5-2019

Getting at the Surface: A Promoter and Coding Sequence Characterization of an Odorant Receptor

Irena G. Parvanova

The Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Biology Commons

Recommended Citation

Parvanova, Irena G., "Getting at the Surface: A Promoter and Coding Sequence Characterization of an Odorant Receptor" (2019). CUNY Academic Works.
https://academicworks.cuny.edu/gc_etds/3242

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
Getting at the surface –
a promoter and coding sequence characterization of
an odorant receptor

by

Irena G Parvanova

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements
for the degree of Doctor of Philosophy, The City University of New York

2019
Getting at the surface –
a promoter and coding sequence characterization of
an odorant receptor
Getting at the surface –
a promoter and coding sequence characterization of an odorant receptor

by

Irena G Parvanova

This manuscript has been read and accepted for the Graduate Faculty in Biology
in satisfaction of the dissertation requirement for the degree of Doctor of
Philosophy

________________________________________
Date Dr. Cathy Savage-Dunn
Executive Officer

________________________________________
Date Dr. Paul Feinstein
Chair of Examining Committee

Supervisory Committee:

________________________________________
Date Dr. Diana Bratu

________________________________________
Date Dr. Carmen Melendez

________________________________________
Date Dr. Thomas Bozza

________________________________________
Date Dr. Duanchen Wen

THE CITY UNIVERSITY OF NEW YORK
ABSTRACT

Getting at the surface –
a promoter and coding sequence characterization of an odorant receptor

by

Irena Parvanova

Advisor: Dr. Paul Feinstein

Odorant receptors (ORs) are expressed by mature olfactory sensory neurons (OSNs), in a monogenic and monoallelic fashion (singular gene choice). It has been repeatedly published that the singular expression of ORs is a consequence of epigenetic regulation. The model invokes the depositing of epigenetic marks (off marks) that reflect transcriptional inactivity on all OR genes in immature neurons followed by the activation of one OR allele in maturing neurons. OR expression has been exclusively studied via transgenic animal models, due to the lack of cell line system that can undergo singular gene choice. Here, we used chromatin immunoprecipitation-quantitative polymerase chain reaction assay (ChIP-qPCR) with antibodies against histone modifications representing epigenetic off and on marks in mouse embryonic stem cells (mESCs) to show that the M71 OR promoter, where M71 is thoroughly studied OR gene in ours and other labs, contains only off marks in mESCs. We established a high-throughput approach that scored for the expression of CRE protein and for the conversion of fluorescence (TdTomato→GFP fluorescent shift) as a reporter in case of M71 OR activation. To attempt M71 OR expression, we administered thousands chemical compounds to the reporter mESCs, yet none led to the expression of our CRE reporter revealing the powerful repression of the M71 gene in mESCs. To further increase the opportunity for M71 expression, we transfected into these reporter cells an
M71-IRES-CRE transgene with OR gene choice enhancers, but we observed no expression of the reporter vis-à-vis random integration. In addition, a set of epigenetic chemical compounds did not induce expression of stable CRE protein.

We hypothesize that the results might be due to negative regulatory sequences used in OR genes that act as suppressor elements in any cell type including mESCs, such as the OR enhancers, OR promoter, or coding sequence. Based on this reasoning, we utilize our mESC-transgenic approach to test for the presence of “negative” DNA sequences that are specific for the OR genes: the enhancer element, the 461bp M71 OR promoter, and the 1kb M71 OR coding sequence.

ORs are G protein-coupled receptors (GPCRs) and like many GPCRs cannot express at the cell surface outside of their normal cell type. The second part of my thesis focuses on how GPCRs traffic to the cell membrane in the heterologous cell line OP6. To explore this peculiarity, we performed a structure-function study between the β2-adrenergic receptor (β2AR), which traffics well to the plasma membrane and M71 OR, which does not. Using OP6 cells, we determined that the N-terminus and the first transmembrane domain of β2AR (β2AR Nt→TM1) trafficked to the cell surface on its own, similarly to the full length of the receptor, while the M71 Nt→TM1 still did not. We eventually deleted the Nt from the β2AR Nt→TM1 protein, which initially appeared to abolish plasma membrane trafficking. A detailed analysis of this mutant revealed that it had in fact trafficked to the plasma membrane, but in “C-terminus out/N-terminus in orientation” rather than the conventionally expected “N-terminus out/C-terminus in” orientation. A re-analysis of the M71 Nt→TM1 “non-trafficking” mutant revealed that it too was in a C-terminus out/N-terminus in orientation.
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my mentor, Dr. Paul Feinstein, for accepting me in his lab and guiding me through the last seven years. While I was working on my Ph.D. dissertation, he provided endless source of knowledge and ideas. He has taught me how to approach a scientific questions and design experiments to pursue the required answers, but also how to think about a scientific problem on my own.

I would like to thank my committee members – Dr. Diana Bratu, Dr. Carmen Melendez, Dr. Shuibing Chen, and Dr. Thomas Bozza for the numerous inputs and the encouragement during the course of my studies. Diana and Carmen have been not only supportive in terms of my work, but also kind and caring through productive and difficult moments of my doctorate studies. Since Dr. Chen was not able to attend my pre-defense seminar, I would also like to thank Dr. Duancheng Wen for his willingness to help and offer his scientific input on a short notice.

I’d like to thank my lab members, in particular, Dr. Sophie Jamet and Dr. Charlotte D’Hulst, who have helped me throughout the entire course of my Ph.D. by giving me critical advices, guiding me through the process, and offering scientific and moral support. Sophie is the most positive, thoughtful, and wise scientist and person, I have ever encountered. I’d also like to thank Raena Mina, Eugene Lempert, and Alessandro Rossa, fellow students in the lab, who have helped me along the way with my studies.

Special thanks to Sergio Bernal, who have collected the majority of the GPCR trafficking data and has helped me with its analysis.
I’d also like to thank Dr. Melissa Rosso for generously offering her ChIP-PCR assay protocol for me to use, as well as, Dr. Gayathri Raghupathi for assisting with its optimization, and the entire Ortiz lab for helping me with the FACS data analysis.

Also, I’d like to thank all of the graduate students from the Hunter Biology Club and Hunter Biology department, for their endless support. Livia, Raena, Jordana, Zayd, and Joey have been helpful, caring, and extremely insightful in their suggestions. The Biology faculty and stuff have been incredibly accommodating and understanding. Dr. Pat Rockewell have been guiding and encouraging me for many years. In addition, I would like to mention my appreciation for having Razib as Genetics and Biochemistry CLT. He made the teaching experience easier by being incredibly accommodating, kind, and nearby when I had any questions.

Duygu Kayaoglu has been a close friend and roommate for an extremely long time and I couldn’t have been able to work long hours, attend conferences, or simply proceed through the day without her being there for me and sharing my experiences.

Last, but not least, I’d like to thank my family: my mom, dad, sister, aunt, Elaine, and my boyfriend Vadim for always being the best gift in my life and for offering endless amount of unconditional love and support. You are and will always be the constant that keeps me grounded and brings me back regardless given situation. I love you more than anything!
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
</tr>
<tr>
<td>CHAPTER 1. INTRODUCTION</td>
</tr>
<tr>
<td>1.1 The sense of smell</td>
</tr>
<tr>
<td>1.2 Overview of odorant receptors</td>
</tr>
<tr>
<td>1.3 Odorant receptors as G protein-coupled receptors.</td>
</tr>
<tr>
<td>1.4 G protein-coupled receptors expression in heterologous cells</td>
</tr>
<tr>
<td>1.5 Concept of choice and epigenetic regulation in the current model of OR expression.</td>
</tr>
<tr>
<td>1.6 Description of regulatory choice elements in the olfactory system</td>
</tr>
<tr>
<td>1.7 Embryonic stem cells line to study OR expression</td>
</tr>
<tr>
<td>1.8 Chemical library approach to activate M71 OR in the mESC reporter system</td>
</tr>
<tr>
<td>1.9 M71 OR mutant series to investigate the lack of expression of ORs in the mESCs</td>
</tr>
<tr>
<td>1.10 GPCR-trafficking assay in heterologous cells</td>
</tr>
<tr>
<td>CHAPTER 2. MATERIAL AND METHODS</td>
</tr>
<tr>
<td>2.1 ChIP-qPCR assay to investigate epigenetic marks on ORs in mESCs.</td>
</tr>
<tr>
<td>2.2 High-throughput screenings assay, using small chemical compound libraries.</td>
</tr>
<tr>
<td>2.3 Plasmid constructs</td>
</tr>
<tr>
<td>2.4 Transfection assays in mESC reporter and olfactory placode cells</td>
</tr>
<tr>
<td>2.5 Neomycin frameshift in mESCs by CRISPR.</td>
</tr>
<tr>
<td>2.6 Laser scanning confocal microscopy for imaging purposes</td>
</tr>
<tr>
<td>2.7 Membrane non-permeabilization antibody-staining assay to detect C-terminal GFP (in vitro)</td>
</tr>
<tr>
<td>2.8 Membrane non-permeabilization antibody-staining assay to detect C-terminal GFP (in vivo)</td>
</tr>
<tr>
<td>CHAPTER 3. Investigating chromatin histone modifications of odorant receptor promoters in mouse embryonic stem cells. Establishing a chemical screening assay to differentiate mouse embryonic stem cells into mature olfactory sensory neurons</td>
</tr>
<tr>
<td>3.1 ChIP-qPCR assay exploits the epigenetic landscape on the M71 OR gene promoter in ROSA-MTMG;M71-IRES-CRE mouse embryonic stem cells</td>
</tr>
<tr>
<td>3.2 Effect of small compound chemical libraries on the epigenetic regulation of the M71 OR promoter in mouse embryonic stem cells</td>
</tr>
<tr>
<td>CHAPTER 4. Reasons for M71 OR gene to be silenced in mouse embryonic stem cells</td>
</tr>
<tr>
<td>4.1 Expression of M71 OR transgenes in ROSA-MTMG;M71-IRES-CRE reporter system</td>
</tr>
<tr>
<td>4.2 Expression of M71 OR transgenes under antibiotic selection in ROSA-MTMG;M71-IRES-CRE reporter system</td>
</tr>
<tr>
<td>4.3 Neomycin frameshift in mESCs by CRISPR</td>
</tr>
<tr>
<td>CHAPTER 5. Investigating GPCR ability to traffic to the plasma membrane in an olfactory placode-derived cell line</td>
</tr>
<tr>
<td>5.1 In vitro filopodia assay to monitor GPCR expression in OP6 cells.</td>
</tr>
<tr>
<td>5.2 ΔNt, ΔCt and ΔNt/ΔCt β2AR FL trafficking expression on the membrane in OP6 cells</td>
</tr>
</tbody>
</table>
5.3 Transmembrane domains significance for GPCR trafficking in OP6 cells 101
5.4 TM1 and N-terminus significance for GPCR trafficking to the cell surface 102
5.5 Membrane non-permeabilization antibody-staining assay to test GPCRs orientation 108

CHAPTER 6. DISCUSSION 116
CHAPTER 7. CONCLUSIONS 125
REFERENCES 126
LIST OF FIGURES

Figure 1. Architecture of the olfactory system. 5
Figure 2. M71 OR backbone was preserved in the created series of transgenes. 7
Figure 3. G-protein cycle and GPCR signaling cascade 7
Figure 4. M71 OR and β2AR full-length trafficking in OP6 cells. 10
Figure 5. Schematic representation of type I, II, III, and IV orientation for single-membrane spanning proteins in bacteria. 14
Figure 6. OR mRNA transcription triggers negative feedback loop via unfolded protein response to prevent the expression of a second OR. 14
Figure 7. 4x21 M71 transgene expression in comparison to endogenous M71 OR expression. 18
Figure 8. 4x21 vs. 5x21 M71 OR expression in vivo. 18
Figure 9. Schematic depiction of genes integration within ROSA 26 and M71 locus in ROSA-MTMG;M71-IRES-CRE mouse line. 21
Figure 10. M71 glomeruli formation in ROSA-MTMG;M71-IRES-CRE mouse line. 22
Figure 11. Oct3/4 immunostaining in ROSA-MTMG;M71-IRES-CRE cell line. 22
Figure 12. Schematic representation of M71 OR transgenes to determine a possible regulatory sequence within the OR. 27
Figure 13. β2AR behaves as an odorant receptor when gene-targeted in an OR locus in mice. 30
Figure 14. Filopodia assay to investigate M71 OR and β2AR FL trafficking in OP6 cells. 30
Figure 15. A single-well imaging area for a 384-well plate. 39
Figure 16. Exogenous CRE expression allows for the presence of GFP-positice cells in ROSA-MTMG;M71-IRES-CRE mESCs. 43
Figure 17. ROSA-MTMG;M71-IRES-CRE cells morphology and fluorescence before and after CRE plasmid expression. 44
Figure 18. Protein expression of CRE recombinase in ROSA-MTMG;M71-IRES-CRE. 44
Figure 19. ChIP-qPCR assay. 45
Figure 20. Anti-H3K4me3 immunoprecipitation of GAPDH and M71 OR in ROSA-MTMG;M71-IRES-CRE mESCs. 48
Figure 21. Anti-H3K9me3 immunoprecipitation of β-globin and M71 OR in ROSA-MTMG;M71-IRES-CRE reporter mESCs. 49
Figure 22. Anti-H4K20me3 immunoprecipitation of β-globin and M71 OR in ROSA-MTMG;M71-IRES-CRE mESCs. 50
Figure 23. M71 OR/CRE expression allows for excision of TdTomato in ROSA26 locus and TdTomato→GFP shift in fluorescence in ROSA-MTMG;M71-IRES-CRE mESCs. 57
Figure 24. Prestwick chemical compounds library pie chart, based on the mechanisms of activity of the chemical compounds. 57
Figure 25. MicroSource chemical compounds library pie chart, based on the mechanisms of activity of the chemical compounds. 58
Figure 26. LOPAC chemical compounds library pie chart, based on the mechanisms of activity of the chemical compounds. 58
Figure 27. Control images for ROSA-MTMG;M71-IRES-CRE mESCs, taken through mCherry and GFP fluorescent filter. 59
Figure 28. Control images, post CRE transfection, of ROSA-MTMG;M71-IRES-CRE mESCs, taken through mCherry and GFP fluorescent filter. 59
Figure 29. Control images post CRE transfection and fixation of ROSA-MTMG;M71-IRES-CRE mESCs, taken through mCherry and GFP fluorescent filter.

Figure 30. A scatter plot with a bar graph of the TdTomato/GFP fluorescence ranking in Table 3 for ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 31. A 384-well plate composite image of ROSA-MTMG;M71-IRES-CRE mESCs, treated with LOPAC chemical compounds library (GFP filter).

Figure 32. A 384-well plate composite image of ROSA-MTMG;M71-IRES-CRE mESCs, treated with LOPAC chemical compounds library (mCherry filter).

Figure 33. A 384-well plate composite image of ROSA-MTMG;M71-IRES-CRE mESCs, treated with LOPAC chemical compounds library (GFP filter).

Figure 34. A 384-well plate composite image of ROSA-MTMG;M71-IRES-CRE mESCs, treated with LOPAC chemical compounds library (mCherry filter).

Figure 35. A template of the compounds positions in 384-well plate.

Figure 36. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in chemical compound treated ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 37. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in chemical compound treated ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 38. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in chemical compound treated ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 39. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in chemical compound treated ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 40. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in chemical compound treated ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 41. TdTomato and GFP fluorescence morphology after the treatment of ROSA-MTMG;M71-IRES-CRE mESCs with selected chemical compounds.

Figure 42. A bar graph, depicting the TdTomato and GFP fluorescence values for 55 chemical compounds in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 43. ROSA-MTMG;M71-IRES-CRE mESCs morphology, after the treatment by six selected chemical compounds.

Figure 44. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence expressed in compound treated ROSA-MTMG;M71-IRES-CRE mESCs (GFP fluorescence is below the established baseline of log10=3.5).

Figure 45. GFP expression in ROSA-MTMG;M71-IRES-CRE mESCs, where the GFP fluorescence levels were below the log10=3.5 cut-off baseline.

Figure 46. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in ROSA-MTMG;M71-IRES-CRE mESCs after Cayman library chemical treatment.

Figure 47. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence values in enhancer-expressing ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 48. Fluorescence and morphology epigenetic compounds treated 5x21 M71 OR ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 49. Expression of endogenous Lhx2 in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 50. TdTomato/GFP expression and morphology in ROSA-MTMG;M71-IRES-CRE mESCs before and after exogenous CRE expression.

Figure 51. Detection of CRE protein, using an anti-CRE antibody 24h post transfection in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 52. EF1α promoter transgenes, transiently in ROSA-MTMG;M71-IRES-CRE cells.
Figure 53. Quantification of the number of express green cells post EFα transgenes transfection in ROSA-MTMG;M71-IRES-CRE reporter.

Figure 54. nx21 enhancer transgenes expression in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 55. FACS quantification of GFP-positive in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 56. Puromycin kill curve to determine antibiotic concentration in ROSA-MTMG;M71-IRES-CRE mESCs

Figure 57. EF1α promoter-transgenes expression under puromycin selection in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 58. FACS quantification of number of green cells expressing EF1α- transgenes after puromycin selection in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 59. FACS quantification of number of green cells expressing EF1α-transgenes before and after puromycin selection in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 60. nx21 enhancer transgenes expression under puromycin seven-day selection in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 61. EF1α/M71 promoter hybrid transgenes expression under antibiotic selection in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 62. FACS quantification of green cells expressing EF1α/M71 promoter hybrid in ROSA-MTMG;M71-IRES-CRE mESCs.

Figure 63. Expressing GPCRs in OP6 cells filopodia assay

Figure 64. ΔNt, ΔCt, and ΔNt/ΔCt full length β2AR::GFP constructs, expressed in OP6 cells.

Figure 65. M71 OR WT Nt→TM1, M71 OR codopt Nt→TM1, and β2AR Nt→TM1 membrane expression demonstrated by a filopodia assay in OP6 cells.

Figure 66. β2AR Nt→TM1 to β2AR Nt→TM7 truncation series expression based on filopodia assay in OP6 cells.

Figure 67. M71 codopt Nt→TM1 to M71 codopt Nt→TM7 truncation series expression based on filopodia assay in OP6 cells.

Figure 68. β2AR Nt→TM1 and ΔNt β2AR→TM1 truncations expression in OP6 cells.

Figure 69. Control experiments for establishing membrane non-permealization antibody-staining assay to test GPCR orientations in OP6 cells.

Figure 70. Membrane orientation of ΔNt β2AR→TM1 truncation in OP6 cells.

Figure 71. Non-permeabilization antibody-staining assay to investigate β2AR Nt→TM1 to β2AR Nt→TM7 truncation series expression in OP6.

Figure 72. Non-permeabilization antibody-staining assay to investigate M71 codopt Nt→TM1 to M71 codopt Nt→TM7 truncation series expression in OP6 cells.

Figure 73. GPCRs expected orientations in the cells.

Figure 74. Investigating the Nt/Ct orientations of GPCR TM1s in OP6 cells.

Figure 75. Investigating the Ct::GFP orientations of β2AR Nt→TM1 in vivo.

Table 1. Schematic representation of transgenic constructs to determine a possible regulator of the M71 OR expression in ROSA-MTMG;M71-IRES-CRE mESCs.

Table 2. ROSA-MTMG;M71-IRES-CRE mouse embryonic stem cells karyotyping results.

Table 3. Log10 values of red fluorescence (TdTomato/AreaA, µm²) and green fluorescence (GFP/AreaB, in µm²).

Table 4. A sample of Cayman epigenetic chemical library compounds.

Table 5. List of DNA plasmids used in OP6 cells assays.
Table 6. M71 OR::GFP (wild type and codon optimized versions) and β2AR::GFP truncated series of plasmids.

Chart 1. Membrane non-permeabilization antibody-staining assay to investigate the extracellular presence of a C-terminus GPCR::GFP in OP6 cells.
LIST OF ABBREVIATIONS

\textbf{0x21\_M71 promoter\_M71-IRES-CRE}: M71 promoter, driving M71 odorant receptor sequence-IRES driving CRE recombinase plasmid (without nuclear localization signal)

\textbf{4x21}: four multimers olfactory enhancer element

\textbf{4x21\_M71 promoter\_M71-IRES-CRE}: four multimer olfactory enhancer element\_M71 promoter driving M71 odorant receptor-IRES driving CRE recombinase plasmid (without nuclear localization signal)

\textbf{4x21\_M71-IRES-TauCherry}: four multimer olfactory enhancer element, M71 odorant receptor sequence-IRES driving tau tagged mCherry fluorophore (transgene)

\textbf{5x21}: five multimers olfactory enhancer element

\textbf{5x21 V1rb2-IRES-taumCherry}: five multimer olfactory enhancer element, V1rb2 sequence-IRES driving tau tagged mCherry fluorophore

\textbf{7x21\_M71 promoter\_M71-IRES-CRE}: seven multimer olfactory enhancer element, M71 promoter driving M71 odorant receptor-IRES driving CRE recombinase plasmid (without nuclear localization signal)

\textbf{9x21\_M71 promoter\_CRE}: nine multimer olfactory enhancer element, M71 promoter driving CRE recombinase plasmid (without nuclear localization signal)

\textbf{9x21\_M71 promoter\_M71-IRES-CRE}: nine multimer olfactory enhancer element, M71 promoter driving M71 odorant receptor-IRES driving CRE recombinase plasmid (without nuclear localization signal)

\textbf{nx21}: n multimer olfactory enhancer element

\textbf{Adcy3}: adenylate cyclase 3

\textbf{β2AR}: β2-adrenergic receptor

\textbf{β2AR Nt\rightarrow TM1-6::GFP}: N-terminus β2-adrenergic receptor transmembrane domain 1to 6 truncation GFP fusion

\textbf{bp}: base pair

\textbf{CDS}: coding sequence

\textbf{ChIP-qPCR}: chromatin immunoprecipitation-quantitative polymerase chain reaction

\textbf{Ct}: C-terminus (protein structure)

\textbf{ΔCt β2AR}: C-terminus deletion full length β2-adrenergic receptor

\textbf{DH}: Deletion OR-Venus::MS2;OMP-H2B::mCherry mouse line

\textbf{DRD2}: type 2 dopaminergic receptor

\textbf{Ebf2}: Early B Cell Factor 2 transcription factor

\textbf{EF1α promoter}: elongation factor 1-alpha promoter

\textbf{EF1α promoter\_β2AR-IRES-CRE}: EF1α promoter driving β-galactosidase sequence-IRES driving CRE recombinase (without nuclear localization signal)

\textbf{EF1α promoter\_CRE}: EF1α promoter driving CRE recombinase (without nuclear localization signal)

\textbf{EF1α promoter\_M71-IRES-CRE}: EF1α promoter driving M71 odorant receptor sequence-IRES driving CRE recombinase (without nuclear localization signal)

\textbf{EF1α/M71 promoter full-size CRE}: EF1α and M71 promoter (full length promoter sequences) driving CRE recombinase (without nuclear localization signal)
**EF1α/M71 truncated promoter_CRE:** EF1α (full length promoter) and M71 promoter (truncated length promoter) driving CRE recombinase sequence (without nuclear localization signal)

**EF1α/M71 full-size promoter_M71-IRES-CRE:** EF1α and M71 promoter (full length promoters) driving M71 odorant receptor-IRES driving CRE recombinase sequence (without nuclear localization signal)

**EF1α/M71 truncated promoter_M71-IRES-CRE:** EF1α (full length promoter) and M71 promoter (truncated length promoter) driving M71 odorant receptor sequence-IRES driving CRE recombinase (without nuclear localization signal)

**eIF2:** eukaryotic initiation factor 2

**FC:** filopodia count

**FL:** full length CDS

**FR1α M71GFP:** 5x21_M71::GFP-IRES-tau mCherry mouse line

**FR1α M71-IRES-tauLacZ:** 5x21_M71::GFP-IRES-tau mCherry,M71-IRES-tauLacZ mouse line

**FR3 β2AR-GFP:** 5x21_β2AR::GFP-ires tau mCherry mouse line

**::GFP:** GFP fusion of truncated GPCR CDS

**GPCR:** G protein-coupled receptors

**HBB:** beta-globin gene

**HD:** homeodomain

**hiPSCs:** human induced pluripotent stem cells

**H3K4me3:** three methylation of the fourth lysine in histone 3

**H3K9me3:** three methylation of the ninth lysine in histone 3

**H4K20me3:** three methylation of the twentieth lysine in histone 4

**5-HT2:** type 2 serotonin receptor

**LBR:** lamin B receptor

**Lhx2:** LIM/homeobox 2 protein

**LSD1:** lysine-specific histone demethylase 1A

**mESCs:** mouse embryonic stem cells

**M71-IRES-CRE:** M71 odorant receptor sequence-IRES driving CRE recombinase (without nuclear localization signal) gene-targeted mouse

**M71-IRES-LacZ transgene:** M71 odorant receptor sequence-IRES driving β-galactosidase transgene

**M71-IRES-TauGFP:** M71 odorant receptor sequence-IRES driving tau-tagged GFP (gene targeted construct)

**MEF:** mouse embryonic fibroblast

**M71 promoter_CRE:** M71 promoter driving CRE recombinase plasmid (without nuclear localization signal)

**M71 OR promoter_M71-IRES-CRE:** M71 promoter driving M71 odorant receptor sequence-IRES driving CRE recombinase plasmid (without nuclear localization signal)

**M71 codopt Nt→TM1-6:** N-terminus-M71 codon optimized odorant receptor transmembrane domain 1to 6 truncations-GFP fusion

**M71 WT Nt→TM1-6:** N-terminus M71 wild type odorant receptor transmembrane domain 1to 6 truncations GFP fusion

**MOE:** main olfactory epithelium

**MOR-EG OR:** Olfactory receptor 7
MOR23: olfactory receptor 16
nATF5: activating transcription factor 5
NLS: nuclear localization sequence
Nt: N-terminus
ΔNt: N-terminus deletion
ΔNt β2AR: N-terminus deletion of β2-adrenergic full length receptor sequence
ΔNt/ΔCt β2AR: N- and C- termini deletion of β2-adrenergic full length receptor sequence
ΔNt β2AR→TM1: N-terminus deletion of β2-adrenergic receptor transmembrane domain 1 truncation with GFP fusion
OB: olfactory bulb
OCAM: olfactory cell adhesion molecule
Oct3/4: Octamer- Binding Protein 3/Octamer-Binding Protein 4
O/E: Olf1/Early B-cell factor
Olfr: olfactory receptor
OR: odorant receptor
OR1A1: Olfactory receptor 1 A1
OSN: olfactory sensory neuron
OP6: olfactory precursor cells
PERK: protein kinase R (PKR)-like endoplasmic reticulum kinase
PEV: position effect variegation
PKC: Protein Kinase C
pOG231: CMV promoter protamine-Cre recombinase vector
PTM: posttranslational modifications
Rosa-MTmg: Rosa membrane TdTomato/membrane GFP mouse
RTP1/2: receptor transporting proteins 1 and 2
TAARs: trace amine-associated receptors
TAAR4: trace amine-associated receptor 4
TdTomato→GFP: TdTomato to GFP shift in fluorescence
TM1: transmembrane 1 domain
TM1α: 5x2I/2AR::GFP Nt→TM1IrestaumCherry mouse line
UPR: unfolded protein response
VAMP::GFP: vesicle-associated membrane protein fused to GFP
V1rb2: Vomeronasal type-1 receptor
CHAPTER 1. INTRODUCTION

1.1 The sense of smell

Volatile chemicals are detected by odorant receptors (ORs), expressed by several million olfactory sensory neurons (OSNs) in mice. OSNs are arrayed in the main olfactory epithelium (MOE), located inside the nasal cavity, capable of binding odorant molecules (Buck and Axel, 1991; Zhang, 2002; Mombaerts, 2004). The mucosal layer facing the luminal side of the MOE is densely populated with ciliary extensions originating from OSNs. These cilia contain the ORs capable of binding odorant molecules (ligands). The MOE is lined with the cell bodies of OSNs, as well as supportive cells, such as sustentacular cells and two stem cell-like populations. Of those stem-like cells, globose basal cells (GBCs) are the active progenitor cells in mice, while horizontal basal cells (HBCs) are normally quiescent. HBCs become activated and differentiate upon trauma to the MOE and give rise to all major cell types in the epithelium. HBCs differentiate into 1) immature sustentacular cells, 2) mature sustentacular cells, and 3) GBCs. The regeneration of the MOE mostly lays with the GBCs that can differentiate into 1) microvillous cells and, 2) cells of Bowman’s gland, and 3) OSNs (Iwai, 2008; Fletcher, 2017; Wu, 2017). In general, between ~3000-5000 OSNs express a single allele of an OR gene. We refer to this monoallelic and monogenic expression as singular gene expression (Axel and Buck, 1991; Chess, 1994). It is currently unknown if singular gene expression relies on any of the known mechanisms of monoallelic expression.

In mice, 1400 OR genes codify odor space in contrast to 400 intact ORs genes in humans. Axons of OSNs expressing the same OR gene converge to a single location in the olfactory bulb, forming a focal point called a glomerulus. For most OSNs expressing an OR, there are two sets
of glomeruli formations located in each hemisphere of the olfactory bulb (OB) (Mombaerts, 1996; Strotmann, 2000). The OB is a neuronal structure situated in the forebrain of vertebrates and the first connection in the brain, where information received by the ORs is processed; it is the location where odorant stimuli are amplified and then deconstructed. Within the glomeruli, the OSN axons signal to the dendrites of higher order neurons, such as those of the mitral and tufted cells, which later project to the cortex and amygdala. The OB not only relays the odor analysis information to the higher centers of the brain, but also receives information from those centers. The OB contains the glomeruli for each OR, located within a broad characteristic position. The current model suggests that OSNs expressing the same OR project into the same glomeruli and find each other not via chemical cues, but by interacting with similar axons while projecting through the OB and migrating together into the same glomerulus.

Glomeruli formation depends on both OR transcriptional expression level and identity. A single base pair change within the receptor coding sequence leads to an amino acid change and alternation of the identity of the axons which shifts the projecting axons/glomeruli position in the OB (Feinstein, 2004; Feinstein and Mombaerts, 2004). The OR expression level also changes the size of the glomeruli (D’Hulst, 2016). Glomeruli for different ORs in the mouse model arise at different developmental stages, but they are first observed from E13 to E16 (Ressler, 1993; Vassar, 1993; Iwema, 2004). There are more than 1600-1800 glomeruli per bulb in the mouse brain, which is consistent with most of the 1400 ORs converging to at least one glomerulus located in the OB (Figure 1). Many of the OR lead to convergence of two glomeruli per bulb, suggesting that not all intact ORs lead to stable glomerular formation (Feinstein et al., 2004).
1.2 Overview of odorant receptors

ORs were first cloned in 1991 (Axel and Buck, 1991). These receptors elucidate the molecular basis for odor perception by binding odorant ligands and conveying the information to higher brain levels to be processed. They are found exclusively in eukaryotes, including yeast. Each OSN expresses only one OR, and does so via a singular mechanism, giving the choice of ~2800 ORs in mice and ~800 ORs in humans.

One embodiment of the mechanism of singular OR choice suggests that multiple genes are initially expressed and eventually one is stably expressed (Magklara, 2011). The convergence of like axons from OSNs into glomeruli depends on the context of the rest of the projecting axons, (Feinstein and Mombaerts, 2004). A portion of the OR genes, in all species, constitute pseudogenes, where the genes sequences are present, but they are out-of-frame; therefore, their expression either leads to a defunct OR and ultimately cell death or a nonsense protein that allows the system to express a second functional OR. Feinstein has previously showed that a functional OR, in addition to detecting odors, is also responsible for the maturation and the identity of the OSNs, proper glomeruli formation, and assistance in gene choice (Feinstein, 2004; Feinstein and Mombaerts, 2004).

ORs represent more than 4% of all genes in the mouse. They are found in clusters throughout nearly all chromosomes genome, but primary on chromosomes 2, 7, and 9 (Johnson, 2012; Malnic, 2016). Within the clusters, ORs do not have high homology amongst themselves, except for paralogous genes, such as olfr151 and olfr160 (M71 and M72 OR), which have about 96% homology in mice, where they differ by 11 out of 309 amino acids (Feinstein, 2004). However,
ORs have a similar gene structure, where every OR has at least one noncoding 5’ exon and single coding exon that often contains the 3’UTR. All OR coding sequences (CDS) are between 900-1100bp in length (Figure 2).

ORs are segregated in three classes of receptor types: Class I and II ORs and Class III - trace amine-associated receptors (TAARs) in mice. Class I represents about 10% of the ORs in mice. Axons from OSNs of Class I ORs project to glomeruli in the dorsal OB. The genes encoding for Class I ORs are located in a cluster on chromosome 7. In contrast, axons from OSNs expressing Class II ORs project both in the ventral and dorsal OB. Genes encoding for them are spread amongst multiple chromosomes. In addition, the expression of Class II ORs depends on a transcription factor called LIM/homeobox protein (Lhx2). The Class III, TAARs, are also expressed in a monoallelic fashion, by OSNs that do not express Class I or II ORs. They, analogously to ORs Class I and II, also have a single coding exon and ~1000bp CDS. In mice, TAARs are known to detect aversive odors, such as the smell of urine of carnivores. There are 15 known TAARs in mice, with a single pseudogene amongst them. TAARs are expressed in a pattern similar to ORs, except for TAAR1, which is expressed in the brain. There are six functional TAARs and two TAAR pseudogenes in humans. Like Class I and II OSNs, TAAR OSNs coalesce in a few glomeruli per OB in mice. Most of the TAARs’ glomeruli project to the dorsal OB and express olfactory cell adhesion molecule (OCAM). A subset of the TAARs express more ventrally and thus project to the ventral OB (Johnson, 2012).
Figure 1: Architecture of the olfactory system.

ORs are expressed by axons of OSNs and located in the olfactory epithelium. The axons, expressing the same OR, converge together in a focal point called glomeruli, located in the OB. The OSNs axons relay the odor information to the mitral and tufted cells, which project to higher centers of the brain, such as granule cells and the olfactory cortex (adopted from https://opentextbc.ca/anatomyandphysiology/chapter/14-1-sensory-perception)
1.3 Odorant receptors as G protein-coupled receptors

ORs and TAARs belong to the G protein-coupled receptors (GPCRs) family of proteins, also known as seven-transmembrane receptors. In mice, ORs and TAARs are enriched in the cilia, where they participate in an odor G protein-mediated transduction cascade (detecting odors). Heterotrimeric G protein alpha, beta, and gamma subunits participate in a canonical GPCR signaling pathway as delineated in the literature. G proteins include 20α subunits, five β subunits, and 12γ subunits (Figure 3) (Sathyanesan, 2013). The current model for OR transduction and odor analysis features G protein activation of Adenylate Cyclase and the production of cyclic AMP (cAMP), a secondary messenger ultimately leading to the opening of ion channels, followed by depolarization of the axons of the OSNs (Chess, 1994). However, ORs do not appear to regulate axonal maturation via cAMP dependent mechanism, but may signal through a non-G coupled, non-canonical signaling, assisting OSN formation and providing individuality of non-identified factors involved in similar interactions (Feinstein, 2004; Movahedi, 2016).

1.4 G protein-coupled receptors expression in heterologous cells

There are four main classes of GPCRs, based on conserved amino acid sequences: A, B, C, and F/S, with class A, “rhodopsin-like” receptors, being the largest one amongst them. On the other hand, based on the nature of their ligands, there are about 350 endoGPCRs (e.g. β2AR), which respond to internal signals such as peptides, hormones, and endocrine signals, and 1500 chemosensory receptors (csGPCRs) (e.g. ORs and vomeronasal receptors) that respond to external signals such as odors and pheromones. TAARs are also represented within the csGPCRs family (Vassilatis, 2003). Except for TAARs, all other csGPCRs, such as ORs, are poorly traffic to the plasma membrane when expressed by in vitro assays.
Figure 2: M71 OR backbone was preserved in the created series of transgenes.

M71 OR consists of M71 5’UTR, intron, CDS and 3’UTR, where the 3’UTR is attached to the end of the CDS.

Figure 3: G-protein cycle and GPCR signaling cascade.

Ligand binding to the GPCR induces a cascade of changes in the receptor and in the bound G-alpha subunit of the G-protein. GTP-bound G-alpha assumes conformation that has low affinity for the GPCR as well as the Beta-Gamma complex, so all three dissociate from one another. The GTP-binding site of G-alpha is also capable of hydrolyzing GTP to GDP. Upon GTP hydrolysis the G-alpha subunit is considered inactive and once again may bind G-beta/gamma and form a heterotrimeric G-protein, which will often be found associated with an unliganded GPCR (adopted from https://en.wikipedia.org/wiki/G_protein-coupled_receptor).
Even though ORs have been characterized for a few decades now, only a small number of their ligands have been identified, due to their inefficient trafficking to the plasma membrane with *in vitro* systems. By contrast, β2AR is a well-studied GPCR, because unlike csGPCRs, it localizes to the plasma membrane in all cell lines. For that reason, β2AR is used as a surrogate to study ORs as a representative of GPCRs in mice. Remarkably, when β2AR is expressed from an OR locus in mice, it fulfills all OR specific functions *in vivo*: permitting the maturation of the OSNs, leading to axon guidance and identity, and assisting in gene choice and glomeruli formation (Feinstein, 2004). Recently, it has been shown that other non-ORs, such as Mcr4 and Drd1, can also direct OR specific functions, albeit at a reduced efficiency (Katidou, 2018).

β2AR is not the only endoGPCR that can traffic to the cell surface in heterologous cells. The type 2 serotonin (5-HT2) and the type 2 dopaminergic receptor (DRD2) also localize on the cell surface when expressed in heterologous cell lines. Due to its ability to robustly traffic to the membrane in cells, we have been using β2AR to study GPCR trafficking in an attempt to understand why β2AR protein readily localizes on the cell surface, while ORs do not.

We have previously described the ability of β2AR to traffic to the plasma membrane in comparison to the M71 OR, in three previous research articles from our lab. These publications relied on a plasma membrane trafficking assay, where plasmids encoding a GPCR green fluorescent protein fusion (GPCR::GFP) were transfected into cells from the olfactory-like cell line OP6 (olfactory placode origin) (Bubnell, 2013; Bubnell, 2015; Jamet, 2015). The olfactory progenitor line are temperature sensitive clonal cells, derived from E10 olfactory placode. The cells are locked in a developmental stage prior to mature olfactory neurons and have a distinctive
morphology with numerous membrane extensions called filopodia.

Our assay reveals that plasma membrane trafficking of GPCR::GFP fusion protein correlates with GFP fluorescence in the filopodia. Thus, one read-out of our trafficking assay is the number of the GFP-labeled filopodia extensions, which are always present in the OP6 cells, but only GFP-labeled when the fluorescently tagged protein is localized on the cells surface (Bubnell, 2013). For instance, when β2AR::GFP traffics to the membrane, many filopodia are observed. In comparison, almost no filopodia are observed when the OR, M71::GFP is transfected in OP6 cells (Figure 4) (Jamet, 2015).

Based on the literature, we knew that the first transmembrane domain (TM1), as well as the N-terminus, are essential for GPCR trafficking to the plasma membrane (Alken, 2002; Köchl, 2002; Wicher, 2008). In bacteria, there are four types of single-transmembrane spanning receptors that are classified based on the location of the N- and the C-termini in with respect to the plasma membrane. Type I receptors are targeted into the membrane by cleavable sequences and anchored with the N-terminus on the extracellular face of the membrane and the C-terminus on the intracellular face of the membrane. Type II receptors have the opposite orientation, with the C-terminus projecting outside of the cell membrane. Type III receptors are also called “type I without cleaved signal,” with a shorter N-terminal domain. Finally, type IV receptors are similar to type II receptors in orientation, but with a larger N-terminal domain (Figure 5) (Spiess, 1995).
The established filopodia assay is used to account for trafficking vs. non-trafficking GPCRs::GFP, based on the presence of GFP-labeled filopodia as result of the GPCR expression on the cell surface. Filopodia are always a part of the OP6 morphology, but only visible when the expressed fluorescent protein traffics to the cell surface (Jamet, at el, 2015). 

A. M71::GFP FL fusion protein, where the C-terminus of the protein was tagged with GFP and transfected in OP6 cells. The protein was trapped in the endoplasmic reticulum and not able to traffic the cell membrane. A’. CellMask Red staining (plasma membrane stain) showed GFP-labeled filopodia in the OP6 cells. Schematic structure of the GPCR – M71 OR FL, fused to GFP (M71 OR FL::GFP). B. mβ2AR FL::GFP was transfected in OP6 cells. Unlike M71 OR, β2AR trafficked to the cell surface; therefore, gGFP-labeled filopodia were observed. B’. CellMask Red was used to co-label the filopodia, stained by mβ2AR::GFP. Schematic representation of mβ2AR::GFP.
1.5 Concept of choice and epigenetic regulation in the current model of OR expression

With GPCRs being a potential therapeutic targets in cancer, diabetes, heart failure treatment, metabolic disease, and other diseases (Desimine, 2018; Ma, 2018; Sloop, 2018; Sun, 2018), it is essential to understand the mechanism by which they are expressed. All csGPCRs are regulated in a singular fashion. Singular mechanism of expression of ORs has been studied for a long time, using transgenic mice. Recently, single cell sequencing has been implemented to gain insights into singular gene choice. These data suggest that multiple ORs can be simultaneously expressed at mRNA level, but that simultaneous expression only occurs in immature neurons and not in mature OSNs. This model suggests that multiple OR mRNAs are initially expressed, but a single one eventually is transcribed and translated (Hanchate, 2015; Monahan, 2015; Tan, 2015).

Analysis of several mice knockouts out for particular genes suggest that for an immature OSN to develop into a mature one and express an OR, lamin B receptor (LBR) is needed to assist with the nuclear reorganization of OR loci. It was found that ORs, which are mainly found in scattered clusters, are re-located to areas of highly repressed constitutive heterochromatin within the nucleus. The presence of LBR protein in OSNs leads to the maturation of the neurons by stabilizing the expression of the chosen OR and re-locating it from heterochromatin where all of the rest ORs remain. If LBR is deleted in mice, the OR singular expression by OSNs is negatively influenced, which allows multiple ORs to be chosen for expression simultaneously. The activation of one OR leads to the expression of a single OR, while the rest of the ORs remain silent (Clowney, 2012).
To further ensure that all OR genes remain transcriptionally silent until one or several are chosen for expression, OR genes are thought to be kept silenced through epigenetic modifications (Magklara, 2011). Examples of epigenetic modifications are DNA methylation and histone posttranslational modifications (PTMs). DNA methylation is correlated with gene imprinting through the maintenance of silenced genes, but there are exceptions (Roux-Rouque, 2000; Reik, 2001; McCarthy, 2009). Secondly, histone modifications include a multitude of PTMs, such as histone methylation, acetylation, phosphorylation, ubiquitylation, and SUMOylation (Berger, 2007). The PTMs are reversible and have complex impacts on gene expression. Marks of transcription initiation and repression are mainly located in the promoter regions and transcription start sites of the genes.

The addition of three methyl groups to the fourth lysine of histone 3 (H3K4me3), for instance, is a hallmark modification of transcription initiation. Guenther et al. showed that over 75% of the genes in human embryonic stem cells (ESCs) are marked with H3K4me3, even if not expressed. In comparison, genes located in clusters on the chromosomes, such as ORs for instance, completely lack that mark. Consistent with the lack of initiation marks, Magklara et al. showed that OR-encoding genes carry constitutive heterochromatin hallmarks, H3K9me3 and H4K20me3, in both the liver and MOE.

Magklara et al. interpreted the repressive marks on the OR loci as a mechanism to ensure the repression in precursor OSNs and the subsequent de-repression of one or more ORs in mature OSNs. Therefore, the presence of inactive histone marks (H3K9me3 and H4K20me3) on ORs is understood as silencing of all OR loci until one or more alleles are de-repressed. It is believed that
H3K9me3 is added to ORs followed by H4K20me3 marks. Once the negative marks are removed from an OR allele, it is expressed and receives transcriptionally active epigenetic marks (Magklara, 2011).

Hence, epigenetic switch mechanisms are believed to ensure singularity. Lysine-specific histone demethylase 1A (LSD1) is believed to play a role in singular OR expression where it acts on both intermediate H3K9me2 and H3K4me2 methylation marks in genes. By demethylating either mark, LSD1 is able to push the gene expression in either direction – transcription or repression. If the gene needs to be activated, LSD1 demethylates H3K9me2, which results in transcription activation. On the contrary, if LSD1 demethylates H3K4me2 to H3K4me1 or H3K4me0, the transcription of the gene is repressed (Lyons, 2014; Markenscoff-Papadimitriou, 2014).

The mechanism of OR gene expression is believed to require a negative feedback loop on LSD1 to prevent the de-repression of any additional OR gene. This LSD1 mechanism requires that a OR protein translated in the ER and triggers unfolded protein response (UPR) pathway (Dalton and Lomvardas, 2015; Monahan and Lomvardas, 2015). In this model, an expressed OR protein causes stress in the endoplasmic reticulum (ER) and activates PERK kinase which phosphorylates eukaryotic initiation factor 2 (eIF2). Phosphorylated eIF2 increases the translation of the nuclear activating transcription factor 5 (nATF5), which in turn induces the expression of adenylate cyclase 3 (Adcy3) and OR chaperones (receptor transporting proteins, RTP1/2). An OR and Adcy3 expression leads to downregulation of the LSD1 and relieves the stress in the ER, ensuring the stable expression of the chosen OR (Figure 6) (Lewcock and Reed, 2004; Shykind, 2004; Serizawa, 2005; Dalton, 2013; Lyons, 2013).
Figure 5. Schematic representation of type I, II, III, and IV orientation for single-membrane spanning proteins in bacteria.

Depiction of four different orientations of single-transmembrane proteins. For instance, the N-terminus of the protein can be oriented outside, while the C-terminus is inside of the bacterial cell. On the contrary, there is a possibility for the N-terminus of the protein to be inside the cell, while the C-terminus is outside of the cell surface. Both conformations exist in bacteria (adopted from Spiess, 1995).

Figure 6. OR mRNA transcription triggers negative feedback loop via the unfolded protein response (UPR) to prevent the expression of a second OR.

According to the mechanism of expression of ORs: 1. LSD1 activity leads to an OR transcription; 2. The unfolded protein response (UPR) leads to the eif2A phosphorylation by PERK; 3. ATF5 is translated; 4. ATF5 activates Adcy3 expression. 5. ADCY3 represses LSD1 activity. 6. ATF5 activates Rtp1. 7. RTP1 inhibits further UPR activation (adopted from Dalton, 2015).
Remarkably, other publications revealed that in the MOE, TAAR genes, Class III ORs, not only lack active marks, but also the inactive marks typically associated with the other two OR gene classes prior to gene expression (Johnson, 2012). The lack of both activating and repressing marks on TAARs begs the question if epigenetic repression is a main regulatory factor of singular OR expression and if other unifying mechanisms of regulations exist to explain monoallelic expression of all ORs, including TAARs.

1.6 Description of regulatory choice elements in the olfactory system

Singularity is clearly regulated by multiple enhancer elements that lay within close proximity of OR promoters in cis- (Vassalli, 2002; Bozza, 2008), while some elements such as the P- and H-elements can lay over 20 kilobases away from the transcription start sites (TSSs) (Serizawa and Vassalli 2011). Deletion of these elements affects the loci which they regulate (Khan, 2011). P- and H-elements are also long distance acting olfactory enhancers for multiple OR genes. By a contrast, a 161bp sequence upstream of the transcription start site of M71 OR primarily regulates its expression. This 161bp contains two homeodomain (HD) sites and an Olf1/Early B-cell factor (O/E) site that have been shown to be necessary and sufficient to drive M71-like expression of an M71-IRES-LacZ transgene inserted in the M71 locus in mice (Rothman, 2005). Many OR promoters contain the HD generic motif of TAATXX, but HD of the P-element, H-element share 13 conserved base pairs (bp) (for instance, the conserved motif sequence often is TTTTTAATGA). If the conserved sequences are added upstream of an OR promoter, then they increase the number of neurons that chooses the OR promoter. For instance, if a HD from the P-element is multimerized and cloned upstream of the MOR23 gene (Olfr16), a dramatic increase is observed in the number of MOR23 transgene-expressing OSNs (Vassalli, 2002).
Similar results are observed when an OR transgene, such as M71 OR, containing a HD multimerized sequence of the H-element, is integrated into the mouse genome. The number of M71 expressing OSNs in those transgenic mice is significantly amplified. Based on our lab’s results, adding more H-element multimers further increases the probability for the OR gene to be expressed \textit{in vivo}; each additional copy in the multimer increases the potency of the enhancer. By introducing enhancer elements derived from the H-choice element, we can increase the probability for a single OR to be chosen by many orders of magnitude, which is different than simply increasing the number of the mRNA molecules for that OR. The higher probability of choice results in higher expression of the OR, because the OR has been chosen more frequently. A single OR is typically expressed in 0.1% of the OSNs in average, which equals about 10,000 OSNs, expressing the same ORs. With the presence of enhancer element, we have achieved the expression of the same OR by millions of OSN (D'Hulst, 2016).

In order for the HD-elements to effect the OR expression, they likely bind the appropriate transcription factors. Lhx2 is a transcription factor thought to bind the HD elements and contributes to the differentiation of OSNs expressing Class I ORs. Zhang et al. have shown that Lhx2 influences the OR expression and maturation of the OSNs (Zhang, 2016; Monahan, 2017). Emx2 is another transcription factor, which possibly binds HDs, but its function in OSNs is unclear. Olf/EBF transcription factors bind the O/E site of an OR promoter.
In a mouse line, a four times multimerized H-element HD was integrated in a M71 transgene, where $4xH = 4x21bp$ (21bp=the conserved 13bp + flanking sequences) (D’Hulst, 2016). The random integration of the transgene in the mouse allows the increased expression of the M71 OR in the epithelium. Simultaneously, an $M71$-IRES-$TauGFP$ reporter was gene-targeted to the endogenous M71 locus of a separate mouse line to monitor endogenous M71 OR expression. We used a CAP-independent Internal Ribosome Entry Site (IRES), allowing bicistronic expression of multiple genes, where a single mRNA is transcribed and translated simultaneously through CAP dependent and CAP-independent mechanisms, resulting in two different peptides (Jang, 1988; Pelletier, 1988).

We visualized the two M71 expressing transgenes in the same mouse: $4x21$ M71 OR allele ($4x21\_M71$-IRES-$TauCherry$) and endogenous M71 allele ($M71$-IRES-$TauGFP$). The high number of red fluorescent OSNs in the overlay image indicates the increased expression of M71 OR in the presence of the $4x21$ enhancer in comparison to the single green OSN. Based on the number of OSNs expressing the M71 OR in the $4x21\_M71$-IRES-$TauCherry$;$M71$-IRES-$TauGFP$ animals in comparison to the $M71$-IRES-$TauGFP$ animals, we concluded that the presence of $4x21$ enhancer significantly increases the expression of M71 (Figure 7). The $4x21$ enhancer is the lowest copy-number multimer tested in animals that increases the number of M71 OSNs. The number of OSNs expressing the same OR is also increased and the resultant glomeruli are typically bigger. Eventually, a glomerulus saturates by the number of axons and in some cases additional glomeruli are formed. Based on the size of the glomeruli and the presence of additional glomeruli, $5x21$ enhancer gives rise to even more M71 OSNs and axons (Figure 8).
Figure 7. 4x21 M71 transgene comparison to endogenous M71 OR expression in mice.

In a single plane confocal image of the main olfactory epithelium in mice, the red fluorescence panel represents the increase of the M71 OR expression due to the presence of a 4x21 M71-IRES-TauCherry transgene in vivo. The green panel shows a single green axon from the expression of endogenous M71 OR. The green arrow focuses on that single endogenous M71 OSN. The Overlay image clearly displays the increased expression of the M71 OR due to the presence of 4x21 M71 transgene, which is represented by the high presence of red fluorescence in addition to a single green endogenous M71 OR. Scale bar = 100µm.

Figure 8. 4x21 vs. 5x21 M71 OR transgenes expression in vivo.

The increased number of glomeruli represents the increased probability of M71 OR expression in 5x21 in comparison to 4x21 M71 OR transgenes in mice. Scale bar = 500µm.
Markenscoff-Papadimitriou et al. have gone on to identify additional to the H- and P-like enhancer elements. However, these elements were mostly defined by an epigenetic signature. The Lomvardas lab has now identified over 60 enhancer elements referred to as “Greek islands.” According to the authors, these particular Greek Islands interact with each other through chromatin looping in a trans-conformation, activating and stabilizing the expression of a single OR out of the total OR pool (Serizawa, 2003; Lomvardas, 2006; Khan, 2011; Clowney, 2012; Markenscoff-Papadimitriou, 2014). Recent work from the Lomvardas lab claims that LHX2 forms a dimer with EBF1 (also known as O/E1) transcription factor; the dimer binds the “Greek islands,” allowing the expression of a single OR (Le Gross, 2016; Lomvardas, 2016; Monahan, 2017; Dekker, 2017).

1.7 Embryonic stem cell reporter line to study OR expression

To study OR expression, we attempted to develop an in vitro assay to differentiate embryonic stem cells into OSNs. Currently, ORs gene expression cannot be studied outside of olfactory neurons. Even more remarkable, OR proteins cannot be expressed in heterologous cells, because they remain stuck in the ER and thus do not traffic to the plasma membrane. Because of these difficulties to exogenously express ORs, we attempted to generate expression of an OR through the mechanisms of OR gene choice. For this purpose, we developed a mouse embryonic stem cells (mESCs) reporter line from blastocysts of a mouse which is a cross between ROSA membrane TdTomato/membrane GFP mouse (ROSA-MTMG) and our M71-IRES-CRE gene-targeted mouse (M71-IRES-CRE) (Bryja, 2006; Li, 2004), which resulted in the ROSA-MTMG;M71-IRES-CRE line.
In the *ROSA-MTMG;M71-IRES-CRE* mice, a reporter has been targeted to the ROSA 26 locus that consists of the ROSA 26 and actin promoters driving a floxed (flanked by loxP sites) membrane bound red fluorophore, TdTomato (a tandem dimer of two Tomato copies), followed by the coding region of green fluorescent protein (GFP). LoxP sites, in the same orientation, flank the TdTomato sequence. The ROSA 26 locus shows broad expression across most cell types in mice including olfactory neurons. On a different chromosome, the M71 OR locus was gene-targeted with the IRES-CRE sequences (Cre recombinase lacks nuclear localization sequence, NLS). This version of the CRE enzyme is shown to be tightly regulated and to have a minimum level of leaky expression. The activation of the M71 gene causes the expression of M71 OR and CRE proteins, resulting in loxP-mediated excision of TdTomato and subsequent expression of membrane-bound GFP, in the *ROSA-MTMG;M71-IRES-CRE* animals. This event causes cells to switch from red to green fluorescence (Figure 9).

Prior to deriving the mESC line, we examined the proper expression of M71 OR in mice. As expected, a single green glomerulus of axons from OSNs expressing M71 is observed in each hemisphere of the olfactory bulb of a gene-targeted animal. Notably, this shows that CRE expression does not interfere with the expected expression of M71 (Figure 10) (Chen, 2006).

To confirm that the new reporter line is in fact mESC line, we identified the ESC marker protein Oct3/4 with a primary antibody (Figure 11). In addition to the antigen being present our mESC reporter colonies had a globular shape and immortal, consistent with what is known about ESC lines.
Figure 9. Schematic representation of gene integration within ROSA 26 and M71 locus in *ROSA-MTMG;M71-IRES-CRE* mouse line.

LoxP-TdTomato-loxP-GFP construct is gene targeted into the ROSA 26 locus on chromosome 6 in *ROSA-MTMG;M71-IRES-CRE* mouse line. Its expression is driven by ROSA26 and actin promoters. IRES-CRE knock-in is inserted in the M71 OR locus on chromosome 9 in the *ROSA-MTMG;M71-IRES-CRE* mouse line.
M71 OR axons co-converge into GFP-expressing glomeruli in the left and the right bulb of ROSA-MT
g;M71-IRES-CRE mouse line. The olfactory bulb is red, due to the presence of TdTomato in the ROSA 26 locus. A single M71 corresponding green glomerulus is observed in each bulb, due to the GFP axons, which result of the M71 OR-IRES-CRE expression. Scale bar = 100µm.

Oct3/4, an embryonic stem cell marker, immunofluorescent staining of blastocyst-derived ROSA-MT
g;M71-IRES-CRE mESCs. Red staining shows the presence of TdTomato at the membrane of the ROSA-MT
g;M71-IRES-CRE mESCs, when no recombination event has occurred, while green fluorescence accounts for the secondary antibody staining, demonstrating the presence of Oct3/4 in the mESCs. Scale bars = 20µm (the cells were extracted by Eugene Lempbert).
1.8 Chemical library approach to activate M71 OR in mESC reporter system

We used the newly created mESC reporter to set up a high-throughput assay, where we attempted to force M71 OR expression. Although ORs have been characterized for a few decades now, only a small number of their ligands have been identified. By characterizing the epigenetic landscape of OR genes, the Lomvardas lab has outlined a model of OR expression, according to which all 1400 ORs are epigenetically silenced by H3K9me3 and H4K20me3 negative marks at the progenitor level (basal cells), until a single OR is chosen to express. Thus, the singular OR allele expressed becomes devoid of the epigenetic repressive marks and transcription initiation ones are now present in the promoter region. An example of a positive epigenetic mark is H3K4me3. OR genes should only express in the olfactory epithelium, therefore, H3K4me3 epigenetic mark should not be present at the OR promoter areas anywhere else in the genome. In support of the absence of H3K4me3 marks an analysis of all genes in hESCs found OR promoters to be devoid of such marks (Guenther, 2007; Magklara, 2011). Since OR genes in hESCs lack H3K4me3 transcription initiation marks, we were interested to confirm this result in mESCs reporter line. We decided to further investigate the epigenetic landscape of an M71 OR in mESCs, by testing if the so-called basal cell specific repressive marks (H3K9me3 and H4K20me3) are present. Using the \textit{ROSA-MTMG;M71-IRES-CRE} mouse line reporter cells, we optimized a chromatin immunoprecipitation-quantitative PCR assay (ChIP-qPCR), which allowed us to investigate the epigenetic regulations of the M71 OR promoter (Rosso, 2015).

The results revealed that the M71 promoter has H3K9me3 and H4K20me3 and an absence of H3K4me3 marks in mESCs. These results led us to consider an experiment, where we would
attempt to force the M71 OR expression by interfering with its repressive marks in the *in vitro* reporter. A development of a successful *in vitro* assay would have allowed to obtain OSNs expressing ORs. To interfere with the epigenetic regulations of the M71 OR promoter, we utilized a high-throughput small compound chemical libraries and epigenetic chemical library, where we administered approximately 4860 small chemical compounds and 175 epigenetic compounds, in two concentrations to *ROSA-MTMG;M71-IRES-CRE* mESCs. In case of M71 OR expression in the mESCs, CRE would excise the TdTomato sequence, allowing for GFP membrane expression. Therefore, if any chemical compound de-repressed M71 OR by removing the negative epigenetic marks of its promoter, the read-out would be a TdTomato→GFP shift.

### 1.9 M71 OR mutant series to investigate the lack of expression of ORs in mESCs

To circumvent the fact that a single OR out of 1400 ORs in mice was targeted for expression, we also applied MouSensor technology, where we used M71 transgene, containing a strong enhancer element (D’Hulst, 2016). The transgene expression was followed by administering of 175 epigenetic compounds to the reporter system. To our surprise, no convincing reporter expression was observed in any of our experiments.

Perhaps the lack of expression of M71 was due to the absence of specific choice molecules in the mESCs. Another strong possibility could be the presence of a negative regulatory sequence within the transgene itself that did not allow our chemical high-throughput assay to work.

To study the effect of a possible “negative” regulatory sequence over the M71 OR expression, we created a series of transgenes to swap out M71 enhancer, promoter, and coding sequences.
with non-OR sequences. In these new transgenes, we utilized MouSensor technology, where
\( nx21 \) enhancer \( (n=0,5,6,7,8, \text{ and } 9) \) was used to increase the probability for M71 OR expression.
Also, the OR promoter was substituted with a heterologous promoter, EF1\( \alpha \), which is often used in ESC stable cell lines and/or the OR coding sequence (CDS) was either omitted or replaced with the \( \beta2AR \) (Figure 2 and 12) (Table 1) (Wang, 2008). For all transgenes, we preserved the M71 backbone: M71 5’UTR, introns, 3’UTR, and polyA tail and the CRE recombinase. Our modified transgenes were integrated without selection into our mESC reporter cell line and green fluorescence was followed by confocal microscopy imaging and FACS. Due to the regrettably small number of green cells after the transgenes integration without selection, we attempted to increase the frequency of observed events creating stable cell lines for the transgenes.

One of the main reasons for the lack of an \textit{in vitro} assay to study and express ORs is due to their inefficient trafficking to the cell membrane in heterologous cells. When the Nt of the OR proteins is modified by a Flag-RHO tag (Shephard, 2013), then a small amount of the protein traffics to the cell surface, while the majority of it remains in the endoplasmic reticulum.
Table 1. Schematic representation of transgenic constructs to determine a possible regulator of the M71 OR expression in ROSA-MTMG;M71-IRES-CRE mESCs.

The name of the construct, enhancer element, promoter type, 5’UTR and intron type, coding sequence type, and cistron sequence are noted in the table. *Multiple clones are listed per transgene, due to the need for re-sequencing of the nx21 enhancer clones, as they tend to resolve into fewer nx21 repeats after re-transformation.
Figure 12. Schematic representation of M71 OR transgenes to determine a possible regulatory sequence within the OR.

To investigate the presence of a regulatory sequence within the M71 OR, we examined the role of the 0x2I, 5x2I-9x2I enhancer elements, M71 OR promoter (by keeping or replacing it with EF1α promoter), and M71 coding sequence (by keeping, omitting, or replacing it) in the mESC reporter.
1.10 GPCR-trafficking assay in heterologous cells

The lack of an OR in vitro assay is the fact that similarly to many GPCRs, ORs do not efficiently traffic to the cell membrane in heterologous cells. An insignificant amount of the protein traffics to the cell surface, while the majority of it remains in the endoplasmic reticulum. ORs represent the largest family of GPCRs, also known as seven-transmembrane (TM) domains receptors. In comparison to ORs, β2-adrenergic receptor (β2AR), a canonical GPCR, traffics to the cell surface in multiple cell lines, making it a good candidate to study GPCR trafficking. In addition, similarly to an OR, β2AR is expressed in an OR locus in mice, β2AR is expressed by OSNs, helps with the maturation of the neurons, and guides OSNs to glomeruli in the olfactory bulb in vivo (Feinstein, 2004).

To contrast the divergent abilities of the M71 OR and β2AR to traffic to the plasma membrane in cells, we used a trafficking assay, where plasmids expressing M71::GFP, β2AR::GFP, or chimeric::GFP fusion proteins of interest were introduced into OP6 cells, a neuronal precursor cell line (supplied by Jane Roskams) (Bubnell, 2013; Bubnell, 2015; Jamet, 2015). These olfactory progenitor cells, locked in a stage prior to mature OSNs, have a distinctive morphology with many membrane/filopodia extensions, where the GPCR::GFP fluorescence is robust. Using the OP6 filopodia, we monitored GPCR::GFP fusion proteins expression on the plasma membrane, because the GPCR trafficking to the membrane correlates with the GFP fluorescence found in the filopodia (Bubnell, 2013). Thus, one measurement of trafficking is the number of GFP-labeled filopodia extensions after GPCR::GFP transfection in OP6 cells. As previously mentioned, filopodia are always present in the OP6 cells, but only labeled with GFP when the fluorescently tagged protein traffics to the cells surface (Jamet, 2015) (Figure 4 and 14).
To understand why β2AR traffics in cells, while M71 OR does not, we focused on the role of the transmembrane domains (TMs). We created truncation series of β2AR::GFP and M71::GFP, where we fused the Ct portions of TM1 to TM5, TM1 to TM3, and TM1 with GFP to study their role in trafficking: β2AR Nt→TM5::GFP, β2AR Nt→TM3::GFP, and β2AR Nt→TM1::GFP. These particular truncations were designed based on the assumption that TM5, TM3, and TM1::GFP fusions would all have their cytosolically located GFP tag: N-terminus out/C-terminus in (Nt out/Ct in) orientation in the plasma membrane. We followed up this truncation series with β2AR Nt→TM6, β2AR Nt→TM4 and β2AR Nt→TM2 that should have an extracellularly located GFP tag.

From all truncations, β2AR Nt→TM1 robustly expressed on the cell surface, which prompt further investigation in the involvement of TM1 in GPCR trafficking in heterologous cells. We also went on to modify the Nt, by Nt deletion (∆Nt) β2AR Nt→TM1, which radically reduced the number of GFP-labeled filopodia in OP6 cells, while ∆Nt in β2AR FL was not affected. It is thought that the TM1 domain alone is the “signal domain” necessary for plasma membrane targeting. So, why did not ∆Nt β2AR Nt→TM1 traffic to the membrane? We guessed that maybe it did in a different orientation with the Ct localized outside of the cell surface. ORs are in Nt in/Ct out orientation in the cell membrane in Drosophila melanogaster (Wicher, 2008). In order to check if ∆Nt β2AR Nt→TM1 was in Nt inside of the cell/Ct outside of the cell orientation, we established an immunostaining “non-permeabilized” membrane assay, where we introduced ∆Nt β2AR→TM1 plasmid in the OP6 cells (Zhuang, 2008) and probed for the presence of GFP with an antibody. This non-permeabilization analysis set forth a large series of experiments that revealed the role of TM1 in proper OR trafficking to the plasma membrane.
A. M71 OR is gene targeted with IRES-LacZ sequence. The presence of the OR is detected via ex vivo X-gal staining. M71 ORs are expressed dorsally in the mouse olfactory bulb. The OSNs, expressing the ORs, coalesce together in glomeruli. B. If β2AR replaces the M71 OR coding sequence in its endogenous locus, β2AR behaves as an OR. Since the β2AR in the OR locus is also gene-targeted by IRES-LacZ knock-in, it can be detected by X-gal staining. The β2AR expresses dorsally, similarly to M71 OR, with a lower number of cells than M71 ORs. Similarly to an OR, β2AR neurons form glomeruli (Feinstein et al, 2004). Scale bar = 500µm (Vassalli, 2002).

We established a filopodia assay to investigate GPCRs trafficking in OP6 cells (FC=filopodia count). A. M71 OR FL was expressed in OP6 cells. The OR protein did not traffic to the plasma membrane, since no GFP-labeled filopodia were observed. The lack of GFP-labeled filopodia extensions are represented in the zoomed-in image and by the red color-coding of the M71 OR panel. B. β2AR FL was transfected in OP6 cells. The β2AR protein was localized on the cellular membrane, depicted by the large number of GFP-labeled filopodia in the OP6 cells. The zoomed-in image emphasizes the large number of GFP-labeled filopodia. Scale bar = 20µm.
CHAPTER 2. MATERIAL AND METHODS

2.1 ChIP-qPCR assay to investigate epigenetic marks on ORs in mESCs

*ROSA-MT>M71-IRES-CRE* mESCs were grown on 10cm plates (100x20 mm Tissue Culture Dish, Corning). 12-16 10cm plates, at confluency, provided around 30-45×10⁶ cells. The cells were fixed with 1% PFA (PFA, Electron Microscopy Sciences) for 10min at 37°C. The PFA was added to the cell growth media at a final concentration of 1% PFA, taking into account the combined volumes of media and the fixating agent for 10min. The 10min incubation was followed by quenching the reaction with 0.125M glycine for 5min at RT on shaker. The glycine was added to the cell growth media at a final concentration of 0.125M taking into account combined volumes of media and the PFA. Next, the cells were washed 4 times with cold 1X Phosphate Buffered Saline (PBS, Corning). The washing step was followed by scraping the cells from the plate surface with 1ml ice cold 1X PBS four times and collecting them into 15ml conical tube (15ml Centrifuge Tube, Celltreat). The cells were spun down for 15min at 2000 rpm at 4°C. The cells were frozen down at -80°C before the next step.

The cells were resuspended in RIPA buffer (Sigma) with PIC (cOmplete™, Mini, EDTA-free, Roche) and PMSF (32%, Electron Microscopy Sciences) (50×10⁶~2.5ml RIPA buffer). The cells were incubated in 1XRIPA buffer (Sigma) for 30min on ice with vigorous vortexing every 5min to lyse the cells (the cell lysis can be checked under microscope). 1ml aliquots of the samples were transferred in microcentrifuge tubes. The cell lysate was sonicated in a cup-horn digital sonicator with ice to shear the DNA. Each sonication was 1min followed by 1min rest on ice at high amplitude for 30 times. Next, the cells were spun down at 13,000rpm for 20min at 4°C. The supernatant was transferred into new tubes. If there was more than one aliquot for the same
treatment, the samples were pooled. BCA assay was done for protein concentration (Pierce® BCA Protein Assay, Pierce).

Aliquots were made and stored at -80°C. The protein samples were as followed: 40µg protein (10%) for input, 800µg protein for IP (400µg aliquot is used for immunoprecipitation, 400µg aliquot is used for a negative control with a non-specific antibody, IgG), and 100µl or less to save for DNA shearing check.

400µg of protein/sample were pretreated with Protein A/G PLUS-Agarose beads (Santa Cruz) for 30min at 4°C. The beads were removed and an antibody was added to the protein sample. The following antibodies were used for the experiments: H3K9me3 (Anti-Histone H3 (tri methyl K9) antibody – ChIP Grade ab8898, Abcam), H4K20me3 (Anti-Histone H4 (tri methyl K20) antibody – ChIP Grade ab9053, Abcam) and H3K4me3 (Anti-Histone H3 (tri methyl K4) antibody (mAbcam1012) - ChIP Grade ab1012, Abcam), and IgG (negative control) (Normal rabbit and mouse IgG, Santa Cruz Biotechnology). The concentration of the antibody can vary depending on the sample. A titration curve was performed and 6µg antibody concentration was used. The samples were incubated overnight at 4°C.

Protein A/G PLUS-Agarose beads (Santa Cruz) were used for the pull-down. They were blocked with 0.3mg/ml of salmon sperm DNA for 30min at 4°C, then washed 3 times with cold 1XPBS. For the washing, the beads were spun down for 2min at 3000rpm at 4°C. There were approximately 25% of bead slurry in the 50µl beads/sample. 50µl of 25% bead slurry was added to the immunoprecipitation (IP) sample (a cell lysate and antibody) and incubated for 2h at 4°C.
After 2h, the samples were spun down for 2min at 3000rpm at 4°C; the supernatant was removed. The samples were washed with 1ml of the following solutions and spun for 2min at 3000rpm at 4°C between washes: Wash 1, Wash 2, Wash 3, and TE buffer twice (pH=8). After the washes, 100µl of 1X Elution buffer (freshly made)/5µl of Proteinase K (20mg/ml, ThermoFisher Scientific) were added. For input samples, the 40µg aliquots volume of the sample were adjusted to approximately 80µl with 1XRIPA buffer. 10µl of 10X Elution buffer/5µl of Proteinase K (20mg/ml) were added to the samples. 11µl of 10X Elution buffer/5µl of Proteinase K (20mg/ml) were added to the DNA shearing sample.

**ChIP Wash Solution:**

For 500µl adjust volumes with H₂O and filter:

**Wash 1:** 0.1% SDS (5ml of 10% SDS)
1% Triton-X (5ml of Triton-X 1000)
20mM Tris pH8 (10ml of Triton-X 100)
150mM NaCl (15ml of 5M stock)

**Wash 2:** 0.1% SDS (5ml of 10% SDS)
1% Triton-X (5ml of Triton-X 100)
20mM Tris-Cl pH8 (10ml of 1M stock)
500mM NaCl (50ml 5M stock)

**Wash 3:** 0.25M LiCl (5.3g LiCl mol. Weight 43.39g/mol)
1% IGEPAL/NP-40 (5ml)
1% Deoxycholate – 5g NaDeoxycholate
1mM EDTA (1ml of 0.5M stock)
10mM Tris pH8 (5ml of 1M stock)

**10X Elution Buffer:**

The liquid was kept at 65°C to avoid viscosity. For 500µl buffer, we added 10% SDS and 0.042g sodium bicarbonate (NaHCO₃, Sigma). For 1X, we diluted with H₂O.
After adding the elution buffer, the tubes were incubated at 65°C overnight. The next day, the DNA was purified, using a QIAGEN PCR Purification kit. qPCR was used to analyze the results. The qPCR primers were initially applied in a regular PCR (dNTPs, Invitrogen; Titanium Taq Polymerase, Clontech Laboratories), using input DNA as a template and an agarose gel to assess the presence of a product. The qPCR primers were tested on various dilutions of genomic DNA template from *ROSA-MT/M71-IRES-CRE* cells to determine their efficiency (Rosso, 2015).

2.2 High-throughput screenings assay, using small chemical compound libraries

Blastocyst-derived mESCs were isolated from *ROSA-MT/M71-IRES-CRE* mouse line (Chen, 2006). The cells were split onto 384-well plate (Corning®, cat. #3683), with the seed density of 0.3x10⁴ cells/well. 24h after plating, the cells were treated with chemical compounds from Prestwick chemical library (1280 compounds Prestwick Chemical, provided in DMSO), MicroSource library (2300 compounds, MicroSource Discovering System, Inc., provided in DMSO), and LOPAC (1280 compounds, Sigma-Aldrich-Millipore, Lopac®) (Cong, 2013), at 1µM and 10µM concentration. The *ROSA-MT/M71-IRES-CRE* cells with the chemicals were grown at 37°C, at 5%CO₂ overnight. The following day, the cells were fixed in 1% Paraformaldehyde solution EM Grade (PFA, Electron Microscopy Sciences) and washed three times in 1X Phosphate Buffered Saline (PBS, Corning). The cells remained in PBS before imaging.

Data were analyzed at the Cell Screening Core of Weill Cornell Medical College. An ImageXpress MICRO imaging system from Molecular Devices equipped with a Photometrics CoolSnapHQ camera from Roper Scientific were used to acquire images. TdTomato images were acquired using 543/22 nm excitation and 593/40 nm emission filters with a 569-dichroic
long-pass filter. GFP images were acquired using 472/30 nm excitation and 520/35 nm emission filters with a 502-dichroic long-pass filter. Plates were transported from plate hotels, using a CRS CataLyst Express robot (ThermoFisher Scientific). Images were acquired at four sites per well (nine sites per well in two experiments). Sites were centered in the wells with 200µm spacing between sites. 696x520 pixel images (897x670 µm) were acquired at 12 intensity bits per pixel. Each pixel is 1.29x1.29 µm in the object (Weill Cornell Medical College) (Figure 15).

2.3 Plasmid constructs

2.3.1 M71 OR transgenes

M71 OR backbone plasmids were created for a screening assay to study the significance of OR enhancer elements, a promoter, and/or a coding sequence for the expression of an OR (Table 1).

2.3.2 GPCR constructs, used in a trafficking assay in OP6 cells

To create β2AR and M71 OR transgenes for a featured *in vitro* assay, a peGFP-N1 (ClonTech) expression vector was used. A PCR of the gene of interest was performed, such that the sequence was flanked by 5’ EcoRI and 3’ PacI containing primers. The PCR products were originally subcloned into a pGEM-T Easy (Promega) vector system, followed by EcoRI – PacI restriction digest, and subsequently cloned in the expression vector.
2.4 Transfection assays in mESC reporter and olfactory placode cells

2.4.1 *ROSA-MTMG;M71-IRES-CRE* cells were grown in T-25 flasks to confluency (2500mm², BP-Falcon, cat# 353109). The cells from a single flask were used for the total of two transfections. The cells were grown in Dulbecco’s modified Eagle’s medium (DMEM, 1X EmbryoMax, Millipore, cat# SLM-220-M), supplemented by 10% fetal bovine serum (FBS, Millipore, embryonic stem cells grade, cat# ES-009-B), 1% penicillin/streptomycin (100X, Millipore, cat# TMS-AB2-C), 1% non-essential amino acids (100X, Millipore, cat# TMS-001-C), 1% nucleosides (100X, Millipore, ES-008-D), and L-Glutamine (100X, Millipore, TMS-002-C) at 37°C and 5%CO₂. *ROSA-MTMG;M71-IRES-CRE* cells were passaged every day using 0.05%Trypsin (Gibco, cat# 25300-054). The cells were transfected with 10µg DNA, using Amaxa nucleofactor technology (Lonza; COS-7, DSMZ setting), using 100µl of Ingenio® Solution (Mirus) and plated on 60mm tissue culture plates (Falcon). Half way through the project, the mouse embryonic fibroblast (pMEF-n, Mitomycin-C inactivated, neomycin-resistant, Millipore) was replaced with by adding inhibitors to the media to maintain the mESCs undifferentiated: Stemolecule™CHIR99021, inhibiting GSK-3β and Stemolecule PD0325901, targeting MAPK/ERK kinase (Stemgent).

2.4.2 The mouse olfactory placode cell line, OP6, supplied by Jane Roskams, was grown in Dulbecco’s modified Eagle’s medium (DMEM, 1X Invitrogen), supplemented by 10% fetal bovine serum (FBS, Millipore, embryonic stem cells grade) and 1% penicillin/streptomycin at 33°C and 5%CO₂. OP6 cells were passaged every two days using 0.05% Trypsin (Millipore). The cells were transfected with 10µg DNA + 100µl of Ingenio® Solution (Mirus), using Amaxa nucleofactor technology (Lonza; COS-7, DSMZ setting), and plated on either 60mm tissue
culture plates (Falcon) or 35mm optical bottom plates (MatTek). Imaging of the non-fixed cells was performed 24h post transfection, using a scanning confocal microscope Zeiss LSM 510.

2.5 Neomycin frameshift in mESCs by CRISPR

Using CRISPR database the following guide line plasmids targeting neomycin were designed and cloned in X330 plasmid (Zhang lab, 2016):

CRISPR guide 1 kill neo top: cacgcTGCACAAGGAACGCCCGTCGTGG
CRISPR guide 1 kill neo bottom: aaacCCACGACGGCGTTCCTTGCAC
CRISPR guide 4 kill neo top: cacgcGCTCTGAAGGCGCGCTTGG
CRISPR guide 4 kill neo bottom: aaacCCGGAACAGGCGGACTAGAGCc
CRISPR guide 7 kill neo top: cacgcACGGCTTCTCCGCGGCCTTGG
CRISPR guide 7 kill neo bottom: aaacCCAACGGCCGGAGAACCTGCGTc.

These plasmids were transfected in ROSA-MTMG;M71-IRES-CRE cells to target and destroy neomycin by cleaving its sequence. The ESCs were trypsinized to a single cell suspension. The cells were counted using hemocytometer. 3x10^6 cells were used for a single transfection. The cells were re-suspended in the Lipofectamine®2000 mixture:

**Tube 1:** 2µg of DNA + 1.5ml Opti-MEM (RT)

**Tube 2:** 60µl of Lipofectamine®2000 + 1.5ml of Opti-MEM (RT).

The tubes were incubated at room temperature for 5min separately, then mixed, and incubated for 20min. The mixture was added to a 10cm dish and incubated for 3h, which was followed by adding cell media up to 10ml and incubating the cells at 37°C, 5%CO₂ overnight.

2.6 Laser scanning confocal microscopy for imaging purposes

We imaged OP6 cells on a Nikon A1 laser scanning confocal microscope using a 60X oil immersion objective. GFP was excited at 488nm and collected at BP500-545nm, while mCherry was excited at 561nm and collected at LP575nm (Bubnell, 2015).
2.7 Membrane non-permeabilization antibody-staining assay to detect Ct GFP (in vitro)

OP6 cells were transfected, using Ingenio® Solution and Amaxa nucleofactor technology (COS-7, DSMZ setting). The cells were plated on 35mm optical bottom plates (MatTek). After 24h, the media from the transfected cells was removed. The cells were stained with an anti-GFP antibody (Millipore, AB10145, rabbit polyclonal, 1/1000 dilution). The staining buffer contained 1X Opti-MEM® reduced serum media (Gibco by Life Technologies), 1.5M sodium azide (NaN₃, Sigma), and 1X HEPES buffer solution (Gibco by Life Technologies). The cells were incubated for 1h on ice. After 1h, the staining solution was removed and the cells were washed three times with washing solution, containing 1X Hanks’ Balanced Salt Solution (HBSS, Cellgro), 1.5M sodium azide, and 1X HEPES solution. Next, the cells were stained with secondary goat anti-rabbit 546 antibody for 30min on ice (Alexa546, 1/500 dilution). The cells were washed three times with washing solution. The OP6 cells were fixed for 15min on ice, using 1% PFA EM Grade (PFA, Electron Microscopy Sciences) and washed three times in 1X Phosphate Buffered Saline (PBS, Corning). The cells were imaged in 1X Live Cell Imaging Solution HEPES buffered physiological saline (Gibco by Life Technologies), using LSM scanning confocal microscopy.

2.8 Membrane non-permeabilization antibody-staining assay to detect Ct GFP (in vivo)

We performed the non-permeabilization staining protocol, using a wholemount sample from the following mouse lines: FR1α M71GFP (5x2I_M71::GFP-IRES-tauCherry); FR1α M71-IRES-tauLacZ (5x2I_M71::GFP-IRES-tauCherry,M71-IRES-tauLacZ); FR3 β2AR-GFP (5x2I_β2AR::GFPrestaumCherry); DH (Deletion OR-Venus::MS2 ; OMP-H2B::mCherry), and TM1α (5x2I_β2AR::GFP Nt→TM1irestaumCherry) (Chart 1).
Figure 15. A single-well imaging area for a 384-well plate.

In the high-throughput screening of chemical compounds libraries activity on M71 OR in ROSA-MTMG;M71-IRES-CRE reporter mESCs, the imaging area included four images taken within the well. For a single well, that area covered approximately 34% of the well within a 384-well plate setting. All colonies residing within the area were captured, imaged, and analyzed for a TdTomato\textrightarrow{}GFP shift in fluorescence.

Chart 1. Membrane non-permeabilization antibody-staining assay to investigate the extracellular presence of a C-terminus GPCR::GFP in OP6 cells.

We established a membrane non-permeabilization imaging assay, where we used an anti-GFP antibody (EMD Millipore) to detect the C-termini GFP presence of the M71 OR and β2AR truncation series in OP6 cells. We used a membrane non-permeabilization protocol to ensure the GFP detection outside of the cell (Zhuang, 2008).
CHAPTER 3. Investigating chromatin histone modifications of odorant receptor promoters in mouse embryonic stem cells. Establishing a chemical screening assay to differentiate mouse embryonic stem cells into mature olfactory sensory neurons

Studying singular OR choice by the Lomvardas lab in the mouse system, led to a current model of OR mechanism of expression, according to which ORs are suppressed by “off” repressive epigenetic marks. In this model, all ORs are epigenetically inhibited at a neuronal stage prior to maturation at a progenitor level (basal cells), before a single OR is chosen to be expressed by an OSN (Magklara, 2011). Two examples of repressive epigenetic modifications are H3K9me3 and H4K20me3 post-translation modifications (PTMs), added on the histones of OR genes (heterochromatic regions of the OR clusters). OR genes, which make up 5% of all genes in mice, should not be active outside of the olfactory system. However, “off” marks are observed on OR genes in non-olfactory tissues, such as liver cells, albeit at a lesser intensity (Magklara, 2011). Consistent with that notion, the “on” mark, H3K4me3, is absent in hESCs, despite being found in 75% of all other genes (Guenther, 2007).

Currently, epigenetic studies of OR clusters are performed on OSNs extracted from olfactory tissue (no olfactory sensory neuronal cell lines exist). For that reason, we were intrigued by Guenther et al. studies of hESCs, since stem cells can be grown in tissue culture for prolonged periods of time. To study epigenetic landscape of OR promoters, we chose mESCs, which are even more robust in tissue culture, for epigenetic analysis of OR clusters.

To that end, we developed a high-throughput assay, that utilizes a reporter system, where we derived a mESC line from blastocyst of ROSA-MTMG;M71-IRES-CRE transgenic mouse line
(Bryja, 2006; the extraction of the mESCs was performed by Eugene Lempert). Originally, we created the *ROSA-MT MG;M71-IRES-CRE* transgenic mouse by crossing ROSA-MT MG animal (LoxP-TdTomato-LoxP-GFP is gene targeted in the ROSA26 locus) with a *M71-IRES-CRE* animal (IRES-CRE is gene targeted in the M71 OR endogenous locus, after the OR) (Figure 9). In the newly created *ROSA-MT MG;M71-IRES-CRE* mouse, GFP fluorescent axons are observed to converge onto the appropriate location glomeruli in the dorsal bulb (Figure 10) (Ressler, 1994). CRE recombinase co-expresses with M71 OR through the use of an Internal Ribosome Entry Site (IRES). *ROSA-MT MG;M71-IRES-CRE* mESCs express membrane bound TdTomato fluorescent protein in all cells. When M71 OR/CRE recombinase is transcribed and translated, TdTomato is excised and a membrane bound GFP fluorescence is expressed.

The newly extracted mESCs were karyotyped to determine the chromosomes morphologies. The results showed that the *ROSA-MT MG;M71-IRES-CRE* mESCs were overall healthy. However, 30% of the cells have trisomy 1 (Table 2) (D'Hulst, 2013). However, trisomy 1 is not known to harm the proper functions of a mESC lines. To further assess the properties of the newly made mESC line, we performed an immunostaining experiment by using Oct3/4 antibody, a well-known marker for ESCs. According to our results, Oct3/4 gene is active, transcribed, and translated in the *ROSA-MT MG;M71-IRES-CRE* mESCs (Figure 11). In addition, we also examined the morphology of the mESC line and concluded that it is consistent with our experience of growing mESCs in the lab (Figure 16). At this point, these mESCs have been passage over the course of seven years, suggesting a sustained pluripotent cell status.
In both mice and mESCs, the CRE/LoxP system allows CRE recombinase to recognize a DNA region of interest, flanked by LoxP sites (Sauer, 1988). In the absence of CRE expression, all mESCs are red due to the presence of TdTomato. When M71 OR is expressed, TdTomato is excised, allowing the presence of GFP fluorescence in the mESCs. Upon observation, we were surprised that none of the native ROSA-MTMG;M71-IRES-CRE mESCs ever expressed GFP, suggesting that ORs were under tight control, i.e. no leaky expression (Figure 17A). To test the functionality of the ROSA reporter, we transfected a CMV promoter-driven CRE plasmid (pOG231, a gift from Geoff Wahl, Addgene plasmid #17736) in ROSA-MTMG;M71-IRES-CRE mESCs (O'Gorman, 1997). The exogenous CRE resulted in TdTomato→GFP shift in fluorescence at a mosaic GFP expression pattern (Figure 17B, 18).

Using the ROSA-MTMG;M71-IRES-CRE mESC line, we employed several strategies to understand the expression of OR genes. First, we set up a chromatin immunoprecipitation-quantitative polymerase chain reaction (ChIP-qPCR) assay (Rosso, 2015) to examine the types of epigenetic chromatin marks on M71 OR. We used antibodies against H3K4me3 (transcription initiation on-mark) and H3K9me3 and H4K20me3 (heterochromatin off-marks) to determine the epigenetic landscape of the M71 OR promoter in mESCs (Figure 19). Using MIQE qPCR guidelines (Bustin, 2009), our results were normalized against GAPDH and Oct3/4 as positive controls for on-marks (genes known to be actively transcribed in mESCs) (Nichols, 1998) and beta-globin gene as a reference for off-epigenetic marks (beta-globin gene is known not to be on in mESCs) (Magklara, 2011).
Table 2. *ROSA-MTMG;M71-IRES-CRE* mouse embryonic stem cells karyotyping results.

<table>
<thead>
<tr>
<th>Cell Line Name</th>
<th>Type of Trisomy</th>
<th>% of Cells with Trisomy</th>
<th>Number of Karyotyped Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROSA-MTMG;M71iresCre</td>
<td>Trisomy 1</td>
<td>31</td>
<td>54</td>
</tr>
</tbody>
</table>

To determine the chromosomal morphology, we cytogenetically karyotyped 54 cells of the *ROSA-MTMG;M71-IRES-CRE* line. The experiment was performed at Memorial Sloan Kettering Cancer Center. Based on the results, the newly established *ROSA-MTMG;M71-IRES-CRE* mESC line contains 31% cells exhibiting Trisomy 1. Trisomy 8, which often causes difficulties with using the mESC lines, however, is not observed in the *ROSA-MTMG;M71-IRES-CRE* cell line.

Figure 16. Exogenous CRE expression allows for the presence of GFP-positive cells in *ROSA-MTMG; M71-IRES-CRE* mESCs.

When exogenous CRE recombinase excises the TdTomato sequence between the loxP sites within *ROSA-MTMG;M71-IRES-CRE* mESCs, we observe a switch from red to green fluorescence in a mosaic pattern (some red cells are preserved within the colony). In addition, the morphology of the ESC colony is preserved, where the cells sustain their trapezoid shape and the colony maintains a globular shape. Both the switch in the fluorescence and the morphology of the colony are factors which allow us to determine that the mESCs have undergone through a successful CRE expression. Scale bar = 100µm.
A. In native ROSA-MT MG;M71-IRES-CRE mESCs, where M71 OR is not expressed, endogenous CRE protein is not produced. All cells are red due to the presence of the TdTomato in the ROSA26 locus. B. In a control experiment, where exogenous CRE is transfected in the ROSA-MT MG;M71-IRES-CRE mESCs, the TdTomato gene is excised and the reporter system switches from red to green fluorescence. The morphology of the mESC colonies remain maintained, where the mESCs retain trapezoid shape, while the colonies maintain globular shape. The switch in the fluorescence and morphology of the colony determine successful recombination. Scale bar = 100 µm.

Figure 18. Protein expression of CRE recombinase in ROSA-MT MG;M71-IRES-CRE.

We transfected a CRE plasmid in ROSA-MT MG;M71-IRES-CRE mESCs. 24h later, we performed immunofluorescence assay, using anti-CRE antibody against the CRE protein expressed in the cells. Red membrane staining indicates the presence of the TdTomato and the lack of CRE expression in the cells. When CRE excises the TdTomato sequence, we observed GFP fluorescence. The blue staining demonstrates the presence of CRE protein within the cells. Scale bar = 20 µm.
Figure 19: ChIP-qPCR assay.

The mESCs are lysed and the anti-epigenetic mark antibody is crosslinked to samples. The samples + the antibody are bound to A/G agarose beads. DNA is washed and eluted. A qPCR is performed, using primers designed for the genes of interest. (https://wolfe4e.sinauer.com/ch/14/act1401/osnpic.jpg)
3.1 ChIP-qPCR assay exploits the epigenetic landscape of M71 OR gene promoter in *ROSA-MTMG;M71-IRES-CRE* mouse embryonic stem cells

Due to the lack of leaky M71 OR expression in our reporter system, we decided to investigate if M71 OR is epigenetically maintained in a repressive state in mESCs. We were interested in examining the histone epigenetic marks on the M71 OR promoter in the *ROSA-MTMG;M71-IRES-CRE* mESCs. We were particularly interested in looking at the H3K4me3 on-epigenetic mark and H3K9me3 and H4K20me3 heterochromatin off-marks. We used GAPDH and Oct3/4 gene as a positive control/reference gene for the presence of positive epigenetic marks, since GAPDH and Oct3/4 genes are expressed in mESCs. In similar fashion, we used β-globin gene (HBB) as a reference gene for H3K9me3 repression mark in mESCs. HBB is localized within an OR locus on chromosome 7 and is transcribed after birth (Kiefer, 2008). Furthermore, we used a non-specific antibody, IgG, as a background pull-down reference.

We examined the H3K4me3 and IgG pull-downs, collected from the mESCs lysates, where the H3K4me3 and IgG antibodies immunoprecipitated the proteins of interest. We used primers against the M71 OR promoter region and GAPDH CDS to verify their enrichment for the transcription initiation mark (according to our data, the GAPDH gene retains higher levels of H3K4me3 than Oct3/4). Similar assay was set up to study the presence of H3K9me3 and H4K20me3 off-marks at the M71 promoter and HBB gene. Each experiment was repeated twice and samples were loaded in triplicates. The calculations were performed as an average fold of enrichment of the epigenetic mark vs. the non-specific IgG antibody pull-down.

When GAPDH was used as a positive control for the presence H3K4me3, we observed a 30-fold enrichment for H3K4me3 for GAPDH in comparison to the IgG pull-down. We did not see
enrichment for the H3K4me3 mark for the M71 OR promoter. We performed a one-way ANOVA and our results were statistically significant for the GAPDH H3K4me3 enrichment in comparison to the IgG non-specific pull-down and M71 OR H3K4me3 pull-down (*p<0.05) (Figure 20).

H3K9me3-enriched HBB and M71 OR were immunoprecipitated, when pulled down by the H3K9me3 antibody in comparison to the non-specific IgG HBB and M71 OR pull-down. We ran a one-way ANOVA test and determined that there is statistically significant H3K9me3 HBB and M71 OR H3K9me3 enrichment vs. the IgG background reading (****p<0.0001) (Figure 21). Also, we examined the H4K20me3 HBB enrichment in comparison to the IgG non-specific HBB pull-down. We observed a 12.5-fold H4K20me3 enrichment for HBB in comparison to the background. When we inspected the H4K20me3 at the M71 OR promoter region, we observed a 7-fold increase to the IgG pull-down. We ran a one-way ANOVA, which disclosed that the H4K20me3 HBB fold increase to IgG pull-down is statistically significant (****p<0.0001). In addition, there is statistical significance in the fold change of the H3K20me3 enrichment for the M71 OR promoter in comparison to the IgG non-specific pull-down (**p<0.01) (Figure 22).

The ChIP-qPCR results of the M71 OR promoter epigenetic screens suggest that OR promoters are repressed in mESCs. In addition, OR genes are epigenetically silenced and not expressed in other non-olfactory tissues, such as the liver and hECSs. These findings suggest that epigenetic silencing might not an olfactory specific phenomenon. However, we have only examined the lack of H3K4me3 transcription initiation mark and the presence of H3K9m3 and H4K20me3 marks on a single OR promoter. It might be beneficial to investigate the epigenetic landscape of promoter regions of other ORs in our reporter system.
We used anti-H3K4me3 antibody to pull-down H3K4me3 methylated genes in *ROSA-MTMG;M71-IRES-CRE* reporter mESCs. We normalized the pull-down enrichment for GAPDH and M71 OR to IgG non-specific pull-down. There is a 30-fold H3K4me3 enrichment for GAPDH in comparison to the IgG pull-down. There was a lack of fold change of the H3K4me3 enrichment for M71 OR gene in comparison to the IgG non-specific pull-down. (*p<0.05; a one-way ANOVA; error bars indicate mean +/-SD).
Figure 21. Anti-H3K9me3 antibody immunoprecipitation of β-globin and M71 OR in ROSA-MT MG; M71-IRES-CRE reporter mESCs.

We used anti-H3K9me3 antibody to pull-down H3K9me3 methylated genes in reporter mESCs. The pull-down for the HBB and M71 OR enrichments for H3K9me3 is normalized to an IgG non-specific background pull-down. For both genes, the enrichment is 12.5-folds higher than the background. (****p<0.001; a one-way ANOVA; error bars indicate mean +/-SD).
We performed an anti-H4K20me3 antibody pull-down in our reporter mESCs. The HBB gene enrichment for H4K20me3 is 12.5-fold higher than the non-specific IgG background, while the M71 OR enrichment for the same repressive mark is 7-fold higher than the IgG non-specific pull-down. M71 OR H4K20me3 enrichment is statistically significant to the IgG non-specific pull-down. The HBB H4K20me3 enrichment is statistically significant to IgG background (**** p<0.0001, ***, p<0.005, **, p<0.01; a one-way ANOVA, error bars indicate mean +/-SD).
To better understand how epigenetic marks are correlated to the expression of a single OR allele, we attempted to adapt our ChIP-qPCR assay to analyze the epigenetic modifications of OR promoters in transgenic mice. We used *nx21* transgenic lines, where a the probability for the *nx21* OR to be expressed is increased. We investigated if higher expression of an OR might alter the epigenetic landscape for that OR in comparison to the wild type. For this purpose, we used mouse lines overexpressing M71 OR and V1rb2 genes (*4x21* M71-IRES-*tau*Cherry; *5x21* V1rb2-IRES-*tau*Cherry + M71-IRES-*tau*GFP). To the moment, our experiments have not been successful and need to be optimized.

### 3.2 Effect of small compound chemical libraries on epigenetic regulation of M71 OR promoter in mouse embryonic stem cells

M71 OR promoter is negatively marked in *ROSA-MTMG;M71-IRES-CRE* reporter. Using a small chemical library screening assay acting on the OR promoter, we intended to reverse the OR repression and generate OSN lines from our mESCs. We were interested in cells that were “mature” olfactory neurons, which expressed or attempted to express an OR.

Small molecule chemical libraries allow the establishment of high-throughput assays, because the chemical compounds are standardized, cost-effective, and stable. We used chemical compounds libraries from a collaborator of ours, Dr. Shuiping Chen from Weill Cornell Medical School. Currently, small chemical molecules are used to target signaling pathways and influence epigenetic modifications and transcription factors activity. For instance, a molecule, such as vitamin C, reduces the H3K9me2 levels (a precursor of the H3K9me3 repressive epigenetic mark) in native ESCs and increases the possibility for a gene to be expressed (Prestwick screening library by Prestwick Chemicals) (Ebata, 2017; Qin, 2017; Ghazizadeh,
Our goal was to uncover similar compounds that activate an OR gene expression in the reporter mESCs, leading to the expression of the M71 OR, simultaneous CRE expression, and the occurrence of GFP fluorescence (Figure 23).

When applying the chemical library screening platform, we exploited approximately 4860 screening compounds on 384-well plates setup, at two different concentrations (1µM and 10µM) (Prestwick library by Prestwick Chemical, which contains approximately 1280 compounds; MicroSource by MicroSource Discovering System, Inc.: ~2300 compounds; and LOPAC chemical library by Sigma Aldrich: ~1280 compounds) (Chen, 2006) (Figure 24, 25, and 26).

To establish the assay, we initially set up its control parameters. We plated \textit{ROSA-MTMG; M71-IRES-CRE} mESCs onto a 384-well plate. According to our control results, we observed red mESC colonies, lacking green membrane fluorescence and leaky expression for M71 OR (Figure 25). Simultaneously, we imaged a second control 384-well plate, where the first two rows of the plate contained \textit{ROSA-MTMG;M71-IRES-CRE} mESCs, while the third row contained \textit{ROSA-MTMG;M71-IRES-CRE} mESCs, transfected with CRE recombinase plasmid (pOG231). The first two rows of the plate sustained red membrane fluorescence; however, the cells in the third row underwent TdTomato\overset{GFP}{} shift in fluorescence, due to the presence of CRE enzyme and the excision of the TdTomato. The observed colonies, however, were mosaic and remained partially red (Figure 27).

As a next step, we established a fixation protocol, where both untreated and CRE transfected \textit{ROSA-MTMG;M71-IRES-CRE} mESCs were fixed with polyformaldehyde. In such a way, the
reporter cells treated with small chemical compounds prior fixation, could been imaged at a later point of time. We analyzed the fluorescence levels and we observed that the fixed cells maintained their fluorescence post treatment (Figure 28).

After establishing the control parameters, we determined the assay readouts. The data were represented as log10 of the red fluorescence values (Tdtomato area/AreaA in \( \mu m^2 \)) and green fluorescence values (GFP area/AreaB in \( \mu m^2 \)), calculated based on the captured mESC fluorescent images by ImageXpress MICRO imaging system at Weill Cornell Medical School. Using the red and green fluorescence values, we established ranking of the control values (in \( \mu m^2 \)) (Figure 29) in descendant order of the green fluorescence of the cells (AreaB), where the ranking table contained the highest level of green fluorescence (in \( \mu m^2 \)) at the top (Table 3). We plotted the values on a graph (Figure 30). On the graph, a base line value of log10=3.5 (in \( \mu m^2 \)) (y-axis) was the sum of the green background fluorescence in the untreated mESCs (of 3.16) and a standard deviation value (0.3687), depicting the variation of the GFP baseline controls amongst themselves. Any value over the base line of 3.5 had been considered a possible positive result, where GFP fluorescence reflects M71 OR/CRE expression in mESCs.

Referencing the control values for each library screening, we performed a high-throughput screening assay, utilizing Prestwick, MicroSource, and LOPAC chemical libraries. We collected data from every well of the 384-sample plates treated with chemical compounds, and ranked it similarly to the control samples, where the log10 of the highest green fluorescence values from AreaB (in \( \mu m^2 \)) were at the top of the table (see Table 3). Higher value (in \( \mu m^2 \)) corresponds to brighter fluorescent image. The control red and green fluorescence values were used as
references in every graph, created based on the *ROSA-MTMG;M71-IRESCRE* mESCs fluorescence after compound treatment (Figure 30). To check the mESC morphology and fluorescence after compound treatment, we obtained composite images for the 384-well *ROSA-MTMG;M71-IRESCRE* plates treated with LOPAC library compounds at 1µM concentration (Figure 31-34). We repeated the experiments at 10µM LOPAC chemical compound concentration (data not shown). The assay was performed, using two other chemical libraries at 1µM and 10µM concentration (data not shown).

The small compound chemicals, used for the treatments, were stored in 384-plates, where a single plate contained 352 compounds, corresponding to four 96-well plates transferred onto a 384-plate template (columns 1, 2, 11, and 12 were empty, hence, they were background readings). In the 384-well plate, every single compound was given a number, based on the number of the original plate and the chemical’s position in it. For instance, if the number of the 96-well plate was #1 and the chemical compound was located in A2 (A1 was an empty well in the 96-well plate as well), therefore, the compound was located at position A3 in the 384-well plate. The following compound belonged to 96-well plate #2 and was located at position A2. Its new position was A4 in the 384-well plate. Next, the chemical from A2, from 96-well plate #3, became B3 and the chemical from A2, from 96-well plate #4, was B4 in the 384-well plate. The template followed the same pattern for all 384-well chemical containing plates (Figure 35).

To specifically address the epigenetic regulations as a reason for the lack of expression of ORs in mESCs, we also utilized an epigenetic targeting compound library from Cayman Chemical (~175 compounds) (Table 4). These chemicals targeted the activity of proteins involved in
adding epigenetic modifications, such as methyltransferases, demethylases, histone acetyltransferases, and histone deacetylases (Qin, 2017).

M71 OR, however, was a single receptor out of 1400 ORs in the reporter system. To circumvent that, we used MouSensor technology to increase the probability of M71 expression (D'Hulst, 2016). As per MouSensor platform, a conserved enhancer element was multimerized four to nine times and placed in front of an OR of interest: the higher the multimer number, the higher the probability of the OR to be chosen in mice. Therefore, we expressed an M71 OR transgene, baring five multimers of the enhancer element, in the mESC reporter system. 24h after the transfection, we re-tested the epigenetic library on the mESCs, hoping that the probability for M71 OR to be chosen increased due to the presence of an enhancer element multimer.
Figure 23. M71 OR/CRE expression allows for excision of TdTomato in ROSA26 locus and TdTomato→GFP shift in fluorescence in \textit{ROSA-MTMG;M71-IRES-CRE} mESCs.

If a chemical compound triggered the M71 OR expression in the in reporter mESCs, CRE recombinase would be simultaneously activated. The enzyme would excise the TdTomato sequence in between the loxP sites and GFP would be expressed. In this case, we would observe a switch from red to green fluorescence in the reporter system, which would suggest that the compound might be involved in the OR expression in mESCs.

Figure 24. Prestwick chemical compounds library pie chart, based on the mechanisms of activity of the chemical compounds.

Approximately, 33\% of the 1280 compounds (~400 compounds) of the Prestwick chemical library by Prestwick Chemical are represented on the pie chart above. The chemicals were grouped based on their chemical activity mechanisms in cells by Alessandro Rossa.
Figure 25. **MicroSource chemical compounds library pie chart, based on the mechanisms of activity of the chemical compounds.**

Approximately 50% of the 2300 chemical compounds (1040 compounds) from the MicroSource chemical library by MicroSource Discovery Systems, Inc. are included on the pie chart above. The chemicals were grouped based on their chemical activity mechanisms in cells by Alessandro Rossa.

Figure 26. **LOPAC chemical compounds library pie chart, based on the mechanisms of activity of the chemical compounds.**

All 1280 chemical compounds belonging to the LOPAC library are included in the representative pie chart. The chemicals were grouped based on their chemical activity mechanisms in cells by Alessandro Rossa (LOPAC®1280).
Figure 27. Control images of ROSA-MTMG:M71-IRES-CRE mESCs, taken through mCherry and GFP fluorescent filter.

A. mESCs were plated in three rows of a 384-well plate and imaged at the mCherry channel. The forth empty row serves as a control reading (no cells, no feeders, and no media). B. The first three rows of a 384-well plated were plated with untreated mESCs and imaged in the GFP channel; the fourth row is empty (no cells, no feeders, and no media). The red panels indicate the presence of TdTomato fluorescence in the well, while the green ones indicate the GFP fluorescent mESCs. C, D. Zoomed-in cell images in mCherry and GFP channels. These images represent the lack of leaky M71 OR expression/green fluorescence in the mESCs.

Figure 28. Control images, post CRE transfection, of ROSA-MTMG:M71-IRES-CRE mESCs taken through mCherry and GFP fluorescent filter.

A. Control images of mESCs at a 384-well plate setting, taken through a mCherry channel, where the first two rows were untreated mESCs. The third row contained mESCs, transfected with CRE plasmid (pOG231). B. Control images of ROSA-MTMG:M71-IRES-CRE cells, taken through the GFP channel, where the wells for row three contained CRE recombinase transfected mESCs. C, D. Zoomed-in images for mESCs in the mCherry and GFP fluorescent channels.
Figure 29. Control red and green fluorescence images post CRE plasmid transfection and fixation of ROSA-MTMG;M71-IRES-CRE mESCs, taken through mCherry and GFP fluorescent filter.

A, B, C. Fixed ROSA-MTMG;M71-IRES-CRE mESCs, imaged through mCherry and GFP channels in a 384-well plate. D, E, F. Fixed ROSA-MTMG;M71-IRES-CRE cells, imaged through mCherry and GFP channels in a 384-well plate. Scale bar = 100µm.

<table>
<thead>
<tr>
<th>Well</th>
<th>AreaA (Log10)</th>
<th>AreaB (Log10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>5.343588636</td>
<td>5.020650937</td>
</tr>
<tr>
<td>2.</td>
<td>5.414295671</td>
<td>4.97602448</td>
</tr>
<tr>
<td>3.</td>
<td>5.274576741</td>
<td>4.904163375</td>
</tr>
<tr>
<td>4.</td>
<td>5.235615953</td>
<td>4.81402213</td>
</tr>
<tr>
<td>5.</td>
<td>5.144319783</td>
<td>4.741846308</td>
</tr>
<tr>
<td>6.</td>
<td>5.242410077</td>
<td>4.587172533</td>
</tr>
<tr>
<td>Average:</td>
<td>5.28</td>
<td>4.84</td>
</tr>
</tbody>
</table>

Table 3. Log10 values of red fluorescence (TdTomato/AreaA, in µm²) and green fluorescence (GFP/AreaB, in µm²).

Log10 for AreaA values (in µm²) and log10 for AreaB values (in µm²) were established based on the fluorescence readouts for fixed ROSA-MTMG;M71-IRES-CRE mESCs. All untreated mESCs were red, while green fluorescence was only a result of CRE recombination in ROSA-MTMG;M71-IRES-CRE cells. These reference values were used for all follow-up graphs of the chemical compounds treatments of the reporter cells. The ranking was done based on the descendant green fluorescence values in AreaB, starting with the highest one. The bottom row reflects the averages for log10 values in areas A and B.
Figure 30. A scatter plot with a bar graph of the TdTomato/GFP fluorescence ranking in Table 3 for ROSA-MT MG;M71- IRES-CRE mESCs.

CRE plasmid was expressed in ROSA-MT MG;M71- IRES-CRE reporter system. In 24h, the cells were fixed and imaged. The fluorescence values were ranked as a log10 descending green fluorescence values (AreaB, µm²). The control experiment showed that when the green fluorescence increases, the red fluorescence decreases (TdTomato➔GFP shift). A GFP base line values depict the green fluorescence readings in untreated mESCs, (background readings/lack of green fluorescence). A y-axis value of 3.5 in µm² is the sum of the green background value of 3.16 and a standard deviation of 0.3687 (****p<0.0001, *p<0.05; a one-way ANOVA; error bars indicate mean +/-SD).
Figure 31. A 384-well plate composite image of \textit{ROSA-MTMG;M71-IRES-CRE} mESCs, treated with LOPAC chemical compounds library (GFP filter).

\textit{ROSA-MTMG;M71-IRES-CRE} mESCs, at low density, were treated with LOPAC chemical compounds at 1\,\mu\text{M} concentration. GFP images were acquired, using 472/30nm excitation and 520/35nm emission filters. Nine images were taken per well (75\% of the well’s surface). The white/grey color image reflects the presence of GFP (white represents more high green fluorescence), while black – its absence.

Figure 32. A 384-well plate composite image of \textit{ROSA-MTMG;M71-IRES-CRE} mESCs, treated with LOPAC chemical compounds library (mCherry filter).

\textit{ROSA-MTMG;M71-IRES-CRE} mESCs, at low density, were treated with LOPAC chemical compounds at 1\,\mu\text{M} concentration. The TdTomato images were acquired using 543/22nm excitation and 593/40nm emission filters. Nine images were taken per well (75\% of its entire surface). The white/grey color reflects the presence of TdTomato (white represents more intense red fluorescence), while black – its absence.
Figure 33. A 384-well plate composite image of ROSA-MTMG;M71-IRES-CRE mESCs, treated with LOPAC chemical compounds library (GFP filter).

*ROSA-MTMG;M71-IRES-CRE* mESCs, at low density, were treated with LOPAC chemical compounds at 1μM concentration. GFP images were acquired, using 472/30nm excitation and 520/35nm emission filters. Four images were taken/well (34% of the well’s surface). The white/grey color reflects the presence of GFP, while black – its absence.

Figure 34. A 384-well plate composite image of ROSA-MTMG;M71-IRES-CRE mESCs, treated with LOPAC chemical compounds library (mCherry filter).

*ROSA-MTMG;M71-IRES-CRE* mESCs, at low density, were treated with LOPAC chemical compounds at 1μM concentration. The TdTomato images were acquired using 543/22nm excitation and 593/40nm emission filters. Four images were taken per well (34% of the well’s entire surface). The white/grey color reflects the presence of TdTomato; black – its absence.
Commercially available libraries of chemical compounds were provided in 96-well plates setup by Prestwick, MicroSource, and LOPAC companies. For the purposes of 384-well format screenings, our collaborators at Weill Cornell Medical School had combined four 96-well plates onto a single 384-well plate, using the template above. All three chemical compounds libraries (Prestwick, MicroSource, and LOPAC libraries) were provided in the same template.

Table 4. A sample of Cayman epigenetic chemical library compounds.

To activate M71 OR, we treated **ROSA-MT MG;M71- IRES-CRE** reporter mESCs with 175 chemical compounds from Cayman epigenetic library. Same assay setup was used as in the rest of the chemical compound libraries high-throughput screenings.
The results of the three chemical compounds libraries treatments on the M71/CRE expression in \textit{ROSA-MTMG;M71-IRES-CRE} mESCs were ranked based on the red/green fluorescence values (in $\mu m^2$). The impact of the compounds was shown through the descending green fluorescence values, where the highest green values were on the top of the table. The data were collected 48h after the chemical treatments. Based on the ranked data, we created graphs, where we compared the fluorescence values of the compound treated mESCs (for each library) to the control fluorescent values (Figure 30). In Figure 30, we established a cut-off green fluorescence background (log10=3.5), where the compounds above the cut-off line would be re-examined in order to investigate their effect on M71 OR expression mechanism. We did not add the GFP imaging control for every experiment; therefore, we used this 3.5 cut-off baseline as a reference in the data analysis. The analysis of the fluorescence values is only one of the methods to study chemical impact on M71 OR expression in the reporter. We additionally inspected the morphology of the mESC colonies in every image. If the cells differentiated into OSNs expressing ORs, the morphology of the cells might have resembled a neuronal structure. On the other hand, a brief event of OR expression might have resulted in green cells in a preserved globular mESC colony shape (Figures 36-39).

After performing chemical screening experiments for all three chemical libraries, we examined the presence of green fluorescence in the reporter system and analyzed the images of the mESCs. Multiple GFP values were below the established baseline. It might be due to the control values not being established for each treated plate, but read on their own and used as reference values. Approximately 455 compounds were potential hits for activating M71 OR promoter in mESCs. However, when administering these compounds in ESC media without adding stem cells or
feeders, we noticed that many of these compounds were auto-fluorescent. Therefore, green fluorescence due to addition of small chemical compounds to the model system was in many cases due to auto-fluorescence of the compounds themselves (data not shown). About 55 out of 455 compounds remained promising in impacting M71 OR expression in the reporter. We repeated the screening experiment using them in two concentrations (Figure 40 and 41).

After the chemical re-screening of the 55 compounds, we analyzed the levels of green and red fluorescence (in \( \mu m^2 \)) and inspected the morphology of the treated cells. We searched for differentiated cells or globular mESC colonies containing green cells (Figure 42 and 43). In addition, we analyzed the fluorescence levels of mESCs treated with compounds, for which the green fluorescence levels was below the established baseline (\( \log_{10}=3.5 \)). We examined these images, to also analyze the cells morphology (Figure 44 and 45). After all of the analysis of the compounds impacts on the M71 OR expression, none of the \(~4860\) chemical compounds from all three libraries seemed promising in forcing M71 OR expression in \( ROSA-MTMG; M71-IRES-CRE \) mESCs.
**Figure 36.** A scatter plot with a bar graph, of TdTomato and GFP fluorescence after chemical compound treatment of *ROSA-MTMG; M71-IRES-CRE* mESCs.

*ROSA-MTMG; M71-IRES-CRE* reporter mESCs were treated with LOPAC chemical compounds (1µM). The cells were fixed 48h after treatment. The data (log10) were ranked based on descending green fluorescence values (in µm²). The fluorescence averages for the compound treated cells were TdTomato value=5.27 and GFP value=2.72 (in µm²). These results were compared to the TdTomato fluorescence baseline average of 5.53 in untreated mESCs and GFP fluorescence control average=3.73 (in µm²), where the GFP fluorescence was due to CRE plasmid transfection in mESCs. A y-axis value of 3.5 in µm² is the sum of the green background value of 3.16 and a standard deviation of 0.3687 (****p<0.0001, a one-way ANOVA, error bars indicate mean +/-SD).
**Figure 37.** A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in chemical compound treated *ROSA-MTMG;M71-IRES-CRE* mESCs.

*ROSA-MTMG;M71-IRES-CRE* mESCs were treated with Prestwick chemical compounds (1µM). The cells were fixed 48h post treatment. The data (log10) were ranked based on descending green fluorescence values (in µm²). The averages were for TdTomato value=5.05 and GFP values=1.70 (in µm²). These results were compared to the TdTomato baseline fluorescence average in native mESCs=5.53 (in µm²) and GFP fluorescence average in CRE transfected mESCs=3.73. A y-axis value of 3.5 in µm² is the sum of the green background value of 3.16 and a standard deviation of 0.3687 (****p<0.0001, a one-way ANOVA; error bars indicate mean +/- SD).
Figure 38. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in chemical compound treated ROSA-MTMG;M71-IRES-CRE mESCs.

The reporter mESCs were treated with Prestwick chemical compounds (1µM) and fixed 48h post treatment. The log10 data were ranked based on descending green fluorescence values (in µm²). The averages for the treated mESCs were for TdTomato=5.20 and GFP value=2.54. These results were compared to the TdTomato fluorescence baseline average of 5.53 (untreated mESCs) and GFP fluorescence average=3.73 (CRE transfected mESCs). A y-axis value of 3.5 in µm² is the sum of the green background value of 3.16 and a standard deviation of 0.3687 (****p<0.0001, ***p<0.005, **p<0.01; a one-way ANOVA; error bars indicate mean +/-SD).
Figure 39. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in chemical compound treated ROSA-MTMG;M71-IRES-CRE mESCs.

ROSAMMTMG;M71-IRES-CRE mESCs were treated with Prestwick chemical compounds (1µM). The cells were fixed 48h after treatment. The log10 data were ranked according to descending green fluorescence values (in µm²). The averages for the treated mESCs were TdTomato value=5.23 and GFP values=2.36. These results were compared to the TdTomato fluorescence baseline average mESCs=5.53 and GFP fluorescence average for CRE transfected mESCs=3.73. A y-axis value of 3.5 in µm² is the sum of the green background value of 3.16 and a standard deviation of 0.3687 (**p<0.01, ***p<0.005, ****p<0.0001, a one-way ANOVA; error bars indicate mean +/-SD).
Figure 40. A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in chemical compound treated ROSA-MTMG;M71-IRES-CRE mESCs.

**ROSA-MTMG;M71-IRES-CRE** mESCs were treated with 455 selected compounds from three chemical compound libraries at 1µM concentration. The cells were fixed 48h after treatment. The log10 data were ranked according to descending green fluorescence values (in µm²). The averages for the treated cells were for TdTomato value=5.25 and GFP values=3.38. These results were compared to the TdTomato fluorescence baseline average of native mESCs=5.53 and GFP control average of CRE transfected mESCs=3.73. A y-axis value of 3.5 in µm² is the sum of the green background value of 3.16 and a standard deviation of 0.3687 (****p<0.0001, *p<0.05; a one-way ANOVA; error bars indicate mean +/-SD).
Eight chemical compounds from the three chemical libraries (1µM concentration) were amongst the candidates for further experiments due to the high green fluorescence after treatment with these compounds in mESCs (in µm²). We examined the effects of these compounds on the GFP expression and the morphology of the mESCs. The names of the compounds and their effect on cells were listed. A. S(-)-Pindolol (beta-blocker, affects hypertension and angina medication); B. Sunitinib malate (tyrosine kinase inhibitor); C. Idarubicin hydrochloride (the compound has shown activity against breast neoplasms, lymphoma, and leukemia); D. Ethosuximide (an anti-seizure medication); E. Dipyridamole (inhibits a blood clot formation); F. NSC 95397 (a phosphatase inhibitor); G. Metronidazole (an antibiotic and antiprotozoal medication); H. Acetazolamide (treats glaucoma, epilepsy, hypertension). Scale bar = 100µm.

Figure 41. TdTomato and GFP fluorescence morphology after the treatment of ROSA-MT MG: M7-IRES-CRE mESCs with selected chemical compounds.
Figure 42. A bar graph, depicting the TdTomato and GFP fluorescence values after selected 55 chemical compounds treatment *ROSA-MTMG;M71-IRE- CRE* mESCs.

*ROSA-MTMG;M71-IRE-CRE* mESCs were treated with 55 chemical compounds from three commercial libraries (1µM). The cells were fixed 48h after treatment. The log10 data were ranked based on descending green fluorescence values (in µm²). The averages for TdTomato value=5.25 and GFP value=3.37. These results were compared to the TdTomato control average value=5.53 and GFP control average value=3.73. GFP cut-off line of log10=3.5 in µm² was determined by adding together the maximum level of green base line value of 3.16 in the native mESCs and a standard deviation of 0.3687 (****p<0.0001; a one-way ANOVA; error bars indicate mean +/-SD).
Selected 55 chemical compounds were candidates for an additional re-screening (1µM concentration). Here, we featured the effect of six of them on mESCs based on the GFP expression and the morphology of the cells. The names of the compounds and their chemical effect were listed. A. Docetaxel (a chemotherapeutic drug); B. Piperacetazine (an antipsychotic drug); C. Metrizamide (non-ionic contrast medium compound); D. (-)-Eseroline fumarate (a potent analgesic); E. Loratadine (an antihistamine). F. Rosolic acid (an inhibitor of DNA fragmentation). Scale bar = 100µm.
mESCs were treated with chemical compounds from three commercial libraries at 1µM concentration. The cells were fixed 48h after treatment. The log10 data were ranked based on descending green fluorescence values (in µm²). The averages for TdTomato value = 5.12 and GFP value = 2.22. The GFP values for these compounds were much lower than the GFP fluorescence control average for CRE transfected mESCs = 3.73 (****p<0.0001, ***p<0.005; a one-way ANOVA; error bars indicate mean +/-SD).

**Figure 44.** A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence expression in compound treated ROSA-MTMG;M71-IRES-CRE mESCs.

**Figure 45.** GFP expression in ROSA-MTMG;M71-IRES-CRE mESCs, where the GFP fluorescence levels were below the log10=3.5 cut-off baseline.

A-D. The green fluorescence for the mESCs after chemical compounds treatment was below the established GFP cut-off line. Scale bar = 100µm.
According to our ChIP-qPCR data, M71 OR promoter is epigenetically repressed. Thus, we extended the chemical treatment experiments to using a specialized epigenetic library by Cayman Chemical, solely including epigenetic compounds, which interfere with the activity of enzymes necessary for epigenetic modifications, such as methyltransferases, demethylases, histone acetyltransferases, and histone deacetylases, which methylate, demethylate, add acetylation groups, or remove them from the marked genes. For the screening, all compounds were utilized at two concentrations, 1µM and 10µM (Table 4). Post-treatment, we examined our ROSA-MTMG;M71-IRES-CRE mESCs for the appearance of membrane bound GFP (Figure 46). Based on our statistical analysis, we considered the same cut-off GFP fluorescence (y=3.5), where any GFP value above the baseline marks a compound, which possibly interferes with the M71 OR expression in mESCs. In addition, we examined the morphology of the compound treated cells.

Amongst the epigenetic compounds to treat the mESCs, 20 compounds led to base level of fluorescence higher than log10=3.5. However, based on the morphology of the mESCs, none of them was worth re-testing in the reporter. In sum, the epigenetic chemicals were not able to drive the expression of M71/OR CRE protein in the reporter.
**Figure 46.** A scatter plot with a bar graph, depicting TdTomato and GFP fluorescence in *ROSA-MTMG;M71-IRES-CRE* mESCs, after Cayman library chemicals treatment.

*ROSA-MTMG;M71-IRES-CRE* mESCs were treated with 175 epigenetic chemical compounds (1μM concentration). The cells were fixed 48h after treatment. The log10 data were ranked based on descending green fluorescence values (in μm²). The averages for TdTomato value=5.18 and GFP value=3.10. These results were compared to the TdTomato fluorescence control average in untreated mESCs=5.53 and GFP fluorescence control average of CRE transfected mESCs=3.73. GFP cut-off line of log10=3.5 in μm² was determined by adding together the maximum level of green base line value of 3.16 in the native mESCs and a standard deviation of 0.3687 (**p<0.0001; a one-way ANOVA; error bars indicate mean +/-SD).
A drawback to the four chemical library screens to activate M71 OR is that we attempted to observe the expression of a single OR out of a pool of a total of 2800 ORs. A single OR is expressed by 10,000 neurons on the average, which is a small number of the same receptor for experimental and analytical purposes. There was the need of system that allowed us to increase the number of the expressed M71 ORs. Previously in our lab, we have identified an 21bp enhancer element derived from the H-element, a well-known olfactory enhancer. Increasing number of multimers of the 21bp element ($n \times 21$) allows for higher probability for an OR to be expressed in mice (Figure 7) (Vassalli, 2011; D'Hulst, 2016).

Similarly, we attempted to enhance the probability of M71 OR expression in mESCs, by transfecting the cells with M71 OR transgene, containing $5 \times 21$ enhancer element. The $5 \times 21_{-}M71\text{promoter}_{-}M71\text{-IRES-}\text{CRE}$ transfected cells were treated with the Cayman epigenetic chemical compounds. 48h after treatment, we imaged the $\text{ROSA-MTMG;M71-IRES-}\text{CRE}$ mESCs, as we previously did during chemical libraries screenings. The results, however, were similar – no green fluorescence with promising morphology was observed for the mESCs after enhancer element transfection and the epigenetic treatment (Figure 47 and 48).
Figure 47. A scatter plot with a bar graph, depicting TdTomato and GFP values in enhancer-expressing ROSA-MTMG;M71-IRES-CRE mESCs.

5x21 M71 OR transfected ROSA-MTMG;M71-IRES-CRE mESCs were treated with epigenetic chemical compounds (1µM). mESCs were fixed 48h after treatment. The log10 data were ranked based on descending green fluorescence values (in µm²). The averages for TdTomato value=5.11 and GFP value=3.04. The results were compared to the TdTomato fluorescence control average for untreated mESCs=5.53 and GFP fluorescence control average of CRE transfected mESCs=3.73. GFP cut-off line of log10=3.5 in µm² was determined by adding together the maximum level of green base line value of 3.16 in the native mESCs and a standard deviation of 0.3687 (****p<0.0001, a one-way ANOVA; error bars indicate mean +/-SD)

Figure 48. Fluorescence and morphology of epigenetic compounds treated 5x21 M71 OR ROSA-MTMG;M71-IRES-CRE mESCs.

Epigenetic compounds (1µM) were added to 5x21 M71 OR transfected ROSA-MTMG;M71-IRES-CRE mESCs. The names and effect of the 3 example compounds are listed. A. S-(5’-Adenosul)-L-methionine (tosylate) (an ubiquitous methyl donor); B. Splitomicin (plays a role in silencing gene expression); C. Anacardic acid (an anti-inflammatory). Scale bar = 100µm.
CHAPTER 4. Reasons for M71 OR gene to be silenced in mouse embryonic stem cells

4.1 Expression of M71 OR transgenes in ROSA-MTMG;M71-IRES-CRE reporter system

According to the current model, OR genes are actively silenced by repressive epigenetic marks, such as H3K9me3 and H4K20me3, until a single OR is chosen to be expressed. When one OR gene is chosen out of a pool of 1400 OR genes in mice, the negative epigenetic marks are replaced with transcription initiation ones (H3K4me3). However, using ChIP-qPCR assay in ROSA-MTMG;M71-IRES-CRE mESC reporter system, we showed that the epigenetic repression possibly occurs in basal cells or is inherently present, which is earlier than suggested, since H3K4me3 active mark is absent and H3K9me3 and H4K20me3 inactive marks are present on the M71 OR promoter in mESCs. Based on these results, we attempted to de-repress the M71 promoter by removing the negative epigenetic marks by administering small chemical compound and epigenetic chemical libraries to our reporter. Regardless, we were unable to activate the M71 OR in mESCs, possibly due to the small probability of a single OR to be chosen out of 1400. To circumvent this issue, we used MouSensor technology (published in our lab), where we added enhancer elements to M71 OR transgene and introduced it to the reporter system, trying to increase the probability of expression of this particular OR. After the transfection of the transgene, we repeated the epigenetic chemical treatment, but we still failed in our attempts to force M71 OR in mESCs.

At this point, we consider two possibilities for our lack of success: 1) proteins necessary for OR expression, such as Lhx2, are not yet present at mESCs; 2) the presence of a negative regulatory sequence in the M71 OR. Initially, we probed the reporter cells for the presence of Lhx2 protein, using anti-Lhx2 antibody. We also examined a few options for negative regulation within the OR by investigating the role of an OR enhancer, promoter, or coding sequence in the OR expression
regulation in mESCs (Nguyen, 2007; Magklara, 2011). To test the role of the enhancer on the M71 OR expression, we made series of \textit{nx21} constructs to randomly integrate in the genome of our \textit{ROSA-MTMG;M71-IRES-CRE} mESC reporter line. Thus, we sought to use a series of \textit{nx21} enhancers: \textit{0x21, 5x21, 6x21, 7x21, 8x21, and 9x21} enhancer sequences to force the M71 OR choice in the \textit{ROSA-MTMG;M71-IRES-CRE} mESCs. Next, we asked the question if the DNA sequence of an OR can carry a code that represses the ability of ORs to express \textit{in vitro}. The diversity of the OR sequences would make it unlikely, however, homeodomain and OE2 within the OR promoters might have certain influence over the OR expression. For this purpose, we substituted the mouse M71 OR with another mammalian receptor, \textit{EF1\alpha}. Finally, we omitted the M71 OR sequence, by replacing it with a CRE recombinase or \textit{\beta2AR} GPCR sequence, to test if the OR CDS itself had negative impact over the M71 OR expression in mESCs.

To test all of these possibilities, we used transfection via AMAXA system (Lonza), where the created transgenes were expressed in the \textit{ROSA-MTMG;M71-IRES-CRE} reporter. In the reporter cells, if M71 OR was expressed, a TdTomato$\rightarrow$GFP shift in fluorescence would be observed. For quantification of the number of green cells expressed in the system, we used confocal microscopy and FACS analysis.

When we investigated the presence of Lhx2 in mESCs, a known transcriptional regulator involved in singular OR choice, there was a minimal Lhx2 expression (Figure 49). Next, we examined the possibility of the presence of a negative regulator in the M71 OR sequence. M71 OR expression in \textit{ROSA-MTMG;M71-IRES-CRE} mESCs could be forced by interfering with that
sequence. Using series of transgenes, we tried to pinpoint the presence of such a negative regulatory sequence within the M71 OR.

Before investigating the effect of these transgenes in the ROSA-MTMG; M71-IRES-CRE mESCs, as a proof of principle, we tested the expression of an exogenous CMV promoter-driven CRE recombinase with nuclear localization signal (NLS) in the reporter system (pOG23, Addgene plasmid #17736) (O'Gorman, 1997). The CRE recombinase was expressed in the mESCs. In 24h, it triggered the TdTomato→GFP shift in fluorescence (Figure 50). In addition, we performed an immunofluorescence staining assay, where we used an anti-CRE antibody (Novagene, Inc.) to examine the presence of CRE protein in the cells. We discovered that the presence of the protein was not only reflected in the form of green membrane immunofluorescence, but it was also retained inside the cells (Figure 51).
Figure 49. Expression of endogenous Lhx2 in $ROSA$-$MTMG;M71$-$IRES$-$CRE$ mESCs.

A. In $ROSA$-$MTMG;M71$-$IRES$-$CRE$ mESCs, red fluorescence reflects the presence of the membrane TdTomato in the cell membranes. B. Red fluorescence shows the presence of membrane TdTomato. An anti-Lhx2 antibody is used for the detection of the Lhx2 protein in the cells (Santa Cruz Biotechnology, Inc.). The green fluorescence shows a low level of Lhx2 expression. C. Using an anti-Lhx2 antibody, red fluorescence shows the presence of membrane TdTomato in the reporter. Green fluorescence detects the high expression of exogenous Lhx2 in the mESCs. Scale bar = 20µm.

Figure 50. TdTomato/GFP expression and cell morphology in $ROSA$-$MTMG;M71$-$IRES$-$CRE$ mESCs before and after exogenous CRE expression.

A. M71 OR/CRE recombinase are not expressed in untreated $ROSA$-$MTMG;M71$-$IRES$-$CRE$ mESCs. The mESCs are red, due to the presence of the membrane TdTomato in the ROSA26 locus. mESCs are trapezoid shape and grow in globular shape colonies (observed here). B. CRE recombinase was transfected in the $ROSA$-$MTMG;M71$-$IRES$-$CRE$ mESCs; the reporter system was proven functional. TdTomato was excised and GFP was expressed. The cells and colonies maintained their shapes. Scale bar = 20µm.
Figure 51. Detection of CRE protein, using an anti-CRE antibody 24h post transfection in ROSA-MTMG;M71-IRES-CRE mESCs.

We used an anti-CRE antibody (EMD Millipore Corporation) to detect CRE protein 24h post-transfection in mESCs. Red fluorescence shows the presence of membrane TdTomato in the mESCs. Green fluorescence reflected recombination in cells. Blue staining showed the presence of CRE protein within the ROSA-MTMG;M71-IRES-CRE cells. The results reveal that in 24h post-transfection CRE protein was expressed in the mESCs. Scale bar = 20µm.

Figure 52. EF1α promoter transgenes expressed in ROSA-MTMG;M71-IRES-CRE mESCs.

A, B. Untreated ROSA-MTMG;M71-IRES-CRE mESCs lack endogenous expression of the M71-IRES-CRE knock-in gene (absence of green cells in the mESCs). C, D. When EF1α substituted the OR promoter (M71 CDS stays), there was TdTomato→GFP shift. E, F. In EF1α transgenes, when OR CDS was replaced by another GPCR (β2AR_IRES-CRE), green mESCs were expressed. G, H. When OR promoter was replaced by EF1α and OR CDS was substituted by CRE, M71/CRE were expressed (TdTomato→GFP shift). Scale bar = 20µm.
The control experiments revealed that once the CRE gene was transcribed and translated, we were able to detect membrane green fluorescence and CRE protein expression in the mESCs. Next series of experiments aimed to determine the existence of a regulatory sequence within the M71 OR, based on the green cells presence or absence in the reporter system. We first examined the role of the OR promoter in mESCs, since previously we had shown that the M71 OR promoter is negatively epigenetically marked. To determine if by removing the promoter we could force M71 OR in the cells, we replaced the entire M71 promoter with 1kb EF1α heterologous promoter and created EF1α promoter_M71-IRES-CRE no NLS plasmid. The lack of NLS assures tighter regulation of the plasmid expression and diminishes leaky expression. In order to widen the parameters of the assay, in another construct, we replaced the M71 OR promoter with EF1α promoter and M71 CDS with CRE recombinase gene (EF1α promoter_CRE no NLS). Finally, we assembled a construct, where the M71 CDS was replaced by the β2-adrenergic receptor (β2AR), another GPCR (EF1α promoter_β2AR-IRES-CRE no NLS plasmid). Using a CRE recombinase sequence allowed us to determine the role of the promoter in the transgene expression. β2AR GPCR was a valuable candidate to replace the M71 OR CDS, because unlike ORs, β2AR traffics in cell lines and it expresses on the cell surface. Furthermore, β2AR serves as a surrogate for an OR in vivo, since it leads to the maturation of the OSNs, axon guidance and identity, and assists in gene choice; therefore, it is a good model system to study ORs (Figure 52). All transgenes were transfected as a circular or linearized plasmid. There was not a noticeable difference between the two types of constructs in the reporter system.

We confirmed our results by detecting the presence of green fluorescent cells via confocal microscopy and by counting the number of green cells through FACS analysis. The numbers of
green cells seen under the microscope relatively matched the numbers shown by the FACS analysis. Using the data from the FACS readout, we additionally calculated the number of green cells in reference to the total number of cells. The number green cells after exogenous CRE expression (pOG231 plasmid) into the ROSA-MTMG;M71-IRES-CRE reporter line was approximately 44% of the total counted cells: 4458 green cells in 10,000 mESCs. The results of the control experiment suggested that we were able to efficiently integrate the DNA into the genome. However, the number of green cells after the expression of EF1αpromoter_M71-IRES-CRE no NLS was comparatively low: 60 green cells in total of 10,000 cell. When EF1α promoter_CRE no NLS was expressed in the reporter system, a total of 184 green cells were observed. This number was reduced to 111 green cells for EF1α promoter_β2AR-IRES-CRE no NLS expression. There was a statistical significance when the number of green cells after EF1α promoter_CRE no NLS was compared to EF1αpromoter_M71i-IRES-CRE no NLS expression, showing that the presence of the OR promoter and CDS inhibits the OR expression in the reporter system. Even if only the OR promoter was substituted, but the GPCR sequence was preserved – EF1α promoter_β2AR-IRES-CRE no NLS – the number of green cells was almost twice the number of green cells after EF1α promoter_M71-IRES-CRE no NLS expression, pointing at negative influence over the OR expression of both M71 OR promoter and CDS in mESCs (Figure 53). However, a weakness of the assay was that the number of the expressed green cells was too low for us to be able to create stable cell lines expressing the transgenes.
Figure 53. Quantification of the number of express green cells post EF1α-transgenes transfection in ROSA-MTMG;M71-IRES-CRE reporter (no antibiotic selection).

A. After FACS analysis, we compared the number of green cells after the expression of CRE plasmid (pOG231), EF1α promoter_M71-IRES-CRE, EF1α promoter_CRE, and EF1α promoter_β2AR-IRES-CRE in ROSA-MTMG;M71-IRES-CRE mESCs. The biggest number of green cells were observed after pOG231 expression. When the M71 OR promoter is replaced by a heterologous promoter, but the M71 OR CDS is preserved, there was green cells expression, but at much lower level. When both of the OR promoter and the CDS being replaced with a heterologous promoter and either a CRE reporter or another GPCR. The CRE reporter only-based green cells expression was higher than the GPCR one (a larger number of observed green cells, after EF1α promoter_CRE expression in the reporter mESCs in comparison to EF1α promoter_β2AR-IRES-CRE expression) in the mESCs (***p<0.0005, a one-way ANOVA).

B. Same experiment as in panel A is featured in B, but we omitted the pOG231 green cell expression data from the figure. In this case, EF1α promoter_CRE expression (no M71 promoter and CDS) led to the highest number of expressed green cells. Next, EF1α promoter_β2AR-IRES-CRE expression gave a lower number of green cells and EF1α promoter_M71-IRES-CRE led to the lowest number of green cells after expression in the reporter mESCs (** p<0.01, a one-way ANOVA; error bars indicate mean +/-SD).
The low number of green cells led us to the next experiment, where we introduced MouSensor technology modified transgenes in the ROSA-MTMG;M71-IRES-CRE reporter line. Here, we examined the role of an enhancer element in the OR expression and the probability of an OR expression in mESCs. In mice, the larger the number of multimers for the enhancer (nx21), the higher the probability for an OR transgene expression. The lowest number of multimers used in animals contains 4x21 enhancer elements. For the in vitro assay, we used 5x21 multimer as the smallest probability enhancer, because expressing ORs in heterologous cells has been a challenge. Still as a control, we expressed a 4x21_M71 promoter_M71-IRES-CRE transgene in the mESC system, which did not lead to the presence of green fluorescent cells. For the nx21 series of constructs, we only examined the influence of the multimers for the transgenes, where the M71 OR promoter CDS was not omitted. We created two sets of transgenes (M71 OR promoter_M71-IRES-CRE and M71 promoter_CRE), where we had n=0,5,6,7,8, and 9x21 multimer enhancer elements. 9x21 was the highest number of multimer, we utilized in the assay, because the presence of a 9x21 enhancer in an OR transgene leads to over 90% expression of the OR in the MOE. If we forced the expression of the M71 promoter_M71-IRES-CRE or M71 promoter_CRE in the reporter mESCs, we would expect TdTomato→GFP shift in fluorescence. The data revealed that 9x21_M71 promoter_CRE, 9x21_M71 promoter_M71-IRES-CRE, 7x21_M71 promoter_M71-IRES-CRE, and 0x21_M71 promoter_M71-IRES-CRE expression resulted in a low number of green cells. Administering any other nx21 transgene did not lead to TdTomato→GFP shift in the reporter (Figure 5).

The assessment was done using fluorescent microscopy and FACS analysis. According to the FACS readings, 0.5 green cells out of 10,000 cells were observed for both
$0x21\_M71$ promoter_M71-IRES-CRE and $9x21\_M71$ promoter_M71-IRES-CRE expression, without antibiotic selection. There was no statistical significance between the expression of both plasmids (Figure 5). Our results showed that the number of $nx21\_M71$ OR expression led to an extremely small number of green cells in the mESCs. However, $0x21\_M71$ OR led to green cells expression, where $0x21\_M71$ OR is the parental contract, hence we are not certain in the role of the enhancer for M71 OR expression in mESCs. We needed to additionally enhance the system and increase the number of green cells after the transgene expression. One of the options to do this was by creating stable cell lines through puromycin selection.

While attempting to establish stable cell lines, we additionally created constructs, where we generated hybrid promoters between EF1α and M71 OR promoter. These hybrids included full length and truncated versions of the M71 OR promoter. It allowed us to analyze the effect of a strong heterologous promoter on the M71 OR promoter as a negative regulator for the transgene expression, as well as, to examine the effect of the deletion of a repetitive sequence within the OR promoter on the transgenes expression in mESC reporter system.
Figure 5. \( nx21 \) enhancer transgenes expression in \textit{ROSA-MTMG;M71-IRES-CRE} mESCs.

A, B. Lack of endogenous expression of the M71-IRES-CRE knock-in gene in the mESCs is depicted by the lack of green cell. C, D. \( 5x21 \) M71 promoter_CRE transgene transfection resulted in the lack of TdTomato\( \rightarrow \)GFP shift in the fluorescence, while \( 9x21 \) M71 promoter_CRE transfection resulted in a few non-clonal green cells in the mESCs. E, F, G. \( 0x21, 4x21, \) and \( 5x21 \) M71 promoter_M71-IRES-CRE transgenes were transfected in the reporter. There was lack of green cells expression. H, I. \( 7x21 \) and \( 9x21 \) M71 promoter_M71-IRES-CRE were introduced to the reporter system, which resulted in sporadic GFP expression. J. CRE plasmid (pOG231) expression resulted in a high number of green cells in mESCs. All images were taken 24h after transfection. Scale bar = 20\( \mu m \).

![Image of transgene expression in mESCs](image)

Figure 55. FACS quantification of GFP cell expression in \textit{ROSA-MTMG;M71-IRES-CRE} mESCs.

We quantified the number of green cells after \( 0x21 \) and \( 9x21 \) M71 promoter\_M71-IRES-CRE transfection in the reporter mESCs. The expression resulted in a small number of GFP-positive cells. These results differed from the confocal microscope observations, where we did not see any green cells expression in the \( 0x21 \) M71 promoter\_M71-IRES-CRE transfected mESC. The experiment was repeated twice (t-non parametric test).
4.2 Expression of M71 OR transgenes under antibiotic selection in *ROSA-MTMG;M71-IRES-CRE* reporter system

As a first step of creating stable cell lines using our series of transgenes and puromycin co-transfection in the *ROSA-MTMG;M71-IRES-CRE* reporter line, we performed a kill curve experiment, where for ten days, we administered 0.5-1.5µg/ml concentrations of puromycin to the mESCs. On day 11, we had approximately 10% of surviving cells at 1µg/ml concentration (Figure 56). It was the optimal concentration we used to create stable mESC lines. Following the kill curve, as a control experiment, we co-transfected the *ROSA-MTMG;M71-IRES-CRE* reporter line with a CRE plasmid (pOG231) and puromycin selection marker. After 24h, we picked 48 colonies and maintained them under puromycin selection. To prove maintained expression of the transgene in the system, we expanded the colonies and performed immunofluorescence (IF) staining against the CRE protein, using anti-CRE antibody. Our results showed that we achieved stable expression of the CRE protein, based on the green membrane fluorescence and the detection of CRE recombinase protein within the cells (Figure 57).

However, M71 OR transgene/puromycin co-transfection system did not function optimally. *nx21* transgenes expression under antibiotic selection in *ROSA-MTMG;M71-IRES-CRE* mESCs did not lead to a higher number of green cells (Figure 58, 59, and 60).
Figure 56. Puromycin kill curve to determine the puromycin concentration in ROSA-MTGG;M71-IRES-CRE mESCs.

Various puromycin concentrations were administered to the ROSA-MTGG;M71-IRES-CRE mESCs for 10 day period. On day 11, ~10% of the cells survived at 1µg/ml puromycin concentration. This concentration was chosen for co-transfection experiments, where a puromycin plasmid was used along with a M71 OR transgenes.

Figure 57. EF1α promoter-transgenes expression under puromycin selection in ROSA-MTGG;M71-IRES-CRE mESCs.

A. CRE (pOG231) and puromycin were co-transfected in ROSA-MTGG;M71-IRES-CRE mESCs, followed by 7-day antibiotic selection. Post selection, we used an anti-CRE antibody to detect CRE protein in the mESCs. Red fluorescence demonstrated TdTomato presence in the mESCs. Green fluorescence revealed TdTomato and GFP recombination. A. Stable CRE protein expression is detected (blue fluorescence). B. CRE/puromycin co-transfection in the mESCs was followed by antibiotic selection. On the eighth day, we isolated a single entirely green colony. C. EF1α promoter_M71-IRES-CRE/puromycin were expressed in the mESCs. The cells were selected. D. EF1α promoter_CRE transgene/puromycin were transfected into the mESCs. The cells were plated at low density and selected. We were able to detect a few entirely green clonal colonies (TdTomato→GFP shift). Scale bar = 20µm
Figure 58. FACS quantification of number of green cells, stably expressing EF1α-transgenes after puromycin selection in ROSA-MTMG;M71-IRES-CRE mESCs.

We compared the number of green cells after EF1α promoter_M71-IRES-CRE, EF1α promoter_CRE, and EF1α promoter_β2AR-IRES-CRE stable transfection under puromycin selection in ROSA-MTMG;M71-IRES-CRE mESCs. The most green cells were observed when both of the OR promoter and CDS were replaced. When only the M71 OR was replaced by a heterologous promoter, there was a green cell expression, but at much lower level (a one-way ANOVA; error bars indicate mean +/-SD).
We compared the number of green cells after EF1α promoter_M71-IRES-CRE, EF1α promoter_CRE, and EF1α promoter_β2AR-IRES-CRE transfection before and after puromycin selection in the mESCs. The largest number of green cells were result of OR promoter and CDS replacement. When M71 OR promoter was omitted and the CDS was preserved, there were green cells, but at much lower level (a one-way ANOVA; error bars indicate mean +/-SD).

A. CRE/puromycin were co-transfected in ROSA-MTMG;M71-IRES-CRE reporter. Post selection, we used anti-CRE antibody. Red fluorescence demonstrated the presence of TdTomato in the cells. Green fluorescence displayed the TdTomato→GFP shift, after CRE expression. CRE protein expression was proven (blue fluorescence). B. 0x21_M71 promoter_M71-IRES-CRE transgene/puromycin were co-transfected in the reporter and selected, which led to a low number of GFP-expressing cells. The results differed from previous data. C. 5x21_M71-promoter_M71-IRES-CRE/puromycin co-transfection did not result in TdTomato→GFP shift in the reporter, even after selection. D. 9x21_M71 promoter_M71-IRES-CRE/puromycin were expressed and selected. We did not observe green cells. Scale bar = 20µm.
In the last experiment in this assay, we created a hybrid between EF1α and M71 OR promoters (EF1α/M71 promoter). Previously, EF1α substitution of the M71 OR promoter led to a relatively easy transgene expression. We investigated if EF1α/M71 promoter allowed the strong heterologous promoter to overcome the transgene repression by the presence of the M71 OR promoter. We created four constructs. The constructs included EF1α full-length promoter/M71 OR full-length promoter and EF1α full-length promoter/ M71 OR truncated length promoter plasmids. The coding sequences in the four constructs included M71-IRES-CRE and CRE (Table 1). By truncating the M71 OR promoter for half of the constructs, we intended to pinpoint a possible negative regulatory sequence within. We omitted the first 226bp sequence in the N-terminus of the M71 full-length promoter, including a 150bp “dead” repetitive GC rich motif, which could serve as a barrier preventing PEV from occurring. As in previous experiments, the hybrid constructs/puromycin plasmid were electroporated into the reporter mESC line, selected for seven days, and analyzed for the presence of green mESCs. The number of green cells was alarmingly low (the experiment needed to be repeated), however, the initial results were quite interesting. The green cells number was: 13/10,000 cells after pOG231 positive control expression in the mESCs, which was extremely low in comparison to previous experiments; therefore, the results were not completely reliable; 3/10,000 cells after EF1α promoter_CRE expression, 0/10,000 cells after EF1α/M71 full-size promoter_M71-IRES-CRE and EF1α/M71 truncated promoter_M71-IRES-CRE expression; 3/10,000 cells for EF1α/M71 full-size promoter_CRE expression; and 1/10,000 cells for expression (Figure 61). Regardless the low number of green cells in the control experiment, the data of the assay reflected the presence of green cells in the absence of M71 OR CDS, but also that EF1α/M71 truncated promoter_M71-IRES-CRE allowed for GFP expression. It would be a good idea to understand the role of the
omitted repetitive sequence within the M71 OR promoter for the M71 OR expression in mESCs (Figure 62).

The collected data in the transgene assay was promising, however, our attempts to express M71 OR transgenes, in the ROSA-MTMG;M71-IRES-CRE reporter mESCs, met multiple challenges. For instance, we tried to clone the puromycin selection marker in the M71 OR transgenes to avoid co-transfection, but we were unsuccessful. Next, we attempted to add FRT-neomycin-FRT to the M71 OR transgenes. Because of that, we needed to destroy the neomycin sequence within the ROSA-MTMG;M71-IRES-CRE mESCs. We used CRISPR mediated deletion to frameshift the neomycin gene in the reporter mESCs.

4.3 Neomycin frameshift in mESCs by CRISPR

The mechanism of Clustered Regularly Interspaced Short Palindromic Repeats- mediated (CRISPR-Cas9) assay is based on the prokaryotic immune system, where a guide RNA helps CRISPR-associated protein (Cas9) to recognize and cut exogenous DNA (Jansen, 2002). In our assay, we used the system to cleave a portion of the neomycin gene within the ROSA-MTMG;M71-IRES-CRE mESCs. We aimed at using neomycin selection for the transgene expression assay, hence, we needed the neomycin gene to be removed from the mESCs genome
might have 

of green cells 

the 

truncated promoter_CRE led to 

/M71 full 

/M71 truncated promoter_CRE. EF1 

α 

6) EF1 

α 

/M71 truncated promoter 

α 

4) EF1 

plasmid 

W 

in 

Figure 6 

number of 

/M71 promoter truncated size_CRE transgene 

epression of the transgene, regardless the presence of the EF1 

puromycin were transfected 

TdTomato 

IRES 

for seven days 

p

IRES 

A. 

ROSA 

Figure 6 

uromycin were 

compared the number of green cells after 

ROSA 

Lack of 

confocal images data. EF1 

/M71 

- 

CRE 

MTMG;M71 

FACS q 

antibiotic selection in 

ROS A-MTMG;M71-IRES-CRE mESCs.

A. Lack of endogenous M71-IRES-CRE expression was observed in the ROSA-MTMG;M71-IRES-CRE cells (no green cells). B. EF1α/M71 promoter full-size_M71-IRES-CRE transgene/ 

puromycin were co-transfected in the mESCs. The cells were plated at low density and selected 

for seven days with no TdToma to→GFP shift. C. EF1α/M71 promoter truncated size_M71-IRES-CRE transgene/puromycin were transfected and selected in the reporter. No 

TdToma to→GFP shift was observed. D. EF1α/M71 promoter full-size_CRE transgene/ 

puromycin were transfected and selected. The omission of the M71 OR CDS did not lead to the 

expression of the transgene, regardless the presence of the EF1α in the hybrid promoter. E. 

EF1α/M71 promoter truncated size_CRE transgene/puromycin expression leads to a small 

number of green cells in the reporter. Scale bar = 20µm.

Figure 62. FACS quantification of green cells expressing EF1α/M71 promoter hybrid transgenes 

in ROSA-MTMG;M71-IRES-CRE mESCs.

We compared the number of green cells after transfection and antibiotic selection: 1) CRE 

plasmid (pOG231); 2) EF1α promoter_CRE; 3) EF1α/M71 full-size promoter_M71-IRES-CRE; 

4) EF1α/M71 truncated promoter_M71-IRES-CRE; 5) EF1α/M71 full-size promoter_CRE, and 

6) EF1α/M71 truncated promoter_CRE. EF1α/M71 full-size promoter_CRE and EF1α/M71 

truncated promoter_CRE led to TdToma to→GFP shift in the mESCs. The FACS did not match 

the confocal images data. EF1α/M71 full-size promoter_CRE expression led to a higher number 

of green cells, even though using confocal microscopy, we did not detect GFP fluorescence. We 

might have missed the green cells, due to their extremely low number.
We designed and created three different guide RNA sequences into X330 plasmid, using guidance sequences #1, 4, and 7 from the CRISPR-Cas9 neomycin databases (Zhang, 2013). According to our analysis, guidance sequence 1 was chosen since it was the first hit in the database, while guidance sequences 4 and 7 had the most homology to the neomycin sequence.

Once the guide RNA expressing plasmids were created, we use Lipofectamine®2000 transfection reagent (Invitrogen) to introduce them in the ROSA-MTMG;M71-IRES-CRE reporter system. Multiple experiments were performed, where we used the guiding sequence plasmids separately or as a cocktail together against the neomycin sequence.

After transfection, we plated the cells at low confluency and allowed them to recover for one week, in which time they grew from a single cell to a colony. After a week, we picked 96 colonies in 96-well plate. We let the cells become confluent and split them into a set of 96-well plates, where we added neomycin containing media. The concentration of the neomycin was picked based on data from the lab and the literature. Since the cells were not dying in the neomycin-containing media, we repeated the experiment multiple times. In addition, we used multiple 96-well plates to apply various neomycin concentrations. Still, we did not manage to destroy the neomycin gene and the ROSA-MTMG;M71-IRES-CRE mESCs remained fully confluent regardless the neomycin concentration.
CHAPTER 5. Investigating GPCR ability to traffic to the plasma membrane in an olfactory placode-derived cell line

ORs have been first cloned in 1991 (Buck, 1991). Currently, most of our knowledge about OR expression mechanisms, ligand binding, and ligand-receptor interactions is accumulated using a mouse model system due to the lack of an expression *in vitro* assay. Because of the lack of an *in vitro* model, most ORs have not been “deorphanized” yet, where ligands for most ORs remain unknown. An *in vitro* system would speed up the process of identifying new ligands and offer better perspectives of the mechanisms of OR choice.

Since ORs cannot be expressed in cells, in the past, we have worked on establishing *in vitro* assays, where we tried to force the M71 OR endogenous expression. Currently, regardless our previous attempts to express ORs by de-repressing the OR promoter in mESC reporter system, as well as with creating a series of transgenes to activate M71 OR in the reporter, we were unsuccessful in establishing a heterologous system to express ORs. However, in previous work from our lab, we have developed an *in vitro* assay, which allows us to track GPCRs ability to traffic in heterologous cells. All plasmids used for this project are listed in a table (Table 5). ORs belong to the GPCR family (seven-transmembrane receptors, 7TM) (Figure 4). We established the GPCR trafficking assay through transfecting olfactory placode 6 (OP6) cells with GPCR::GFP fusion proteins (Bubnell, 2013; Bubnell, 2015; Jamet, 2015). In the assay, if the GPCRs traffic to the cell membrane, we observe many actin extensions, filopodia, as a result of the trafficking of the GFP-fused plasmids to the surface. On the contrary, if the GPCRs did not traffic to the membrane, the filopodia were present, but not visible. We have previously shown, by staining the filopodia, using CellMask Deep Red Plasma Membrane Stain, that the filopodia are always present in the OP6 cells, but only visible when the GPCR::GFP protein expresses on
the cell membrane (Bubnell, 2015). The readout of the assay was a “positive” or “negative” answer for protein trafficking. If we observed a single digit filopodia number, we determined the protein as non-trafficking or poorly trafficking to the membrane. A double-digit number of visible filopodia indicated the localization of the protein on the cell membrane. We counted the number of visible filopodia in ten cells and averaged them. To account for GPCRs trafficking, we also considered the intensity of the expression of the protein on the cell surface and the morphology of the cells.

5.1 *In vitro* filopodia assay to monitor GPCR FL expression in OP6 cells

We examined the ability of multiple GPCRs FL to traffic to the plasma membrane. For this purpose, we introduced the GPCRs to the OP6 model system by using transfection expression. Ten images were taken per GPCR condition, where the filopodia number for ten cells for each GPCR were counted, averaged, and represented by filopodia count number (FC) average value.

Our results indicated that β2AR, type 2 serotonin (5-HT2a), and the type 2 dopaminergic receptor (DRD2) trafficked into the filopodia in OP6 cells, with an average number of 80, 74, and 55 filopodia respectively. In comparison, ORs, such as Olfr151 (M71 OR) (wild type and codon optimized versions) and OR1A1 (human OR) trafficked poorly to the plasma membrane, with an average number of 1, 1, and 0 filopodia respectively.

Mouse transmembrane-associated receptor 4 (mTAAR4), similarly to ORs, is expressed by OSNs and did not traffic to the plasma membrane. Additionally, vomeronasal type-1 receptor B2 (V1rb2), expressed in the vomeronasal did not traffic to the plasma membrane *in vitro*. ORs and other receptors correlated to the sense of smell, did not traffic to the surface *in vitro* (Figure 63).
Table 5: List of DNA plasmids used in OP6 cell assays.

<table>
<thead>
<tr>
<th>NUMBER</th>
<th>GPCR::GFP (ABBREVIATION)</th>
<th>GPCR NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>β2AR (FL)</td>
<td>Mouse β2 adrenergic receptor</td>
</tr>
<tr>
<td>2</td>
<td>5HT2a</td>
<td>Type 2 serotonin</td>
</tr>
<tr>
<td>3</td>
<td>DRD2</td>
<td>Type 2 dopaminergic receptor</td>
</tr>
<tr>
<td>4</td>
<td>V1rb2</td>
<td>Vomeronasal type-1 receptor B2</td>
</tr>
<tr>
<td>5</td>
<td>mTAAR4</td>
<td>Mouse transmembrane-associated receptor 4</td>
</tr>
<tr>
<td>6</td>
<td>MOR-EG</td>
<td>Olfactory receptor 151 (Ofr151)</td>
</tr>
<tr>
<td>7</td>
<td>M71 FL WT</td>
<td>Olfactory receptor 151 (Ofr151)</td>
</tr>
<tr>
<td>8</td>
<td>M71 FL codopt</td>
<td>Olfactory receptor 73 (Ofr73)</td>
</tr>
<tr>
<td>9</td>
<td>OR1A1</td>
<td>Olfactory receptor family 1 subfamily A member 1</td>
</tr>
<tr>
<td>10</td>
<td>ΔNt β2AR FL</td>
<td>N-terminus deletion β2 adrenergic receptor FL</td>
</tr>
<tr>
<td>11</td>
<td>ΔCt β2AR FL</td>
<td>C-terminus deletion β2 adrenergic receptor FL</td>
</tr>
<tr>
<td>12</td>
<td>ΔNt/ΔCt β2AR FL</td>
<td>N- and C-termini deletion β2 adrenergic receptor FL</td>
</tr>
<tr>
<td>13</td>
<td>M71 Nt→TM1 WT</td>
<td>M71 transmembrane domain 1 wild type</td>
</tr>
<tr>
<td>14</td>
<td>M71 Nt→TM1 codopt</td>
<td>M71 transmembrane domain 1 codon optimized</td>
</tr>
<tr>
<td>15</td>
<td>β2AR Nt→TM1/(D1129)</td>
<td>β2AR transmembrane domain 1</td>
</tr>
<tr>
<td>16</td>
<td>β2AR Nt→TM2</td>
<td>β2AR transmembrane domain 2</td>
</tr>
<tr>
<td>17</td>
<td>β2AR Nt→TM3</td>
<td>β2AR transmembrane domain 3</td>
</tr>
<tr>
<td>18</td>
<td>β2AR Nt→TM4</td>
<td>β2AR transmembrane domain 4</td>
</tr>
<tr>
<td>19</td>
<td>β2AR Nt→TM5</td>
<td>β2AR transmembrane domain 5</td>
</tr>
<tr>
<td>20</td>
<td>β2AR Nt→TM6</td>
<td>β2AR transmembrane domain 6</td>
</tr>
<tr>
<td>21</td>
<td>β2AR Nt→TM7 (FL)</td>
<td>β2AR transmembrane domain 7 (full length)</td>
</tr>
<tr>
<td>22</td>
<td>M71 codopt Nt→TM2</td>
<td>M71 codon optimized transmembrane domain 2</td>
</tr>
<tr>
<td>23</td>
<td>M71 codopt Nt→TM3</td>
<td>M71 codon optimized transmembrane domain 3</td>
</tr>
<tr>
<td>24</td>
<td>M71 codopt Nt→TM4</td>
<td>M71 codon optimized transmembrane domain 4</td>
</tr>
<tr>
<td>25</td>
<td>M71 codopt Nt→TM5</td>
<td>M71 codon optimized transmembrane domain 5</td>
</tr>
<tr>
<td>26</td>
<td>M71 codopt Nt→TM6</td>
<td>M71 codon optimized transmembrane domain 6</td>
</tr>
<tr>
<td>27</td>
<td>M71 codopt Nt→TM7 (FL)</td>
<td>M71 codon optimized transmembrane domain 7 (full length)</td>
</tr>
<tr>
<td>28</td>
<td>gapGFP</td>
<td>Membrane gap GFP</td>
</tr>
<tr>
<td>29</td>
<td>Ni DRD2 FL</td>
<td>N-terminus GFP type 2 dopaminergic receptor</td>
</tr>
<tr>
<td>30</td>
<td>VAMP</td>
<td>Vesicle-associated membrane protein</td>
</tr>
<tr>
<td>31</td>
<td>OR1A1 Nt→TM1</td>
<td>Olfactory receptor family 1 subfamily A member 1 transmembrane domain 1</td>
</tr>
<tr>
<td>32</td>
<td>mTAAR2 Nt→TM1</td>
<td>Mouse transmembrane-associated receptor 2 transmembrane domain 1</td>
</tr>
<tr>
<td>33</td>
<td>mTAAR3 Nt→TM1</td>
<td>Mouse transmembrane-associated receptor 3 transmembrane domain 1</td>
</tr>
<tr>
<td>34</td>
<td>mTAAR4 Nt→TM1</td>
<td>Mouse transmembrane-associated receptor 4 transmembrane domain 1</td>
</tr>
</tbody>
</table>

All DNA plasmids that were used for the *in vitro* assay in this project are listed in the above table. The plasmids have been designed and made by Paul Feinstein, Irena Parvanova, Jacklyn Bubnell, Sergio Bernal, and Delia Tomoiaga.
5.2 ΔNt, ΔCt and ΔNt/ΔCt β2AR FL trafficking expression on the membrane in OP6 cells

We tested the expression of multiple GPCRs FL and determined if they localized on the cell surface. We investigated why some GPCRs traffic, while others do not. Our previous analysis of β2AR FL revealed that individually deleting the N-terminus (Nt) and the C-terminus (Ct) for β2AR did not prevent β2AR from trafficking (Jamet, 2015). However, the FC diminished from 80 to 50 after the omission of either terminus. Nevertheless, simultaneously deleting both regulatory regions abolished trafficking completely, suggesting that neither one of the N- and C-termini influences the trafficking of the protein to the plasma membrane, but the absence of both abolishes it (Figure 64).

5.3 Transmembrane domains significance for GPCR trafficking in OP6 cells

Having observed that the N- and C-termini were possibly involved in trafficking, we intended to further dissect the GPCRs and understand what other aspects of their structure might play role in cell membrane expression. Based on the work from multiple labs, we examined the significance of TM domain truncations for trafficking for β2AR and ORs (Kobilka, 1988).

In Kobilka lab, the scientists investigated various truncation versions of β2AR transmembrane domains in correlation to their functionality. The scheme of their experiment gave us the idea to further understand the role of individual TMs in trafficking and set up a truncation series for β2AR where GFP is fused to the Ct portions of TM5 (β2AR TM1 to TM5), TM3 (β2AR TM1 to TM3), and TM1 (β2AR TM1), with the N-terminus: β2AR Nt→TM5, β2AR Nt→TM3, and β2AR Nt→TM1. We created the same truncations for M71 wild type and codon optimized versions: M71 WT Nt→TM5, M71 WT Nt→TM3 and M71 WT Nt→TM1; M71 codopt Nt→TM5, M71 codopt Nt→TM3, and M71 codopt Nt→TM1 (Table 6). We have observed that
C-terminus GPCR::GFP does not fluoresce outside of the cell (data not shown); therefore, we initially did NOT design and create β2AR, where fused GFP to the Ct portions of Nt→TM2, Nt→TM4 and Nt→TM6, because we expected that those truncations will have the C-terminus outside of the cell and it would not be fluorescent.

Out of the nine truncations for β2AR, M71 WT, and M71 codopt, only β2AR Nt→TM1 trafficked to the OP6 plasma membrane, even though not as well as β2AR FL, β2AR Nt→TM1 plasmid expression (β2AR FL: FC=80; β2AR Nt→TM1: FC=47). M71 codopt Nt→TM1 could not express on the membrane surface (FC=5). M71 WT Nt→TM1 (FC=13) somewhat traffics to the plasma membrane, but not nearly as well as β2AR Nt→TM1 (Figure 65). Since β2AR Nt→TM1 was the only truncation that expressed on the cell surface, we considered that TM1 domains might be important for membrane expression of GPCRs on the cell surface; therefore, we decided to further investigate the structure of the various GPCR TM1s (Figure 66 and 67).

5.4 TM1 and N-terminus significance for GPCR trafficking to the cell surface

β2AR FL traffics in heterologous cells, while M71 OR FL does not. Similarly, β2AR Nt→TM1 expresses on the cells membrane, while M71 Nt→TM1 does not. Based on the data, our observation is that TM1 plays a role in the trafficking and GPCR expression on the cell surface. In terms of cell surface expression, N-terminus, also seems to be a regulatory region. Based on these results, we interfered with the β2AR TM1 sequence or/and the Nt and noted the impact of the plasmid expression in OP6 cells.
Figure 63. Expressing GPCRs in OP6 cells filopodia assay.

A. A schematic representation of the GPCR coding sequence, fused to GFP (GPCR::GFP). B-J. GPCRs FL were transfected in OP6 cells. Each GPCR was represented in a separate panel. The borders of the panels were colored green or red based on the GPCR ability to express or not express on the plasma membrane. The upper right region of the panel represents a zoom-in image of the visible filopodia for each GPCR. Ten images were taken per GPCR; the number of the filopodia for ten cells were counted, averaged, and represented by the FC average value in the left lower corner of each panel. β2AR, 5HT2a, and DRD2 successfully trafficked to the plasma membrane, with an average count of 80, 74, and 55 filopodia respectively. V1rb2, TAAR4, MOR-EG OR, M71 OR (wild type), M71 OR (codon optimized version), and OR1A1 OR, on the other hand, remained in the cell, expressing maximum of FC average = 2. Scale bar = 20µm.
Figure 64. ∆Nt, ∆Ct, and ∆Nt/∆Ct full length β2AR::GFP expression in OP6 cells.

A, B, C. Nt, Ct, and Nt/Ct were deleted (∆Nt β2AR, ∆Ct β2AR, and ∆Nt/∆Ct β2AR) in β2AR FL. The deletion constructs were transfected in OP6 cells. Their ability to traffic to the cell membrane was arbitrated by the FC number, observed after the expression of each construct in OP6. For ∆Nt β2AR and ∆Ct β2AR constructs expression, FC = 50 and FC = 52 respectively (β2AR FL: FC = 80). For ∆Nt/∆Ct β2AR construct expression, FC = 2 in OP6 cells. Scale bars = 20µm.

<table>
<thead>
<tr>
<th>GPCRs</th>
<th>FULL LENGTH</th>
<th>TM1-5</th>
<th>TM1-3</th>
<th>TM1</th>
</tr>
</thead>
<tbody>
<tr>
<td>M71 OR WT::GFP</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>M71 OR codopt::GFP</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>β2AR::GFP</td>
<td>✓</td>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 6. M71 OR::GFP (wild type and codon optimized versions) and β2AR::GFP truncated series of plasmids.

Series of truncated plasmids for the three proteins, fused to GFP, were created. M71 OR (WT and codopt) and β2AR: transmembrane domains1-5 (TM1 to 5 = β2AR Nt→TM5), truncated transmembrane domains1-3 (TM1 to 3 = β2AR Nt→TM3), and truncated transmembrane domain1 (TM1 = β2AR Nt→TM1). Only β2AR FL and β2AR Nt→TM1 trafficked to the cell surface.
As β2AR Nt→TM1 traffics in vitro, we investigated if its trafficking is affected by Nt deletion (ΔNt). ΔNt β2AR→TM1 trafficking was impacted by the Nt deletion, for β2AR Nt→TM1 expression, FC=37, while for ΔNt β2AR→TM1, FC=15 (Figure 72). These results differed from the trafficking patterns of ΔNt β2AR FL in OP6 cells. ΔNt β2AR FL did not drastically affect the ability of the protein to express on the cell surface, FC=50 (Figure 68). Additionally, we asked if ΔNt β2AR→TM1 truly did not traffic in OP6 cells or there was another explanation for the lower number of observed filopodia. From the literature, we know that in other species, such as the fruit fly (Drosophila melanogaster), GPCRs are oriented with the Nt inside and the Ct outside the cells (Wicher, 2008). Our observations for the ΔNt β2AR→TM1 trafficking patterns were surprising, especially since ΔNt β2AR FL expression was not dramatically impacted in cells. Therefore, another option was that ΔNt β2AR→TM1 still trafficked in OP6 cell, but similarly as in Drosophila, the Nt was inside, while the Ct projected outside of the cell surface. However, GPCR::GFP does not fluoresce outside of the cell and we did not observe fluorescent filopodia (data not shown).
Figure 65. M71 OR WT Nt→TM1, M71 OR codopt Nt→TM1, and β2AR Nt→TM1 membrane expression demonstrated by a filopodia assay in OP6 cells.

A, B. M71::GFP Nt→TM1 (WT and codopt) were transfected in OP6 cells. Ct::GFP was a reporter in the assay. OR proteins remained trapped in the cells, unable to express on the plasma membrane. Small number of filopodia were observed for M71 WT Nt→TM1, where FC = 13 and M71 codopt Nt→TM1, where FC = 5. The red outlines of the figure panels represent the lack of trafficking of the truncation, while the zoom-in panel depicts the lack of membrane extensions. C. β2AR Nt→TM1 was transfected in OP6 cells (Ct::GFP) and expressed on the cell membrane, FC = 47. The green fluorescent outlines of the panels signify protein trafficking in OP6 cells. The zoomed-in image shows the presence of the visible filopodia. Scale bars = 20µm.
Figure 66. β2AR Nt→TM1 to β2AR Nt→TM7 truncation series expression based on filopodia assay in OP6 cells.

A-G. β2AR::GFP Nt→TM1 to TM7 (TM7=FL) plasmids (Ct::GFP), were transfected in OP6 cells. Out of all truncations, based on the number of visible filopodia, β2AR Nt→TM1 and β2AR Nt→TM7 (FL) constructs were expressed on the cell surface. The rest of the truncated plasmids were trapped inside the cells. Scale bar = 20µm.

Figure 67. M71 codopt Nt→TM1 to M71 codopt Nt→TM7 truncation series expression based on filopodia assay in OP6 cells.

A-G. M71::GFP codopt Nt→TM1 to M71::GFP codopt Nt→TM7 (TM7=FL) were transfected in OP6 cells (Ct::GFP). All M71 OR were stuck in the endoplasmic reticulum in the heterologous cells. The experiment re-confirms that ORs do not express in vitro. Scale bar = 20µm.
5.5 Non-permeabilization imaging assay to test GPCRs Nt/Ct orientation in the membrane

5.5.1 Establishing a non-permeabilization imaging assay

GPCR trafficking data in *Drosophila melanogaster*, as well as a study conducted on Endothelin B receptor, showed that GPCRs can be in multiple orientations in the plasma membrane, with their Ct inside, but also outside of the cell surface (Wicher, 2008; Köchl, 2002; Alken, 2009; Benton, 2006). For instance, ORs are oriented with Nt out/Ct in the plasma membrane in mammals, while ORs are oriented with the Nt in/Ct out of the plasma membrane in *Drosophila*. β2AR Nt→TM1 traffics in cells, but ΔNt β2AR→TM1 was relatively impaired from expression in the plasma membrane, while the FL versions traffic to the cells surface. Therefore, we asked if we misinterpreted the results for GPCRs when we monitored their trafficking in OP6 cells.

Firstly, we analyzed if ΔNt β2AR→TM1 traffics in OP6, regardless the lack of filopodia in OP6 cells, due to the positioning of the protein, where Nt is oriented in the intracellular space and the Ct::GFP is outside of the cells (Figure 68) (Lundin, 2008; Fukutani, 2011).

For positioning in the cell membrane, Speiss et al. defines Type I, II, III, and IV orientations for single-membrane spanning proteins (Figure 5). We did not use this nomenclature to describe multiple-transmembrane spanning proteins, but used Nt out/Ct in vs. Nt in /Ct for orientation purposes. However, these characteristics are only true when we discuss GPCR TM1s and FLs (not the rest of the truncations in the series). We based the assay on the Ct orientation, because of the Ct::GFP fusion.

To monitor the GPCR Ct::GFP orientation in the OP6 cells, we developed an assay, where we ensured that the plasma membrane of the plasmid transfected cells remained non-permeabilized.
Then, we examined the GFP expression on the cell surface, using a primary anti-GFP antibody and a secondary red fluorescent antibody against the anti-GFP antibody (Chart 1) (Zhuang, 2008).

To establish the assay, we conducted a series of controlled experiments. We tested a membrane bound protein, gapGFP, in the non-permeabilization assay. The gap protein led the GFP fluorophore to the cell membrane, which allowed for visible filopodia, but the gapGFP was not observed outside the cell surface (no red fluorescence was detected by the 2ry antibody). gapGFP was a control for Nt out/Ct in orientation of a plasmid in OP6 cells. Nt::GFP fusion of DRD2 was used as another control for Nt out/Ct in orientation, because we observed both filopodia expression and red fluorescence for Nt out/Ct in (Jeanneteau, 2004). Finally, a vesicle-associated membrane protein fused to GFP (VAMP::GFP), a protein that only traffics in Nt in/Ct out orientation, was tested as a control in the non-permeabilization imaging assay (Figure 69). Unlike GPCR Ct::GFP, VAMP Ct::GFP fluoresces outside of the plasma membrane.

VAMP::GFP fusion allowed for visible filopodia, despite the Ct::GFP fusion for the fluorophore. We do not fully understand why we do not detect Ct::GFP fluorescence, since we know that GFP is present (GFP is detected by the anti-GFP antibody in the non-permeabilization imaging assay).

5.5.2 β2AR and M71 TM truncations tested in the non-permeabilization imaging assay

Once the non-permeabilization imaging assay was established, we investigated the orientation of GPCRs, previously tested for trafficking in OP6 cells. As a result, ΔNt β2AR→TM1, which did not traffic according to the OP6 cells filopodia assay, was detected by the presence of red fluorescence based on the GFP presence outside of the membrane; therefore, when expressed on the cell surface, ΔNt β2AR→TM1 localized in Nt in/Ct out orientation. Hence, certain GPCRs,
that we have tested for cell surface expression in the filopodia assay, might still be trafficking, but in Nt in/Ct out orientation, which is the opposite of the expected Nt out/Ct in orientation for GPCRs (Figure 70).

As mentioned, we did not initially design and create β2AR Nt→TM6, Nt→TM4, and Nt→TM2 truncation plasmids, because GFP fluorescence is not detected in the extracellular side. However, at this point, we created these truncation versions and used the non-permeabilization imaging assay to test their cell surface expression. When tested the GFP presence outside of the cells in the assay for β2AR Nt→TM1 to β2AR Nt→TM7 (FL) expression, the data showed that β2AR Nt→TM1 trafficked to the OP6 cell surface in both orientations (Nt out/Ct in and Nt in/Ct out). β2AR Nt→TM2 and β2AR Nt→TM6 trafficked in Nt out/Ct in orientation. β2AR Nt→TM3, β2AR Nt→TM4, and β2AR Nt→TM5 remained stuck in the endoplasmic reticulum and did not express on the cell membrane in either orientation, while β2AR Nt→TM7 (FL) only trafficked in Nt out/Ct in orientation (Figure 71).

It is unclear why certain truncation proteins traffic and others do not. It might be the protein folding, where β2AR Nt→TM1, 2, and 6 truncations were folded properly and expressed on the membrane, while other truncations, β2AR Nt→TM3 and β2AR Nt→TM5, were misfolded and degraded. Also, there is a possibility that particular domains are required for the localization of the GPCRs to the plasma membrane.
Figure 68. β2AR Nt→TM1 and ΔNt β2AR→TM1 truncations expression in OP6 cells.

A. β2AR Nt→TM1 trafficked to the cell membrane in OP6 cells. A large number of visible filopodia were observed (FC = 37). B. (ΔNt β2AR→TM1 ability to express on the plasma membrane was impacted and the FC was reduced. Scale bar = 20µm.

Figure 69. Control experiments for establishing membrane non-permealization antibody-staining assay.

A, A’, A”. Gap membrane protein allows GFP membrane expression. There were visible filopodia, but no red antibody staining fluorescence. B, B’, B”. Nt::GFP-tagged DRD2 FL was used to study protein trafficking to the plasma membrane (visible green fluorescent filopodia); red fluorescence showed the presence of GFP outside of the OP6 cells. GFP-DRD2 was a gift from Jean-Michel Arrang (Addgene plasmid # 24099; http://n2t.net/addgene:24099; RRID:Addgene_24099). C, C’, C”. Vesicle-associated membrane protein (VAMP) has a Ct, always oriented outside of the cell. Green filopodia accounted for the VAMP::GFP trafficking in OP6 cells, while the red fluorescence shows Ct GFP antibody detection of VAMP protein outside the cells. Scale bar = 20µm (the figure is courtesy of Sergio Bernal).
Sequentially, we investigated M71 codopt Nt→TM1 to M71 codopt Nt→TM7 (FL) orientation in the plasma membrane of OP6 cells. According to the data, we found out that M71 codopt Nt→TM1 trafficked in Nt in/Ct out orientation, whilst M71 codopt Nt→TM2 to M71 codopt Nt→TM7 (FL) remained stuck in the ER (Figure 75). Similar results were observed for M71 WT Nt→TM1, 3, 5, and 7 (FL), where M71 WT Nt→TM1 was the only truncation that trafficked to the cell surface in Nt in/Ct out orientation (data not shown). For the observed proteins, GPCR TM1 seemed to always express on the OP6 cells membrane in both or at least in Nt in/Ct out orientation (Figure 72).

5.5.3 Other GPCR TM1s and FLs trafficking abilities tested in non-permeabilization imaging assay

We further investigated the ability of other GPCR TM1s to traffic in either orientation to the plasma membrane of OP6 cells. We observed that the GPCR TM1s for multiple TAARs, as well as for OR1A1, were all in either both or Nt in/Ct out orientation in the OP6 membrane (Figure 73 and 74). For all the tested GPCR FL, no Nt in/Ct out orientation was observed in the OP6 assay. At this point, we are not sure if the Nt in/Ct out orientation is a GPCR TM1s phenomenon in vitro or if the OP6 machinery is not able to support both orientations of GPCR FLs. Our observations on the GPCR TM1s orientations in the membrane of a heterologous cell line were interesting. However, there are two important questions: 1) do TM traffic in multiple orientations in vivo; 2) do GPCR FL traffic in multiple orientations in vivo?

5.5.4 β2AR Nt→TM1 orientation, based on non-permeabilization imaging assay in vivo:

We conducted in vivo studies, using the same non-permeabilization imaging assay. We used
mouse tissue of β2AR Nt→TM1 expressing animals. After we stained the tissue, we were able to detect the GFP presence on the outside of the cilia of the MOE (Chart 1, Figure 75).

In the image, the green dots depicted the GFP fluorescence inside the section, while the red dots represented the outside staining for GFP presence. The fact that the green and the red dots do not merge can be due to the curvature of the sample. The yellow dots represent the presence of receptors in both orientations in the MOE. We were pleased with the results and are currently adjusting the protocol to investigate the β2AR and M71 OR orientations in the MOE, in order to answer the ultimate question – do GPCR FLs exist in both orientations in vivo?
Figure 70. Membrane orientation of ΔNt β2AR→TM1 truncation in OP6 cells.

A. In comparison to β2AR Nt→TM1 plasmid, ΔNt β2AR→TM1 truncation did not travel to the cell surface in OP6 cells, while β2AR FL, ΔNt β2AR FL trafficked well in heterologous cells. Based on the data, we asked if the ΔNt β2AR→TM1 protein trafficked, but in a different orientation. B. Using an anti-GFP antibody, we examined the orientation of ΔNt β2AR→TM1 truncation in the plasma membrane. The Ct::GFP of ΔNt β2AR→TM1 was detected outside of the cell surface, shown by the red fluorescence of the 2ry anti-GFP antibody. C. Merged image depicting ΔNt β2AR→TM1 truncation trafficking to the cell membrane, where we saw a small number weak filopodia and intense anti-GFP antibody staining. Scale bar = 20µm.
Figure 71. Membrane non-permeabilization antibody-staining assay to investigate β2AR Nt→TM1 to β2AR Nt→TM7 truncation series expression in OP6 cells.

A-G. β2AR::GFP Nt→TM1 to β2AR::GFP Nt→TM7 (FL), where Ct::GFP is a reporter, when transfected in OP6 cells. β2AR Nt→TM1 and β2AR Nt→TM7 (FL) were expressed on the cell surface, based on the number of visible filopodia. The rest of the truncated plasmids did not traffic. A. When we used an anti-GFP antibody, we observed that the β2AR Nt→TM1 plasmids trafficked to the plasma membrane, but in opposite of the expected orientation (Nt in/Ct out). Since it is a non-permeabilization assay, the detected GFP was located outside of the cell surface. Scale bar = 20μm.

Figure 72. Membrane non-permeabilization antibody-staining assay to investigate M71 codopt Nt→TM1 to M71 codopt Nt→TM7 truncation series expression in OP6 cells.

A-G. M71 codopt Nt→TM1 to M71 codopt Nt→TM7 (FL), where Ct::GFP is a reporter, transfected in OP6 cells. M71 OR truncated plasmids and the M71 FL remained stuck in the endoplasmic reticulum. A. When an anti-GFP antibody was used to test the orientation of M71 codopt Nt→TM1 to M71 codopt Nt→TM7 (FL) truncations, we observed that the M71 TM1 plasmid trafficked in the cells, but in opposite of the expected orientation (Nt in/Ct out). In the assay, the detected GFP was located outside of the cell surface. Scale bar = 20μm.
CHAPTER 6. DISCUSSION

Our ChIP-qPCR data showed that the M71 OR is marked by the repressive epigenetic chromatin marks H3K9me3 and H4K20me3 in the ROSA-MTMG;M71-IRES-CRE ESC reporter line (a cell type similar to the inner cell mass cells within blastocysts). We substantiated the presence of the “off” marks with the beta-globin (HBB) gene, which we hypothesized should be “off” as well. We believe that the repressive marks are correlate to the ORs not being expressed at this early developmental stage, and as such, having these marks in such an early cell lineage would suggest that “off” marks are not part of the OR singular gene expression mechanism. To better understand the ChIP-qPCR data, future experiments should investigate the status of promoter methylation for other ORs in ROSA-MTMG;M71-IRES-CRE reporter line. However, the lack of CRE expression after random integration of the nx21-M71-IRES-CRE transgenes suggest that ORs are poorly activated in mESCs.

One reason for the early occurrence of negative epigenetic marks on ORs might result from the clusterization of ORs on chromosomes in MOE, where approximately 30 OR clusters exist in the mouse genome. The presence of OR clusters might trigger the negative repression of an OR until a single one is chosen to be expressed. To further examine this phenomenon, we looked at the published “off” marks amongst nine OR singletons genes (OR receptors not located into clusters). Our analysis suggests that some of the non-cluster ORs possess a degree of repressive marks, while others lack them in the mature OSNs. It has been published that TAAR genes do not possess repressive marks (Johnson, 2012). Could repressive marks have value in the OR clusters? One possibility is that repressive marks are present in the OR cluster to reduce the opportunity of non-chosen ORs to be expressed. The diverse epigenetic landscapes tied to OR expression of different
OR loci suggest a complex mechanism of regulation. Based on our results and inconsistencies in the literatures, it seems unlikely that epigenetic regulation of ORs is directly involved in singular gene choice.

To further examine the rigidity of epigenetic marks on ORs in ESCs, we attempted to de-repress M71 OR. Here we tested the effect of 4860 chemical compounds from small compound libraries on the M71 OR expression by high-throughput chemical screening in ROSA-MTMG;M71-IRES-CRE mESCs. Prestwick, MicroSource, and Lopac®1280 chemical libraries compounds were administered in the reporter cell line. The activation of M71 OR, determined by TdTomato→GFP shift in fluorescence, did not need to be maintained; CRE expression solely needed to occur and delete the reporter gene. Unfortunately, we found no evidence for excision of the reporter gene in mESCs and thus by inference no CRE expression. We additionally attempted to remove or modify the negative marks by using an 175 epigenetic compounds library, purchased from Cayman Chemical; we were still not able to induce CRE expression as measured by TdTomato excision followed by the expression GFP. The lack of activation of M71 OR after treatment with thousands of compounds chemical may be a reflection that even if the “off marks” were removed from the ESCs, this extremely early developmental stage might simply not express the necessary transcription factors to regulate OR gene transcription. A curious experiment would be to perform the high-throughput chemical screen on a later stage developmental cell line, such as olfactory precursor OP6 cells. In addition, multiple compounds might have to be tested together in both mESC reporter system and OP6 cells. If the OR expression could be influence in vitro, then it would be a faster method than in vivo transgenesis to understand transcription and translation mechanisms of OR expression.
Figure 73. GPCRs expected orientations in the cells.

A. A single-transmembrane orientation in Nt out/Ct in the cells. B. A single-transmembrane orientation in Nt in/Ct out of the cell surface.

Figure 74. Investigating the Nt/Ct orientations of GPCR TM1s in OP6 cells.

A-F. Using a membrane non-permeabilization antibody-staining assay, we tested multiple GPCR TM1s orientation in OP6 cells. All of them trafficked in both orientations with Nt out/Ct in and Nt in/Ct out of the cell surface. Scale bar = 20µm.
Figure 75. Investigating the C-terminus orientations of β2AR Nt→TM1 in vivo.

We applied non-permeabilization antibody-staining assay, using MOE tissue. The green fluorescence dots depicted the Ct::GFP inside the cells, the red fluorescence dots presented the antibody staining against the Ct::GFP outside of the cell surface, and the yellow fluorescence showed that β2AR Nt→TM1 was Nt out/Ct in and Nt in/Ct out orientation in the cell membranes. Scale bar = 20µm (the image is courtesy of Sergio Bernal).
In a final attempt to de-repress M71 OR by chemical screenings in mESCs, we introduced an nx21 enhancer driving an ectopic M71-IRES-CRE transgene into the system (where $n = \text{number of } 21\text{bp multimers of the H-element olfactory enhancer}$). The enhancer element is known to increase an OR probability to be chosen in vivo (D’Hulst, 2015). We introduced a 5x21_M71-IRES-CRE transgene in our ROSA-MTMG;M71-IRES-CRE reporter and administered the epigenetic compound library. Surprisingly, after the treatment, no GFP could be visualized in the mESCs. These experiments further supported the idea that ORs are tightly repressed in mESCs and they cannot be activated by a single compound added to the system in our experiments. Our endogenous M71-IRES-CRE allele is located in an OR cluster. Activation of this M71 allele might lead to the activation of some or the rest of the ORs across the OR cluster, or vice versa. Therefore, the reason for the lack of M71 expression in mESCs might be a functional of a cluster regulatory processes rather than at the level of a single OR locus.

In addition to using chemical compound libraries to force M71 OR choice in the ROSA-MTMG;M71-IRES-CRE mESCs reporter, we used series of M71 OR transgenes, containing M71 OR 5’UTR, intron, 3’UTR, and polyA, which transgenes we transfected in the mESCs, aiming again at forcing the expression of M71 OR in the reporter system. The transgenes, tested the role of OR enhancers (nx21 sequence), OR promoter, and coding sequence on OR expression in mESCs. As an example, we replaced the OR promoter with the EF1α promoter and/or the M71 coding sequence in the transgenes with CRE or β2AR-IRES-CRE). If CRE was expressed by the transgene, then we would observe TdTomato→GFP shift, depicted by GFP expressing cells.

We have shown that the nx21 enhancer sequence functions in nearly all integrating loci as nearly 100% of the 5x21 M71 OR transgenes are expressed in mice. These results suggest that even if
the transgene lands in closed chromatin, the nx21 enhancer opens the chromatin and allows gene expression. It is known that neighboring enhancers from other genes alter the normal function of a closely located transgene – an effect commonly referred to as position effect variegation (PEV). Surprising our results were different. During the epigenetic screen on randomly inserted 5x21_M71 promoter_M71-IRES-CRE transgenes in our mESC reporter line, we did not observe any green cells expression and therefore PEV was not observed. One possibility is that a part of the transgene actively leads to its own suppression: an enhancer that attracts “off marks” as a target of silencing in vitro. Our data with the 0x21 enhancer_M71 promoter_M71-IRES-CRE transgene revealed expression in a few cells in our mESC reporter line, while 5x21 and 6x21 transgenes did not reveal any GFP mESC cells. In contrast, our 7x21 and 9x21 transgenes contained enough multimers to overcome the repression as GFP mESCs were more readily found.

In addition to enhancer effects, the OR sequence might have a negative effect over the OR transgene expression in vitro. If this hypothesis is true, 5x21_M71 promoter_β2AR-IRES-CRE transgene for instance might NOT attract “off marks”. For that reason, we also created constructs, where we omitted the OR promoter and/or CDS sequence, trying to force the expression of β2AR in the reporter mESCs. We also investigated if the presence of a strong promoter, such as EF1α mammalian cell promoter, could overcome the lack of M71 OR expression in the ROSA-MTMG;M71-IRES-CRE mESCs. We discovered that EF1α-transgenes expression, with or without the M71 CDS, caused TdTomato→GFP shift in the reporter system. After transient transfections with a EF1α promoter_M71-IRES-CRE Tg, the system revealed 60 green mESCs out of 10,000 cells. Replacing the M71 CDS in the EF1α promoter_CRE CDS Tg,
yielded 184 green mESC out of 10,000 cells, while β2AR CDS instead of the M71 CDS, yielded 111 green mESCs out of 10,000 cells. Higher number cells for CRE vs. β2AR-IRES-CRE CDS could reflect the lower efficiency of the IRES sequence at producing CRE. Overall, the absence of the M71 promoter and CDS allowed for expression of the transgenes in the reporter. The data suggested that both of them play a role in suppressing the OR expression in vitro. In the future it would be interesting to compare a stable transgene for EF1α promoter_β2AR-IRES-CRE in mESC colonies with a stable transgene EF1α promoter_M71-IRES-CRE mESC colonies. Would β2AR and/or M71 protein be stable in mESCs?

The significance of the OR promoter for the OR expression was further shown by creating transgenes, where hybrid EF1α promoter/M71 OR promoter drives the transgenes expression. In the cases of EF1α promoter/M71OR full-length promoter_CRE and EF1α promoter/M71OR truncated promoter M71-IRES-CRE expression, we observed TdTomato→GFP fluorescent shift in the reporter system. The presence of the strong heterologous promoter and omission of the M71 CDS allowed for M71 OR expression in mESCs, regardless of the M71 OR promoter presence in the transgene. In the EF1α promoter and M71 promoter chimeras where a GC-rich repetitive sequence within the M71 OR promoter was deleted, the transgene was expressed, regardless the presence of the M71 CDS. When EF1α promoter was present and the M71 CDS was omitted, GFP positive cells were observed. Thus, there is a clear negative influence between the presence of both the M71 OR promoter and CDS on the repression of the M71 OR expression in the reporter mESCs.
Overall, our transient transfections of M71 transgenes into the mESC reporter yielded small numbers of the GFP-expressing cells. Mixing a puromycin selectable marker with our M71 transgenes did not significantly improve on the number of GFP-expressing cells. The low number of GFP-positive cells did not allow for more detailed analysis of the regulatory roles of the enhancer element, OR promoter and coding sequence.

In addition, while attempting to force the M71 OR/CRE expression in mESCs (TdTomato→GFP shift), untreated ROSA-MTMG;M71-IRES-CRE mESCs remained red; leaky M71-IRES-CRE from our knock-in allele as never observed. In contrast, many genes are characterized with minimal leaky transcription in embryonic stem cells, even if the protein is not translated. However, that was not observed for the M71 OR gene in mESCs. Perhaps OR mRNA is not expressed in mESCs because of instability of the mRNA molecules. Even if some mRNA molecules for the M71 OR are expressed, they might be degraded due to the fact that this particular OR is not expressed in olfactory neurons.

Establishing a transgene-screening assay, not only aimed at understanding how ORs are expressed in a singular fashion, but also at establishing an in vitro system where ORs are forced to express. For the moment, we were unable to set up this particular assay, however, we succeeded in establishing another in vitro assay, using an olfactory placode cell line (OP6). This assay focused on investigating GPCR expression, trafficking and orientation on the plasma cell membrane.

We established a GPCR trafficking assay, using OP6 cells as a model system. OP6 cells possess membrane extensions called filopodia, which are present at the cell membrane, but only visible if
a fluorescent protein is expressed on the plasma membrane (Bubnell, 2013; Jamet, 2015; Bubnell, 2015). Using the assay, we have shown that GPCRs with C-terminus-fused GFP (Ct::GFP), such as β2AR::GFP, traffic to the OP6 cells membranes, which results in a high number of GFP-labeled filopodia, while ORs, such as full length M71::GFP (FL), remain trapped in the cells and do not label the OP6 filopodia. Furthermore, we showed that similar results were observed when β2AR Nt→TM1::GFP and M71 OR Nt→TM1::GFP, where we observed GFP-labeled filopodia for β2AR Nt→TM1 truncation, while we did not for M71 OR Nt→TM1. These results revealed that TM1 might partially be responsible for the β2AR capability to traffic to the cell membrane. Deletion of the Nt of the β2AR→TM1 (ΔNt β2AR→TM1) impacted the β2AR Nt→TM1 ability to traffic in OP6 cells. In contrast, the ΔNt β2AR::GFP FL trafficked to the plasma membrane (Jamet, 2015). The absence of GFP-labeled filopodia of OP6 cells expressing ΔNt β2AR→TM1 led us to examine the GPCR TM1 orientation in the OP6 plasma membrane.

To test the GPCR TM1 orientation in OP6 cells, we established an anti-GFP immunostaining assay. According to the results, ΔNt β2AR→TM1 trafficked in the OP6 cells, but in N-terminus in/C-terminus out orientation. As a consequence, we have defined two orientations for GPCR TM1s in the OP6 membrane: N-terminus out/C-terminus in (Nt out/Ct in) and N-terminus in/C-terminus out (Nt in/Ct out). Understanding TM1 orientation in cells and in vivo, will allow for better insight of the mechanism of GPCR expression on the cell surface. Nt in/Ct out orientation might be a phenomenon true only for TM1 truncations. In addition, we observed that β2AR Nt→TM1 traffics in both orientations in vivo, but at the moment, we do not have any evidence that GPCR full-lengths expresses in multiple orientations in the membrane in vivo or in vitro.
CHAPTER 7. CONCLUSIONS

The main conclusions of my studies are: 1) There is no M71 OR expression in ROSA-MTMG:M71-IRES-CRE mouse embryonic stem cells due to the absence of transcription initiation epigenetic marks, such as H3K4me3, and the presence of negative epigenetic marks, H3K9me3 and H4K20me3; 2) Approximately 5000 chemical compounds in two concentrations, which include 175 epigenetic chemical compounds, were not able to drive OR repression in mESCs; 3) OR promoter and coding sequence negatively impact OR expression in mESC expression system; 4) Surprisingly, there was a weak, but real affect when the olfactory choice enhancer element (x21) was added to an M71-IRES-CRE transgene in vitro; OR expression could now be weakly observed, suggesting that the x21 sequence was involved in OR expression in non-olfactory cells; 5) The first ~60 amino acids of a GPCR (the Nt and TM1) impact trafficking to the plasma membrane of OP6 cells; 6) GPCR TM1s traffic in Nt out/Ct in and Nt in/Ct out orientation in the OP6 cells plasma membrane; 7) β2AR Nt→TM1 expresses in both orientations in vivo.
REFERENCES


