. Gene **236**(2): 281-291.

Strotmann, J., Levai, O., Fleischer, J., Schwarzenbacher, K., and Breer, H. (2004). Olfactory receptor proteins in axonal processes of chemosensory neurons. J Neurosci **24**(35): 7754-7761.

Strotmann, J., Wanner, I., Helfrich, T., Beck, A., and Breer, H. (1994). Rostro-caudal patterning of receptor-expressing olfactory neurones in the rat nasal cavity. Cell Tissue Res **278**(1): 11-20.

Strotmann, J., Wanner, I., Krieger, J., Raming, K., and Breer, H. (1992). Expression of odorant receptors in spatially restricted subsets of chemosensory neurones. Neuroreport **3**(12): 1053-1056.

Sultan, M., Amstislavskiy, V., Risch, T., Schuette, M., Dokel, S., Ralser, M., Balzereit, D., Lehrach, H., and Yaspo, M. L. (2014). Influence of RNA extraction methods and library selection schemes on RNA-seq data. BMC Genomics **15**: 675.

Sun, T., Li, S., Yang, J., Yin, Y., and Ling, S. (2014). Identification of a microRNA regulator for axon guidance in the olfactory bulb of adult mice. Gene **547**(2): 319-328.

Takagaki, Y., Seipelt, R. L., Peterson, M. L., and Manley, J. L. (1996). The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell **87**(5): 941-952.

Takeuchi, H., Inokuchi, K., Aoki, M., Suto, F., Tsuboi, A., Matsuda, I., Suzuki, M., Aiba, A., Serizawa, S., Yoshihara, Y., Fujisawa, H., and Sakano, H. (2010). Sequential arrival and graded secretion of Sema3F by olfactory neuron axons specify map topography at the bulb. Cell **141**(6): 1056-1067.

Takeuchi, H., and Sakano, H. (2014). Neural map formation in the mouse olfactory system. Cell Mol Life Sci **71**(16): 3049-3057.

Tanigawa, N., Fujita, T., and Fujii, H. (2014). Oligoribonucleotide (ORN) interference-PCR (ORNi-PCR): a simple method for suppressing PCR amplification of specific DNA sequences using ORNs. PLoS One **9**(11): e113345.

Tian, B., and Manley, J. L. (2017). Alternative polyadenylation of mRNA precursors. Nat Rev Mol Cell Biol **18**(1): 18-30.

Tian, H., and Ma, M. (2004). Molecular organization of the olfactory septal organ. J Neurosci **24**(38): 8383-8390.

Treloar, H. B., Purcell, A. L., and Greer, C. A. (1999). Glomerular formation in the developing rat olfactory bulb. J Comp Neurol **413**(2): 289-304.

Tsai, L., and Barnea, G. (2014). A critical period defined by axon-targeting mechanisms in the murine olfactory bulb. Science **344**(6180): 197-200.

Vassar, R., Chao, S. K., Sitcheran, R., Nunez, J. M., Vosshall, L. B., and Axel, R. (1994). Topographic organization of sensory projections to the olfactory bulb. Cell **79**(6): 981-991.

Verhaagen, J., Oestreicher, A. B., Gispen, W. H., and Margolis, F. L. (1989). The expression of the growth associated protein B50/GAP43 in the olfactory system of neonatal and adult rats. J Neurosci **9**(2): 683-691.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature **456**(7221): 470-476.

Wang, F., Nemes, A., Mendelsohn, M., and Axel, R. (1998). Odorant receptors govern the formation of a precise topographic map. Cell **93**(1): 47-60.

Whitesides, J. G., 3rd, and LaMantia, A. S. (1996). Differential adhesion and the initial assembly of the mammalian olfactory nerve. J Comp Neurol **373**(2): 240-254.

Wienholds, E., Kloosterman, W. P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H. R., Kauppinen, S., and Plasterk, R. H. (2005). MicroRNA expression in zebrafish embryonic development. Science **309**(5732): 310-311.

Wong, S. T., Trinh, K., Hacker, B., Chan, G. C., Lowe, G., Gaggar, A., Xia, Z., Gold, G. H., and Storm, D. R. (2000). Disruption of the type III adenylyl cyclase gene leads to peripheral and behavioral anosmia in transgenic mice. Neuron **27**(3): 487-497.

Wunderlich, H. (2016). Alternativ Polyadenylierung des olfaktorischen Rezeptors mOR37a in der Wildtyp Maus. Bachelor Thesis, Karlsruhe Institute of Technology.

Young, J. M., Friedman, C., Williams, E. M., Ross, J. A., Tonnes-Priddy, L., and Trask, B. J. (2002). Different evolutionary processes shaped the mouse and human olfactory receptor gene families. Hum Mol Genet **11**(5): 535-546.

Young, J. M., Shykind, B. M., Lane, R. P., Tonnes-Priddy, L., Ross, J. A., Walker, M., Williams, E. M., and Trask, B. J. (2003). Odorant receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels. Genome Biol **4**(11): R71.

Yuen, P. S., Brooks, K. M., and Li, Y. (2001). RNA: a method to specifically inhibit PCR amplification of known members of a multigene family by degenerate primers. Nucleic Acids Res **29**(6): E31.

Zapiec, B., Bressel, O. C., Khan, M., Walz, A., and Mombaerts, P. (2016). Neuropilin-1 and the Positions of Glomeruli in the Mouse Olfactory Bulb. eNeuro **3**(5).

Zhang, H., Lee, J. Y., and Tian, B. (2005). Biased alternative polyadenylation in human tissues. Genome Biol **6**(12): R100.

Zhang, X., and Firestein, S. (2002). The olfactory receptor gene superfamily of the mouse. Nat Neurosci **5**(2): 124-133.

Zhang, Y., Ueno, Y., Liu, X. S., Buller, B., Wang, X., Chopp, M., and Zhang, Z. G. (2013). The MicroRNA-17-92 cluster enhances axonal outgrowth in embryonic cortical neurons. J Neurosci **33**(16): 6885-6894.

Zhao, J., Hyman, L., and Moore, C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. <u>Microbiol Mol Biol Rev</u> **63**(2): 405-445.

Zou, D. J., Feinstein, P., Rivers, A. L., Mathews, G. A., Kim, A., Greer, C. A., Mombaerts, P., and Firestein, S. (2004). Postnatal refinement of peripheral olfactory projections. <u>Science</u> **304**(5679): 1976-1979.

Zufall, F., and Leinders-Zufall, T. (2007). Mammalian pheromone sensing. <u>Curr Opin Neurobiol</u> **17**(4): 483-489.

6. Appendix

Abbreviations

APA	alternative polyadenylation
CDS	coding sequence
EGFP	enhanced green fluorescent protein
FISH	fluorescence *in situ* hybridization
GAP43	growth associated protein 43
GPCR	G-protein coupled receptor
GSP	gene specific primer
IRES	internal ribosomal entry site
lncRNA	long non-coding RNA
miRNA	microRNA
NUP	nested universal primer
PAS	polyadenylation signal
OB	olfactory bulb
OE	olfactory epithelium
OMP	olfactory marker protein
O/N	overnight
OR	olfactory receptor
OSN	olfactory sensory neuron
RACE	rapid amplification of cDNA ends
RT	room temperature
UPM	universal primer mix
UTR	untranslated region
v/v	volume per volume
w/v	weight per volume

List of Figures

List of Tables

PCR primers and modified oligonucleotides

All PCR primers were purchased from Eurofins Genomics (Ebersberg, Germany) and all TaqMan Probes were purchased from Biomers. net (Ulm, Germany).

Primer	Sequence (5'→3')	Usage
37A_intron_-1995F	GAGGGGTGCTGTATGGAAGA	ddPCR
37A_intron_-1822R	GTCTCCTTGCTGACAGTGGT	ddPCR
37A_3'UTR_175F	CCTTTACACCTTTGGCATCATCCA	ddPCR
37A_3'UTR_416R	TTTACCCTTGCTGGTCCCTTT	ddPCR
37A_3'UTR_3429F	GCTCTGTCTCATTCCCACCT	ddPCR
37A_3'UTR_3676R	GTGGTGTGAACTTAGGAGAA	ddPCR
37A_3'UTR_4124R	CGGACTGAACCTCTGACTCT	ddPCR (GSP)
37A_intron TaqMan Probe	HEX-TTCCTACATTCCTGAGAAGCCTGG-BHQ1	ddPCR
37A_prox TaqMan Probe	FAM-TCAACCTTAGTGGCGTCGTAGAAC-BHQ1	ddPCR
37A_dist TaqMan Probe	HEX-ACCCCTGGCTGGATTTACTCTTA-BHQ1	ddPCR
37A_short_fwd	*GCTGAGCTC*CTGTCACAGTGCAGAACTT	ddPCR (cloning)
37A_short_rev	*ACGGGTACCTTTT*TCCATTCATTTCTTTCCTTTA	ddPCR (cloning)
37A_long_rev	*ACTGGTACCTTTT*TCTTGTCAAAGCAGGAACTAG	ddPCR (cloning)
37A_3'UTR_241F	CCTTAGTGGCGTCGTAGAAC	PCR
37A_3'UTR_431R	CCTTGCTGGTCCCTTTTGC	PCR
37A_dd prox 1stpA rev2	CAGGGCTATGTGGTGAGAATC	PCR
37A_intron_-47F	GCTTTGGGAGGAAGACAATGTT	PCR
GFP_539F	CGACCACTACCAGCAGAACA	PCR
Si_mOR37A_prox_fwd	*TATGAGCTC*GGTGGGCCAGAAACACCTA	PCR/FISH/SB
Ki_mOR37A_prox_rev	*ATAGGTACC*GCATCTGTCTCCTAGTTATG	PCR/FISH/SB
Si_mOR37A_dist_fwd	*GTAGAGCTC*CTCCCAAGGCATGTACCACT	PCR/FISH/SB
Ki_mOR37A_dist_rev	*TTAGGTACC*GCATATTTAGGCTCCCACCA	PCR/FISH/SB
Ki_37A_qPCR_rev	*CTCGGTACC*AGACCTCTTGATGGATCGTG	PCR/SB
NUP	AAGCAGTGGTATCAACGCAGAGT	3'/5' RACE
UPM	Long: CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT Short: CTAATACGACTCACTATAGGGC	3'/5' RACE
37A_3'Race_GSP (GFP)	CGCCGACCACTACCAGCAGAACACC	3' RACE
37A_3'UTR_329F	TGTTCACCGTTTGCCCTGC	3' RACE nested PCR
37A_3'UTR_1480F	AGATTCTCACCACATAGCCCTG	3' RACE nested PCR
37A_3'UTR_2414F	GCTGGGGAAATGGGACTATGC	3' RACE nested PCR
37A_3'UTR_4374F	CCAAGGGGAACTGTCAGCATTCTAC	3' RACE nested PCR
37A_s4-s3_fwd	CCGTGGTCTGCAGGTTCGGACACTT	3' RACE nested PCR
37A_cds_5'Race_GSP	CCAGCCATTCTCTCCAAAGGTGTAAAGG	5' RACE
37A_5'UTR_-2291R	GATGTCACCAGCACCAAGTTCAG	5' RACE nested PCR
37C_3'UTR_1233F	GAATGGATGCCCTGGTGTAT	qPCR

| Primer	Sequence (5'→3')	Usage
37C_3'UTR_1392F	GGCTGATGTAAGAAGAGAGTGTG	qPCR
37C_3'UTR_6429F	CTACCACCCTACTCCTCCCTT	qPCR
37C_3'UTR_6592F	GAGAAATCCTGATGATGATGAAG	qPCR
37C_old_primer_F	*GGCGAGCTC*TATCAGTTAAGACCATGGAATGGAT	PCR/FISH/ 3' RACE nested PCR
37C_old_primer_R	*GCAGGTACC*GTCATACATTTTTCTCTTTGGAAGC	PCR/FISH
Si_37C_3'UTR_4760F	*ACCGAGCTC*ATCCCAGCCGAGGTGTGTAA	PCR/FISH
Ki_37C_3'UTR_5381R	*GCAGGTACC*GAGCCCTACAGAAATCACCAC	PCR/FISH
37C_GSP_1018F	AAGGCAATGCTAGGTGAGGAACAA	3' RACE
37C_3'UTR_6428F	CCTACCACCCTACTCCTCCCTTAT	3' RACE nested PCR
Olfr701_prox_qPCR_F	GCAGGAGAAGGTTTAGGGTTT	qPCR
Olfr701_prox_qPCR_R	CTCGTCTACTGTGTTGGGGA	qPCR
Olfr701_dist_qPCR_F	GACTCTCATTGGCATCATC	qPCR
Olfr701_dist_qPCR_R	ATAAAAGTAGGTATCTGGAACA	qPCR
Olfr702_qprox _1071F	CTGCCTCCCTTGAGTGTTGT	qPCR
Olfr702_qprox_1244R	AGTAAACTTCATAAATGTCCCTG	qPCR
Si_olfr701_prox_F	*TTGGAGCTC*ACCAGCCATGTGAGTTTGTG	PCR/FISH
Ki_olfr701_prox_R	*TGAGGTACC*GGAGCCATTCAGCCTGTTTG	PCR/FISH
Si_olfr701_dist_F	*ATAGAGCTC*GACTCTCATTGGCATCATCCA	PCR/FISH
Ki_olfr701_dist_R	*TAAGGTACC*CTGATTGTCACAGCACCACA	PCR/FISH
Olfr701_GSP761F	GTGCCACTTTCATGTATGTGCTGCC	3' RACE
Olfr701_GSP1265F	AGAGAATGGAGAGGACTGGGGGAAG	3' RACE
Olfr701_3'UTR_2094F	ATCAGAGGCAGGGGATAGGGAGC	3' RACE nested PCR
Olfr701_3'UTR_3153F	ACACATTTTAGCACCAGGTTTTCT	3' RACE nested PCR
Olfr701_3'UTR_ 4770F	CATAGATACTTTCACTGTTCTGGCAGC	3' RACE nested PCR
Olfr702_GSP812F	TAAGGTTGTCTCTGTGTTCTACTCA	3' RACE
Olfr702_3'UTR_1014F	CCCTTGTTCCCTTCTTTTCTCATACT	3' RACE
Olfr701_dist_qPCR_R	ATAAAAGTAGGTATCTGGAACA	3' RACE nested PCR
Olfr701_3'UTR_3903R	CTGATTGTCACAGCACCACA	3' RACE nested PCR
T3	AATTAACCCTCACTAAAGG	Colony PCR
T7	TAATACGACTCACTATAGG	Colony PCR
M13 Forward	TGTAAAACGACGGCCAGT	PCR
M13 Reverse	CAGGAAACAGCTATGACC	PCR

Southern blot protocol

Southern gel preparation:
→ 0.7% agarose (Analytical Grade, Promega) in 1x TBE
→ Incubate gel in 300ml 0.25M HCl for 5min (gentle agitation)
→ Incubate gel in 300ml 0.5M NaOH/1.5M NaCl for 30min (gentle agitation)
→ Incubate gel in 300ml 1M transfer solution for 30min (gentle agitation)

Southern transfer (downward capillary transfer):
→ Set up transfer pyramid
  → Make a stack of paper towels of 3cm in height and place a reservoir with transfer solution on top of a supporting object
  → Wet two pieces of Whatman 3MM filter paper with transfer solution and place on top of the paper towels
  → Wet the membrane (Nylon membrane, positively charged, Roche) with transfer solution and place on top of the filter paper
  → Place the gel on the membrane
  → Wet one piece of Whatman 3MM filter paper cut to the same size of the gel and use one piece of wetted Whatman 3MM paper to build a bridge between the gel and the reservoir
→ Cut parafilm to cover the top piece of Whatman 3MM paper
→ Blot for 2h
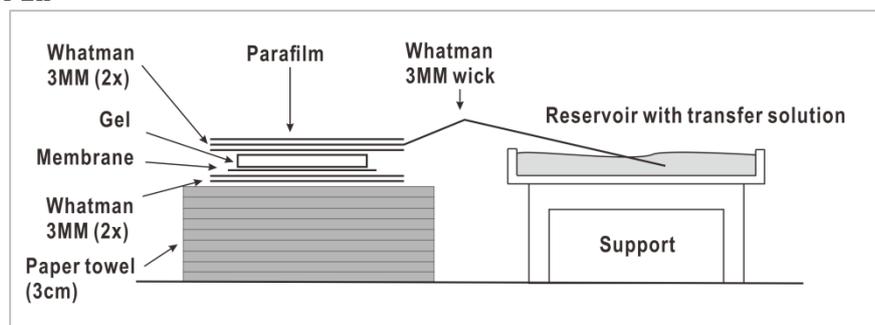
→ Disassemble transfer pyramid
→ Mark the gel size and gel slot position in membrane with a syringe needle
→ Wrap the membrane in a plastic sheet and immobilize the DNA with 70000μJ in a UV crosslinker (UV Crossliner Stratalinker 1800, Stratagene)

Hybridization:
Hybridization is performed in 50ml centrifuge tube in a hybridization oven with rotation
→ Incubate membrane in 10ml pre-hybridization solution for 1h at hyb. temp.[1]
→ Denature RNA probe in 50% formamide for 3min at 100ºC and chill on ice, and add to the membrane in pre-hybridization solution

Stringency washes:
Stringency washes are performed in 50ml centrifuge tube in a hybridization oven with rotation
→ Pre-warm wash buffers to hyb. temp.[1]
→ Wash membrane in 10ml pre-hybridization solution for 2x 10min at hyb. temp.[1]
→ Wash membrane in 10ml 0.1x SSC/0.1% SDS for 2x 15min at hyb. temp.[1]

Antibody detection:
→ Wash membrane in 50ml wash buffer for 3min (gentle agitation)
→ Incubate membrane in 50ml 1% Blocking reagent in 1x MABS for 1h (gentle agitation)
→ Incubate membrane in Anti-Digoxigenin-AP Fab fragments (1:7500 dilution in 50ml 1% Blocking reagent) for 30min (gentle agitation)
→ Wash membrane in 50ml wash buffer for 3x 10min (gentle agitation)

→ Equilibrate membrane in 50ml reaction buffer for 5min

Chemilumiscence:
→ Place membrane on a plastic sheet and incubate in 1ml CDP-Star (1:20 dilution in reaction buffer) for 5min
→ Image the membrane in Imager 600 (GE Healthcare) in chemiluminescent mode

[1]: hyb. temp. is calculated with the following formula:

$$Tm(^{o}C) = 79.8 + 18.5 \cdot (\log Na^{+}) + 58.5 \cdot (\%GC) + 11.8 \cdot (\%GC)^{2} - 820/L - 0.5 \cdot (\%formamide) - 12.5$$

Na^{+}: concentration of Na^{+} ions in moles per liter in the hybridization buffer
%GC: GC content of the probe (number between 0 and 1)
L: length of the probe
%formamide: percentage of formamide in the hybridization buffer

FISH protocol

Step	Procedure	Solution	#	Time /min	Details
Day1					
1	Fixation	4% PFA in PBS	1	10	PFA at RT
2	Wash	PBS[DT]	1	5	
3	Permeabilization	0.2M HCl	1	10	
4	Wash	PBS	1	5	
5	Inactivation of endogenous peroxidase	3% H_2O_2 in PBS	1	25	
6	Wash	PBS	1	5	
7	Acetylation	5mM acetic anhydride in 0.1M TEA	1	10	Mix well
8	Wash	PBS	1	5	
9	Hybridization	150µl per slide 75µl Formamide 37.5µl 20x SSC[DT] 3µl Torula RNA (20mg/ml) 1.5µl tRNA (10mg/ml) each RNA probe (1-3ng/µl) Denature (3min, 99ºC) 15µl 50x Denhardt's	1	O/N	In humid chamber at hyb. temp.[1] under silanized coverslips
Day2					
10	Wash	5x SSC	1	10	Remove coverslips
11	Stringency wash	50% Formamide 2x SSC	1	30	At hyb. temp.[1]
12	Stringency wash	0.2x SSC	1	60	At hyb. temp.[1]
13	Stringency wash	0.2x SSC	1	15	
14	Wash	PBST	1	5	
15	Blocking	2% Blocking Reagent in PBST (pH 7.4)	1	60	
16	Antibody incubation	150µl per slide Anti-Fluo-POD (1:1500) in Blocking Reagent	1	O/N	In humid chamber at 4ºC
Day3					
17	Wash	PBST	3	10	
18	Sensitization	150µl per slide 1mM BHR in PBST	1	10	BHR sensitive to humidity
19	Wash	PBST	3	5	3rd time w/ 15µM DETAPAC
20	Methyldopamine deposition	150µl per slide 300mM mutarot. glucose 15µM DETAPAC 5µM Alexa488-Methyldop. In 100mM Borate buffer 1µl/slide 1:1000 Glucose oxidase	1	60	In humid chamber at RT
21	Wash	PBST	3	5	
22	Inactivation of 1st peroxidase	0.1M Glycin (pH 2)	1	10	
23	Inactivation of 1st peroxidase	6% H_2O_2 in PBS	1	30	

Step	Procedure	Solution	#	Time /min	Details
24	Wash	PBS	1	5	
25	Blocking	2% Blocking Reagent in PBST (pH 7.4)	1	60	
26	Antibody incubation	150µl per slide Anti-Dig-POD (1:1500) in Blocking Reagent	1	O/N	In humid chamber at 4ºC
Day4					
27	Wash	PBST	3	5	
28	Sensitization	150µl per slide 1mM BHR in PBST	1	10	BHR sensitive to humidity
29	Wash	PBST	3	5	3rd time w/ 15µM DETAPAC
30	Methyldopamine deposition	150µl per slide 300mM mutarot. glucose 15µM DETAPAC 5µM Cy3-Methyldop. In 100mM Borate buffer 1µl/slide 1:1000 Glucose oxidase	1	60	In humid chamber at RT
31	Wash	PBST	3	5	
32	DAPI staining	150µl per slide 1:1000 in PBS	1	25	In humid chamber at RT
33	Wash	PBS	1	5	
34	Embedding	Mowiol			Dark at 4ºC

[1]: hyb. temp. is calculated with the following formula:

$$Tm(ºC) = 79.8 + 18.5 \cdot (\log Na^+) + 58.5 \cdot (\%GC) + 11.8 \cdot (\%GC)^2 - 820/L - 0.35 \cdot (\%formamide) - 25$$

Na^+: concentration of Na^+ ions in moles per liter in the hybridization buffer
%GC: GC content of the probe (number between 0 and 1)
L: length of the probe
%formamide: percentage of formamide in the hybridization buffer

Acknowledgement

I am deeply grateful to everyone whom I have had the honor and pleasure to work with during this scientific progression, and everyone who has become part of my unforgettable oversea experience. In the past years, I have learned not only some fascinating aspects of neurobiology, but also the rigorous attitude and the critical thinking necessary in scientific research.

I would like to express my gratitude to Prof. Martin Bastmeyer for involving me in the research group and for cultivating an environment that fosters independency and self-discipline.

I would also like to acknowledge Prof. Jörg Kämper for evaluation of my work as the second referee.

I am especially indebted to my supervisor Dr. Franco Weth for his invaluable professional guidance and academic advice throughout my thesis. He sets an excellent scientist example for me to follow and to strive to.

This research project is generously funded by Karlsruhe School of Optics and Photonics.

I would like to give thanks to my colleagues for the wonderful working atmosphere and their research support.

Special thanks are reserved for my family and Alex for their love and support.