Epigenetic Regulation of Olfactory Receptor Genes and Dedifferentiation of an Olfactory Placode Cell Line

by

Rutesh Niranjanbhai Vyas

Faculty Advisor: Dr. Robert P. Lane

A Dissertation submitted to the Faculty of Wesleyan University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Middletown, Connecticut May 2019
To My Loving Parents, Prafulla and Niranjan and My Dearest Son, Krish...
Acknowledgements

I would like to thank Prof. Dr. Robert P. Lane for his excellent mentorship. It was a lifetime opportunity, best research experience and a great honor to work under his guidance. I appreciate his enthusiastic support and necessary directions to conduct the great science in his laboratory. He provided countless hours helping me develop my scientific writing and professional presentation skills. He helped me to think scientific problems critically, practice experimental precision, and his tremendous efforts shaped my scientific career. I really enjoyed scientific communications with him and learned a lot from him during my graduate studies at Wesleyan University.

I would like to thank my committee members, Dr. Scott Holmes, Dr. Michael McAlear, and Dr. Amy MacQueen, for their time and guidance during laboratory rotations, committee meetings, department seminars, course works and scientific communications. I appreciate their insightful comments and suggestions on annual progress reports that helped me focus on my research, advance my science knowledge, and inspire me to think critically. My special thanks to Dr. MacQueen for microscopy and Dr. Taylor for BSL2 facility.

I am especially thankful to have excellent laboratory members. Diane has collaborated on my projects. Joyce has been very supportive, I enjoyed working with her. Thanks to Dr. Kilinc for teaching techniques and scientific
communication. James and Spencer have been excellent. Thanks to Gazhia for continuing dedifferentiation project.

I am thankful to faculty members who supported and advanced my knowledge during graduate studies: Dr. Donald Oliver, Dr. Helen Treloar, Dr. Rich Olson, and all other faculty members. I appreciate the support provided by administration (Ms Anika Dane, Ms Maureen Snow, Mr Horace Chambers), graduate office (Ms Cheryl-Ann Hagner and Ms Barbara Schukoske), and international student office (Ms Janice Watson). Special thanks to my friends Zeliang, Anand, Anna, Brandon, Lina, Karen, and Raghav for all the fun.

I would like to thank my undergraduate mentor, Prof. Dr. D.J. Sen, ex-mentor Dr. M.K. Bhattacharjee, my ex-boss, Prof. Dr. D.A. Williams, my current boss, Dr. Chay Kuo, my uncle and aunt (Mr L.V. Upadhyaya, Usha Upadhyaya), my family friends (Mr J.A. Patel and Kailash Patel, Mr M.K. Shah and Lina Shah), for their support. I am glad to work with post-doctoral fellows (Dr. David, Dr. Abdi, Dr. Buckley, Dr. Naffa’A, Dr. Ramamoorthy, and Dr. Neves, at Duke University.

I would like to express my deepest and enduring gratitude to my family for their endless support and encouragement throughout my life. My mother, Prafullaben, and my father, Niranjanbhai, have supported me to study abroad, and taking care of my Son, Krish, which allowed me to focus on my research. Krish is full of fun, very loving, and my inspiration. This thesis would not have been possible without my family’s scarifies and support. My special thanks to my brother, Amit, for his support. I love you Krish...
# TABLE OF CONTENTS

**CHAPTER 1: INTRODUCTION**  ............................................................................................................................................. 13

1.1  **Introduction to the Olfactory Sensory System** ........................................................................................................ 13
  1.1.A  Biological Significance of Olfaction .......................................................................................................................... 14
  1.1.B  Chemosensory Systems in Various Animals .............................................................................................................. 15
  1.1.C  The Mouse Olfactory System .................................................................................................................................. 20

1.2  **Structure and Function of the Mouse Olfactory Epithelium** ....................................................................................... 25
  1.2.A  Anatomy of the Mouse Olfactory System .................................................................................................................... 25
    1.2.A.1  Exterior Structures and the Epithelial Layer ........................................................................................................ 25
    1.2.A.2  The Glomerular Layer and Olfactory Processing .............................................................................................. 27
  1.2.B  Non-Neuronal Cells in the Mouse Olfactory System .................................................................................................. 28
    1.2.B.1  Sustentacular cells .................................................................................................................................................. 29
    1.2.B.2  Microvillar cells ...................................................................................................................................................... 29
    1.2.B.3  Bowman's Duct and Gland cells .......................................................................................................................... 30
    1.2.B.4  Other Cell Types of the Olfactory Mucosa ........................................................................................................... 31
  1.2.C  The Olfactory Sensory Neurons of the OE .............................................................................................................. 32

1.3  **Development of the Olfactory Epithelium** ................................................................................................................... 34
  1.3.A  Embryonic Development .............................................................................................................................................. 34
    1.3.A.1  Olfactory Placode (OP) Development ................................................................................................................... 34
    1.3.A.2  Generation of the Olfactory Epithelia ..................................................................................................................... 36
    1.3.A.3  Development of the Olfactory Bulb ........................................................................................................................ 38
  1.3.B  Basal Cells ...................................................................................................................................................................... 40
    1.3.B.1  HBCs and Developmental Potential ....................................................................................................................... 41
    1.3.B.2  Globose Basal Cells and Developmental Potential ................................................................................................ 43
  1.3.C  OSN Maturation ............................................................................................................................................................... 45
  1.3.D  OSN-Glomerulus Targeting ............................................................................................................................................ 48
    1.3.D.1  Axon Convergence as an Organizing Principle ...................................................................................................... 48
    1.3.D.2  Axon Guidance Molecules and the OSN Targeting Process .................................................................................. 50
    1.3.D.3  Glomeruli Development ........................................................................................................................................... 53
  1.3.E  Regeneration of the OE ..................................................................................................................................................... 55
    1.3.E.1  Biological Problem (Environmental Insult) ............................................................................................................ 55
    1.3.E.2  Regeneration Models ................................................................................................................................................ 57
    1.3.E.3  Sequence of Events During Tissue Regeneration ................................................................................................ 58
  1.3.F  Epigenetic Regulation of OSN Specification and Differentiation .................................................................................. 61
    1.3.F.1  Epigenetic Transitions During Neuronal Development ............................................................................................ 61
    1.3.F.2  Elaboration on Two Key Epigenetic Regulators ................................................................................................... 68

1.4  **Odorant Receptors and OSN Development** .................................................................................................................. 74
1.4.A.1 The Odorant Receptor .................................................................74
1.4.A.2 Odor Stimulation and Coding .....................................................75
1.4.A.3 OR-specific Signal Transduction ................................................77
1.4.A.4 OR Genome Organization and Evolution .....................................79
1.4.B OR Gene Regulation is Critical to Olfactory Function ......................81
1.4.B.1 Monogenic OR Expression as an Organizing Principle .................81
1.4.B.2 OR protein Functions in Glomeruli Targeting .............................84
1.4.C The Genetics of OR Gene Regulation ..........................................86
1.4.C.1 Spatial Patterning of OR Expression ..........................................86
1.4.C.2 Temporal Patterning of OR Expression .......................................88
1.4.C.3 OR Promoter Structures and OR Enhancer/LCR like Functions ....89
1.4.D Mutually Exclusive OR Expression as an Epigenetic Problem ..........91
1.4.D.1 Allelic Exclusion ........................................................................92
1.4.D.2 Monogenic Expression ...............................................................97
1.4.D.3 Chromatin Factors: The Role of LSD1 and G9a in Singular Odorant
Receptor Gene Regulation .................................................................105
1.4.D.4 Establishment of Mutually Exclusive OR Expression .................110

1.5 The OP6 Model System for Study of Odorant Receptor Gene
Regulation and Olfactory Sensory Neuronal Development .....................115

CHAPTER 2: LYSINE-SPECIFIC DEMETHYLASE-1 (LSD1) DEPLETION
DISRUPTS MONOGENIC AND MONOALLELIC ODORANT RECEPTOR (OR)
EXPRESSION IN AN OLFATORY NEURONAL CELL LINE ....................... 163

2.1 Introduction .......................................................................................165

2.2 Methods ..........................................................................................170
2.2.A Mouse OP6 Cell Preparation and Immunofluorescence ..................170
2.2.B LSD1 Knockdown by RNAi ............................................................171
2.2.C cDNA Analyses ............................................................................172
2.2.D Random Sampling Simulations to Estimate OR Complexity ..........173
2.2.E Western Blot Analyses ...................................................................174
2.2.F Chromatin Immunoprecipitation (ChIP) ........................................174
2.2.G RNA FISH ..................................................................................175

2.3 Results .............................................................................................176
2.3.A LSD1 Knockdown in OP6 Cells .....................................................177
2.3.B LSD1 Depletion does not Prevent OR Switching during OP6 Clonal
Expansion ............................................................................................180
2.3.C LSD1 Depletion does not seem to Alter OR Representation in OP6 Cell
Populations .........................................................................................184
2.3.D LSD1 Depletion Disrupts Monoallelic and Monogenic OR Expression in OP6 Cells .......................................................................................................................... 186
2.3.E H3K4 and H3K9 Methylation States at OR Loci in LSD1 Depleted cells. 193

2.4 Discussion .................................................................................................................. 197

2.5 Concluding Remarks on Chapter Two ...................................................................... 205

CHAPTER 3: G9A PERTURBATION TRANSFORMS AN Olfactory Neuronal Cell Line To A Pluripotent State Capable of Self-Organization .................................................. 212

3.1 Introduction .................................................................................................................. 215

3.2 Materials and Methods ............................................................................................. 222
  3.2.A OP6 and 293T Cell Cultures .................................................................................. 222
  3.2.B CRISPR-Cas9 Gene Targeting ............................................................................. 222
  3.2.C cDNA Analyses ................................................................................................... 223
  3.2.D Immunocytochemistry ......................................................................................... 223
  3.2.E Image Analysis ..................................................................................................... 225

3.3 Results ........................................................................................................................ 226
  3.3.A Transformation of OP6 Cells Following Introduction of the G9a CRISPR-Cas9 Cassette .............................................................................................................. 227
  3.3.B CRISPR-Cas9 Appears to Perturb G9a Expression without Gene Deletion ......................................................................................................................... 231
  3.3.C G9a Perturbation is Correlated with Reorganization of Heterochromatin ............................................................................................................................... 236
  3.3.D Gene Expression Analysis in Transformed OP6 Cell Populations .............. 240
  3.3.E Immunofluorescence for Stem Marker Proteins .............................................. 243
  3.3.F Initial Observations on the Differentiation Potential of Transformed OP6 Cells ......................................................................................................................... 245

3.4 Discussion .................................................................................................................... 249

CHAPTER 4: FUTURE INVESTIGATIONS OF OR REGULATION AND OSN DEDIFFERENTIATION ........................................................................................................ 272

4.1 Introduction ................................................................................................................. 272
  4.1.A What is the Mechanism of CRISPR-Cas9 Mediated G9a Perturbation? 275
  4.1.B Has G9a Perturbation caused Chromatin Reorganization? ......................... 279
4.1.D Has G9a Perturbation Activated Endogenous Cell Cycling Networks?
286
4.1.E Has G9a Perturbation Activated Stem Cell Networks? ................. 290
4.1.F Do Transformed OP6 Colonies Exhibit Tumorigenic Properties? ....... 295

4.2 Further Characterization of Transformed Colony Potential ............. 297
4.2.A Are Transformed OP6 Cells Pluripotent? .................................... 300
4.2.B Are Transformed OP6 Colonies Capable of Self-organized Development? ........................................................................... 306

4.3 Future Studies on Epigenetic Regulation of OR Genes ................. 311
4.3.A Does G9a Perturbation Impact OR Regulation in OP6 Cells? ......... 314
4.3.B Does LSD1 Overexpression Impact OR Regulation in OP6 Cells? .... 317
4.3.C Additional Predictions of an Iterative Epigenetic Model for OR Selection 320

4.4 Summary of the Thesis .................................................................... 326
List of Figures

Chapter 1

Figure 1. The Mouse Olfactory System ................................................................. 22
Figure 2: Various Cell Types and Markers of the Olfactory Epithelium ..................26
Figure 3: Axonal Connectivity between the OE and the OB: .................................28
Figure 4: Embryonic Development of the Olfactory Placode ...............................35
Figure 5: Timeline of the Olfactory System Development ....................................37
Figure 6: The Olfactory Bulb Development ..........................................................39
Figure 7: HBC Developmental Potential ...............................................................42
Figure 8: GBC Developmental Potential ...............................................................44
Figure 9: OB Axonal Targeting .............................................................................53
Figure 10: OR Activity Dependent Axon Sorting for Glomerular Development .....55
Figure 11: Process and Timeline of the Epitheliogenesis Post-MeBr Injury ..........60
Figure 12: Transcriptional Activator and Repressor Roles of G9a .........................73
Figure 13: The Olfactory Receptor Signal Transduction Pathway .........................78
Figure 14: The “Lomvardas Model” for OR Gene Regulation ..............................112

Chapter 2

Figure 1. LSD1 mRNA and Protein Knockdown in OP6 Cells .............................180
Figure 2. Colony switching assays indicate comparable levels of OR complexity in knockdown versus control colonies .................................................................183
Figure 3. RNA FISH with Individual and Pooled OR Probes ...............................189
Figure 4. Two-color RNA FISH on LSD1-depleted Cells Using Pairs of OR Probes. .........................................................................................................................192
Figure 5. Chromatin Immunoprecipitation (ChIP) to Measure Changes in H3K4 and H3K9 Levels at OR Loci in LSD1-depleted OP6 Cells. .......................................................... 197

Chapter 3

Figure 1. G9a-targeted CRISPR-Cas9 Treatment Transforms OP6 Cell Morphology and Growth. .................................................................................................................. 230

Figure 2. G9a Expression is Heterogeneous within Transformed Colonies........ 233

Figure 3. Correlation between G9a Depletion and Heterochromatin Reorganization. ......................................................................................................................... 239

Figure 4. Gene Expression Analysis in Treated and Untreated OP6 Populations for a Panel of Pluripotent and Cell Progression Gene Markers................................. 242

Figure 5. Transformed Colonies Express Pluripotent Marker Genes, as well as Marker Genes from the Three Canonical Germ Layers........................................ 244

Figure 6. Transformed Colonies Differentiate and Self-organize into Putative Developmental Structures....................................................................................... 247

Chapter 4

Figure 1: 3D Chromocenter Phenotype and Transcriptional Activity of Satellite Repeats .......................................................................................................................... 282

Figure 2: RNA-FISH Analysis using Antisense Large T antigen Probe ............ 289

Figure 3: Transformed Colonies Express Markers from Three Germ Layers...... 301

Figure 4: Differentiation of Stem-like Colonies into Glands and Neurons ........ 305

Figure 5: Overexpression of LSD1 in OP6 and LSD1 KD Cells........................... 319

Figure 6. Impact of LSD1 Knockdown on DNA Methylation and DNMT1 in OP6 Cells.................................................................................................................................. 322

Figure 7. Impact of LSD1 Knockdown on Co-REST Protein Levels in OP6 Cells. 324

Figure 8. Impact of LSD1 Knockdown on Lamin-B Receptor (LBR) Protein Levels in OP6 Cells............................................................................................................... 325
List of Tables

Table 1. Degenerate-OR PCR product complexity in knockdown versus control colonies. ................................................................. 185

Table 2. Degenerate-OR PCR product overlap between knockdown and control colonies. ................................................................. 186

List of Publication

Lysine-specific demethylase-1 (LSD1) depletion disrupts monogenic and monoallelic odorant receptor (OR) expression in an olfactory neuronal cell line.

Vyas RN, Meredith D, and Lane RP.

Molecular and Cellular Neuroscience, 2016.

Volume 74, July 2016, Pages 58-70.
Abstract

Function of the mammalian olfactory system depends on specialized olfactory sensory neurons (OSNs) that each express only one allele ("monoallelic") of one odorant receptor (OR) gene ("monogenic"). The lysine-specific demethylase-1 (LSD1) protein removes activating H3K4 or silencing H3K9 methylation marks in a variety of developmental contexts, and is thought to be important for proper OR regulation. Most of the focus in the field has been on a potential "activating" function for LSD1; e.g., in the demethylation of H3K9 associated with the expressed OR allele. In this thesis, we show that depletion of LSD1 in an immortalized olfactory-placode-derived cell line (OP6) results in a significant increase in multigenic and multiallelic OR transcription per cell, while not seemingly disrupting the ability of these cells to activate new OR genes during clonal expansion. These results are consistent with LSD1 having a role in silencing additional OR alleles, as opposed to being required for the activation of OR alleles, within the OP6 cellular context. We propose a model for initial OR gene selection, whereby management of both H3K9 methylation (by enzymes that add this mark at ORs; e.g. G9a and GLP) and H3K4 methylation (by enzymes that remove this mark at ORs; e.g. LSD1) are required for reduction of the number of competing OR candidate genes for activation. This model predicts that disruption of either enzyme (G9a to add H3K9 methylation or LSD1 to remove H3K4 methylation) might shift the equilibrium towards increased H3K4
/decreased H3K9, thereby increasing the probability of multiple OR activation events per cell. To further test this hypothesis, we perturbed G9a expression using the CRISPR-Cas9 technique in order to test whether perturbation of G9a contributed to an equilibrium shift towards decreased H3K9 methylation (and increased probability of multiple OR activation events per cell, as with the LSD1 perturbation). We made several observations in this experiment that suggest a much broader developmental role for G9a in the OSN lineage. We observed that CRISPR-treated OP6 cells transform/de-differentiate into stem-cell like colonies that exhibit global disruption of heterochromatin, expression of stem-like markers, and ultimately, differentiation and organization of progeny cells into developmental structures. In this thesis, I will review the field of olfactory system development (Chapter 1), discuss published results indicating a role for the LSD1 protein in regulating OR expression during olfactory sensory neuronal (OSN) development (Chapter 2), discuss unpublished results suggesting a role for the G9a protein in regulating OSN development (Chapter 3), and finish with a discussion of questions raised by this work and future experimental directions inspired by preliminary results (Chapter 4).
Chapter 1: INTRODUCTION

1.1 Introduction to the Olfactory Sensory System

Our brain senses a variety of information about its surroundings through sensory systems. Context dependent environmental stimuli activates specific sensory neurons that transform this environmental information into action potentials, which are further interpreted by higher brain centers. In turn, responses to this information in terms of perception, memory, learning, cognition, emotion, and behavior result from context-dependent inputs from sensory neuronal activity. Animals possess numerous sensory systems to probe the physical world; e.g. a visual system for detecting patterns of light, a somatosensory system for detecting physical contact, an auditory system for detecting patterns of sound, a gustatory system for sensing taste, and an olfactory system for distinguishing odors (Brownell et al., 1997; Chaudhari et al., 2010; Firestein et al., 2001).

In this introductory chapter to my thesis work, I describe the anatomy of the olfactory system, including the development and differentiation of olfactory sensory neurons (OSNs), that give animals the capability to recognize, distinguish, and respond appropriately to tens of thousands of biologically relevant odorants in the environment. This introduction will largely focus on the role of odorant receptor (OR) proteins in OSN function, and I provide
background on the genetic and epigenetic mechanisms that are thought to be important for the development of proper OR and OSN function. The main questions addressed in this thesis concern the role of epigenetic factors involved in (a) regulating OR gene expression, critical for specifying OSN function, and (b) regulating OSN development, critical in the context of establishing and maintaining a robust olfactory system. I have addressed these questions using a model cell line that I will introduce as a useful resource for isolating regulatory and developmental stages.

1.1.A Biological Significance of Olfaction

The sense of smell is the primary sense by which many organisms detect food sources, mates, and predators (Shepherd et al., 2004; reviewed in Hoover et al., 2010; Ache et al., 2005). In mammals, olfaction provides the ability to detect and discriminate tens of thousands of chemical stimuli (odors) within the environment (Malnic et al., 1999; Uchida et al., 2000, Zhang and Firestein, 2002). Olfaction elicits a range of critical animal behaviors, such as appetitive, aversive, social, emotional and reproductive behaviors essential for survival (Li et al., 2015, Zou et al., 2015, Yang et al., 2010, Logan et al., 2012). Loss of olfaction leads to mal-adaptive outcomes, such as anosmia (inability to smell), or hyperosmia and hyposmia (increase and decrease ability to smell, respectively), which negatively impacts fitness and quality of life (Croy et al., 2014, Godoy et al., 2015). Loss of olfaction is also an early sign of many neurodegenerative
diseases (Doty et al., 2012), therefore, understanding mechanisms important for
the development and maintenance of the olfactory system may provide insight
into therapeutic intervention in early stages of these diseases.

1.1.B Chemosensory Systems in Various Animals

There are three major chemosensory systems utilized by animals: gustation,
somatosensation, and olfaction (Chandrashekar et al., 2006; Nelson et al., 1996,
Axel et al., 2005; reviewed in Ma et al., 2007). Each of these chemosensory
systems transduces various environmental signals through different families of
chemoreceptors, as well as through different signal transduction pathways
(Zufall and Munger et al., 2016). These systems each have distinct anatomical
structures and distinct connections to the brain, so they elicit different
functional and behavioral responses (Mayerhof and Korsching et al., 2009).
Unlike the visual and auditory sensory systems that detect stimuli from a fixed
receptive field, chemicals have limitless structural possibilities, requiring a
different organization and set of strategies for properly interpreting these
environmental cues (Doty et al., 2015).

Vertebrate gustatory chemosensation is mediated through taste
receptors located in the taste buds. Taste buds transmit chemosensory
information via nearby nerve fibers that relay taste qualities to the brain (Doty
et al., 2015, Chen et al.,). There are five major taste modalities: salty, sour, sweet,
bitter, and savory (or umami) (Carleton et al., 2010). Salty and sour taste
sensations are detected through ion channels; sweet, bitter, and umami tastes are detected through G protein-coupled receptor (GPCR) taste receptors (reviewed in Doty et al., 2015).

Somatosensory chemosensation is mediated through nociceptor and thermoreceptors found in many sites within the body, including the GI tract, arteries, mucous membranes, and muscles (Watson, Paxinos and Puelles, 2012). Chemoreceptors in the muscle respond to acidity (e.g. lactic acid buildup during anaerobic exercise), chemoreceptors in mucous membranes of the skin warn us about irritating chemicals (Guyenet, Stornetta and Bayliss et al., 2010), enteroendocrine cells of the GI tract detect nutrients and toxins, and respiratory chemoreceptors in the arteries of the neck measure circulating CO$_2$ and O$_2$ levels (Kandel, Schwartz, Jassell, Siegelbaum, Hudspeth 2013). The trigeminal somatosensory system in the facial/mouth regions detects pressure, pain, chemical irritants, and temperature sensation (reviewed in Van der Cruyssen, 2017), including detection of menthol, camphor, mustard oil, capsaicin, and diallyl sulfide (onion) (Viana et al., 2010; Gerhold and Bautista et al., 2009; Silver and Finger et al., 2009).

Chemosensory systems in insects (fruit flies, honeybees, mosquitoes, spiders, butterflies, etc.) detect a wide range of volatile and soluble chemicals, including adenine nucleotides, sugars, peptides, amino acids, and gases (e.g., carbon dioxide and oxygen) in picomolar to micromolar concentrations (Mayerhof and Korschling et al., 2009). Blood-feeding insects (e.g., mosquitoes)
express chemoreceptors for detecting carbon dioxide (CO₂) gas, which is an important cue for the presence of its host (Chaisson et al., 2012). Many insects, such as honeybees and ants, are highly dependent on odor identity for discrimination between friend and foe (Rogers et al., 2013) and for orientation between food sources and nest (Pahl et al., 2011). Chemosensation in insects occurs through a hairy sense organ called sensilla (Steinbrecht et al., 2007). Olfactory sensilla are located in the antennae and maxillary palp on the head (Voshall et al., 1999; reviewed in Carlson, 2015). Gustatory sensilla are present on the body surface (e.g. mouth, legs, cerci, wings, tentacles and ovipositor in females) (Falk et al., 1975, Villella et al., 2008; reviewed in Stocker, 1994). Each sensillum is filled with potassium and protein-rich lymph fluid, as well as cuticles containing four types of odorant receptor neurons (OSNs) (Steinbrecht at al., 1996). Chemicals from the environment dissolve in the lymph and reach the dendrites of sensory neurons through small pores in the cuticles. The olfactory system in insect antennae contains olfactory sensory neurons (OSNs) expressing odorant receptors (ORs) that convey odorant information to the antennal lobe (Martin et al., 2011). Antennal lobes are composed of densely packed neuropils (glomeruli), where OSN axons synapse with the projection neurons and local inhibitory interneurons (Nagayama et al., 2014; Shipley et al., 2009; reviewed in Imai, 2014) that project to higher brain centers, such as the mushroom body and lateral horn (Christensen et al., 2002). Insect chemoreceptors are ligand-gated ion channels; as discussed subsequently, this
contrasts the use of G-protein-coupled receptors (GPCRs) utilized in mammalian olfaction (Brody et al., 2000; Wicher et al., 2012). The total repertoire of chemosensory receptors in insects is smaller (<100 ORs/GRs in a typical insect) as compared to mammals (>1,000 ORs in a typical mammal, such as mouse) (Robertson et al., 2006; The Honeybees Genome Sequencing Consortium, 2007).

In fish, chemoreceptors are distributed throughout the body, but are concentrated in the lips and barbels (Hara et al., 1992). Chemosensory-mediated behaviors, such as homing, predator avoidance, reproduction, foraging, and feeding, are central to the lifestyle and life cycle of the fish species (Valentincic et al., 2004). Fish olfactory organs are often inconspicuous, positioned symmetrically on the head, lying just beneath the dorsal surface and near the eyes. This organ consists of an anterior nostril through which water enters and a posterior nostril through which water leaves the olfactory chamber (Hamdani et al., 2007). OSNs are located in the folded, leaf-like structure called the olfactory rosette that resembles a miniature rose (Ngai et al., 1993), consisting of two layers of epithelium and an intervening layer of connective tissue (Fishelson et al., 2010; Kermen et al., 2013).

The fish olfactory epithelium (OE) contains three different types of OSNs: ciliated sensory neurons, the microvillus sensory neurons, and crypt cells (Kasumyan et al., 2004). Each of these specialized cell types expresses different receptor types, elicits different behaviors, and project to different areas of the olfactory bulb. The ciliated sensory neurons express the main ORs (~100 genes),
which mediate reproductive and alarm behaviors. The ciliated OSNs are wired to
the dorsomedial olfactory bulb (Hamdani et al., 2007; Alioto and Ngai, 2005).
The microvillus sensory neurons express V2Rs (~6 genes), which detects amino
acids. The microvillus OSNs mediate feeding behaviors and project to the lateral
olfactory bulb (Bertmar et al., 1981; Yoshihara et al., 2009). The crypt cells
express V1Rs (~56 genes), which respond to amino acids, bile acids, and
reproductive pheromones. The crypt OSNs mediate mating and alarm behaviors
and project to the ventromedial olfactory bulb (Hino et al., 2009; Sato et al.,
2005, 2007; Hamdani and Doving, 2006; reviewed in Karmen et al., 2013).

Fish taste buds are also distributed throughout the body, including the
lips, gills, pharynx and body surface, as well as within the oral cavity (Fishelson
et al., 2004; Kapsimali et al., 2011; Krueger et al., 1976; Barlow et al., 2015).
Typically, fish possess a fleshly, muscular area on the roof of their mouth, called
the palatal organ, which contains millions of taste buds (Reutter et al., 1974).
Fish express two types of taste receptors, T1R and T2R, that like the mammalian
taste system, signal the presence of specific food qualities (Ishimaru et al., 2005;
Oike et al., 2007). Taste reception on the body surface is supported by
microvillus cells innervated by the trigeminal nerve, which signals the presence
of competitors, predators and food sources (Sbarbati et al., 2003).

Amphibians (e.g. Xenopus laevis) have two nasal cavities, one exposed to
air (principal cavity) and the other exposed to water (middle cavity) (Parsons et
al., 1967; Rehorek et al., 2000). While there are a few noteworthy differences
concerning post-synaptic connectivity within higher olfactory processing centers (Woodley et al., 2014; Ache et al., 2005), the organization and development of amphibian chemosensory systems are similar to mammals, which also contain two principle organs – the nose (for volatile chemosensation) and the mouth (for detection of soluble substances). Mammals, like their aquatic cousins, also utilize two types of taste receptors (T1Rs, T2Rs) and three types of odorant receptors (ORs, V1Rs, V2Rs), however, mammals possess an additional organ, the vomeronasal organ (VNO), in which the bulk of V1R/V2R receptors are expressed that mediate social (e.g., pheromonal) types of behaviors (Dulac et al., 2003; Stowers et al., 2015; Keverne et al., 2002).

1.1.C The Mouse Olfactory System

The mouse olfactory system is located within the nasal cavity and is responsible for detecting volatile odorants in the environment (Figure 1). It consists of the olfactory epithelium (OE) containing olfactory sensory neurons (OSNs). OSNs are bipolar neurons consisting of multiple dendrites (cilia), each with a single large axon that connects to the brain (Cuschieri and Bannister., 1975a; Graziadei and Monti Graziadei, 1979; Buck and Axel, 1991). The dendrites of OSNs line the surface of the nasal cavity where they come in direct contact to the air and environmental odorants (Carr and Farbman, 1993). OSNs are vulnerable as a consequence of exposure to the environment, and these neurons must be replaced throughout the life of the organism. In order to replenish OSNs, a
population of stem-like cells, the Globose Basal Cells (GBCs), is maintained in the olfactory epithelium (Moulton et al., 1974; Graziadei and Monti Graziadei, 1979, Schwob, 1995). As described in much greater depth below (1.3: Development of the Olfactory Epithelium), GBCs differentiate into immature neuronal progenitors (INPs), which further give rise to fully differentiated, mature OSNs (Graziadei et al., 1978; Schwob et al., 2002). The axons of OSNs are extended to the first olfactory processing center of the brain, the olfactory bulb (OB), where each type of OSN connects to interneurons at a targeted glomerulus (Ressler et al., 1994; Vassar et al., 1994; Mori et al., 1999; Mombaerts et al., 1996; Wang et al., 1998). Thus, OSNs provide a direct connection between the environment and the brain.

How does the mouse brain make sense of the large range of odors from its environment? As with other animals, the odorant molecules from the environment are detected by OSNs through odorant receptor (ORs) proteins (Malnic et al., 1999). ORs are G protein-coupled receptors (GPCRs) with seven trans-membrane domains embedded in OSN cell membranes (Buck and Axel, 1991). The odorant molecules physically interact with ORs in the OE, and these receptor-ligand interactions are coupled to signal transduction that leads to a depolarization of stimulated OSNs (Firestein et al., 2001; Balluscio et al., 1998). Depolarization via OR stimulation is therefore the mechanism by which the nose communicates the presence of the specific odorants to the OB (Buck et al., 1999; reviewed in Krupp, 2010; Ronnett, 2002).
The range of odorant recognition is enabled by a large repertoire of ~1,400 OR genes that are clustered at multiple chromosomal locations in the mouse genome (Zhang et al., 2008; Young 2002). Remarkably, each olfactory sensory neuron is functionally specialized by virtue of expressing a single allele (“monoallelic”) of one and only OR gene (“monogenic”) (Chess et al., 1994). Therefore, there are ~1,400 different types of OSNs in the mouse nose, each expressing one type of OR and therefore specialized to detect a narrow range of odorant chemistry specified by its expressed OR.

Figure 1. The Mouse Olfactory System

(A) Schematic of the mouse head. The main olfactory epithelium (MOE) located dorsal-posterior axis of nasal cavity. Neurons from MOE project axons to the olfactory bulb (OB). Neurons from the Vomeronasal Organ (VNO) project axons to accessory olfactory bulb (AOB). The septal organ (SO) and Grueneberg ganglion (GG) are located anterior to the MOE (Figure published in Spehr et al., 2009). (B) The nasal cavity stain with eosin. Nasal lumen (white), nasal septum (middle), nasal turbinate (purple stain), MOE is bathed in the nasal turbinate and medial septum (Figure adapted from Herbert et al., 1999).
In order to maximize sensitivity to a particular odor quality, the OSNs expressing a given OR gene are broadly dispersed across the surface of nasal epithelia. One interesting developmental question is: how do OSNs otherwise broadly distributed in the nose, yet responsible for the detection of a specific odorant, communicate as a group the presence of that odor? The answer that has emerged is that the axons of OSNs expressing the same OR converge to a single common glomerulus in the OB, thus provide a cumulative signaling of that odor to one physical location in the brain (Momaberts et al., 1996; Wong et al., 1998; reviewed in Axel et al., 2005). A complex “smell” will consist of a unique combination of odorant chemicals, each of which will stimulate their own subset of OSNs, which in turn, leads to a distinct combination of target glomeruli receiving these stimulatory inputs (Ressler et al., 1994; Vassar et al, 1993; Barnea et al., 2004; Vassali et al., 2002). Therefore, the internal representation of a complex “smell” will be the distinct combination of activated glomeruli in the bulb, a pattern of activity that the mouse brain can recognize as a specific odor stimulus (Figure 2).

A second chemosensory organ in mice is the accessory olfactory system (AOS), which is located at the anterior base of the nasal septum (Figure 1). It consists of vomeronasal epithelium, which is structurally similar to the OE. Axons of the VSNs project to the accessory olfactory bulb (AOB) that is located at the posterior-dorsal side of the OB (Figure 1). VSNs express two types of vomeronasal receptors, V1R and V2R, that detect non-volatile pheromones
involved in social and sexual behaviors (Tirindelli et al., 2009; Silva et al., 2017; Dalton and Lomvardas, 2015).

The mouse olfactory system also contains two additional olfactory structures: the septal organ (SO) and the Grueneberg ganglion (GG) (Figure 1). The SO is located at the ventral base of the nasal septum and posterior to the opening of the nasopalatine duct that delivers fluids from the mouth to the nose (Weiler et al., 2003). The major function of the SO is to sense the chemicals found in fluids such as urine, sweat, and the endocrine gland, as well as mucous secreted in genitals liberated with licking behaviors (Jolly et al., 1966; Steel et al., 1984). The SO mediates innate behaviors such as fear, sex and social responses (Ma et al., 2003). The GG is located at the entrance of the nasal cavity and detects maternal cues, as well as alarm pheromones from predators (Brechbuhl et al., 2008). It consists of ciliated GG sensory neurons ensheathed by glial cells. The axons of the GG project to the most anterocaudal part of the OB forming a “beads on string” structure called necklace glomeruli (Lou et al., 2008; Fuss et al., 2005).
1.2 Structure and Function of the Mouse Olfactory Epithelium

1.2.A Anatomy of the Mouse Olfactory System

1.2.A.1 Exterior Structures and the Epithelial Layer

The inner lining of the nose contains cartilaginous flaps called nasal turbinates that serve to increase surface area of the nasal epithelium (Chen et al., 2014). Dorsal to the posterior turbinate is the olfactory mucosa (olfactory neuroepithelium) situated on either side of the nasal septum (Cuischieri and Bannister, 1975a). The olfactory epithelium (OE) is the primary sensory organ in the mouse olfactory system, as it is the location where environmental odors are detected. The OE contains pseudostratified layers of columnar neuroepithelium cells that are organized from apical (outer) to basal (inner) in the following order: sustentacular cells (outer apical layer), mature olfactory sensory neurons, immature olfactory sensory neurons, globose basal stem cells, and horizontal basal stem cells (inner basal layer) (Schwob et al., 2002, Leung et al., 2007). All cells of the OE contact a basement membrane called the lamina propria (LP), a thin layer of loose connective tissue that resides beneath the sensory epithelium (Frisch et al., 1967; Farbman et al., 1992). The LP contains fibroblasts, Bowmen’s gland, and bundles of olfactory axons coated with ensheathing Schwann cells, as well as myelinated fibers of the trigeminal nerve, blood vesicles, and connective tissue (Tome at al., 2009; Chen et al., 2014; Choi and Goldstein, 2018). Both LP and OE together constitute the olfactory mucosa membrane, which sits on an
ethmoidal bony surface called the cribiform plate (Borgmann-Winter et al., 2015). The Bowmen’s gland is a branched tubuloalveolar structure that secrets serous and mucous (mucopolysaccarides), which baths the olfactory epithelium, dissolves the odorants, facilitates odor diffusion, and protects against harmful toxins/pathogens (Solbu et al., 2012). Among other components, the mucous layer contains olfactory binding proteins (OBPs), which are small, water soluble proteins that facilitate the delivery of volatile, organic, hydrophobic odorous molecules within this aqueous environment in order to enhance their access to odorant receptor sites (Mariella et al., 2000) (Figure 2).

**Figure 2: Various Cell Types and Markers of the Olfactory Epithelium**
The OE contains Bowman’s G/D cells, HBCs, GBCs, immature and mature OSNs, Sus cells. Cell bodies of Sus cells, OSNs, and duct of Bowman’s gland reside in the OE. OSN axons, acinar structure of the gland, OECs, fibroblasts, blood vessels, macrophages, mesenchymal cells reside in the lamina propria. Cell markers use for immunohistochemical identify are depicted under each color-coded cell types on the right panel (Figure adopted from Lavoie et al., 2017).
1.2.A.2 The Glomerular Layer and Olfactory Processing

The olfactory nerve consists of a collection of OSN axons extending from the OE to the olfactory bulb (OB). Olfactory axons project to a spatially confined area in the OB, each neuron connecting to a single glomerulus (Vassar et al., 1993). There are >2,000 distinct glomeruli in the mouse OB, with one on each side of the bulb dedicated to a specific OR (Royet et al., 1988, Potter et al., 2001). In the periphery, the OE itself has crude spatial organization whereby each OR is expressed within one of four spatially distinct zones (Ressler et al., 1993). Within each zone, OSNs expressing the same OR are seemingly randomly distributed, yet are programmed to extend their axons to a single glomerulus in the OB from these various starting positions (Vassar et al., 1993, Vassalli et al., 2002) (Figure 3). I will discuss the OSN-glomeruli targeting problem in greater detail in section 1.3 below, which focuses generally on the development of the olfactory system.

Decoding “smells” is a consequence of the specific subset of glomeruli receiving input from environmental cues. For example, the “smell” of lemon is due to volatile chemistry inhaled into the nose, whose structural qualities bind to a combination of OR proteins of specific OSNs, which in turn, send action potentials to the corresponding combination of glomeruli to which these OSNs converge (Rubin and Katz, 1999; reviewed in Mori and Sakano, 1999). Therefore, OSNs that respond to lemon odor activate a subset of OSNs to generate a topographic map of activated glomeruli in the OB, and the spatial
organization of this topographic map for lemon odor is conserved among individuals within the species (Bulfone et al., 1998; Imai et al, 2010; Ache et al., 2005).

Figure 3: Axonal Connectivity between the OE and the OB:
The OE is divided into four zones (zone I-IV) depending on the type of ORs expressed within each zone. OSNs expressing zone specific ORs are distributed randomly within each zone. OSNs expressing same receptor converge their axons to the fix glomeruli on the OB (Figure adopted from Mori et al., 1999).

1.2.B Non-Neuronal Cells in the Mouse Olfactory System

In the following subsections, I describe morphological and functional properties of each of the major non-neuronal cell types that comprise the olfactory
epithelium: the sustentacular cells, microvillar cells, and gland and duct cells (Figure 2).

1.2.B.1 Sustentacular cells

Sustentacular (Sus) cells form the most apical layer of the OE (Graziadei and Monti Grazieadei, 1979). Sus cells provide structural and metabolic support to the OSNs (Breipohl et al., 1974; Ding and Coon, 1988). Sus cells are mucous-producing and surround the dendrites of the OSNs to insulate and regulate the extracellular ionic environment (Hegg et al., 2009). Sus cells produce cytokeratins (CKs) and detoxifying enzymes such as cytochrome p450 and glutathione S-transferase, to detoxify the OE environment (Yu et al., 2005; Whitby-Logan et al., 2004; Ling et al., 2004; Gu et al., 1998). Sus cells facilitate OSN replacement by becoming phagocytic during OSN cell death (Suzuki et al., 1996).

1.2.B.2 Microvillar cells

Microvillar cells (MVCs) are rare and are found interspersed among Sus cells (Hansen and Finger, 2008). MVCs are morphologically different than Sus cells in that they are bipolar, oval/spindle flask-like shaped, and contain a microvillar/tufted apical membrane with short basal processes that do not reach the basement membrane (Miller et al, 1995; Moran et al., 1982; Rowley et al., 1989). Immunohistochemical analyses revealed two major types of MVCs in the OE. Type I MVCs express ankyrin (integral membrane scaffolding protein that
stabilizes ion channels in the membrane), cytokeratin 18 (CK18, filamentous protein) and villin (actin stabilizing protein commonly found in epithelial cells) (Asan et al., 2005). Type II MVCs express Na+ and K+ ATPase (Asan et al., 2005), E-cadherin, TrpM5 channel (Hansen et al., 2008), CD73 (Pfister et al., 2012), IP3R3 receptor (Hegg et al., 2010), and the neuregulin type I signaling components ErbB2 and ErbB3 (Gilbert et al., 2015). Therefore, MVCs contain many components also expressed in a typical neuron, however, the function of MVCs in the OE remains unclear (Gilbert et al., 2015).

1.2.B.3 Bowman’s Duct and Gland cells

Bowman’s Duct and Gland (D/G) cells are branched tubuloalveolar serous-secreting glands whose acinus (gland) structure is located in the lamina propria with a secretory duct that extends through the OE (Frisch et al., 1967; Graziadei and Monti Graziadei., 1979; Moulton et al., 1970). G/D cells produce mucus that protects the OE from toxins, as well as dissolves odor-containing gases (Brittebo et al., 1997). Mucus contains protective antimicrobial and antiviral immunoglobulins, such as lactoferrin (Getchell et al., 1991). D/G cells express Pax6 and Sox9 transcription factors along with structural and cell adhesion markers, such as keratin 18 (K18) and E-cadherin. These molecular markers are often upregulated in mesenchymal stem cells and cancer stem cells, and accordingly, D/G cells appear to be a third stem cell type of the OE (Huard et al., 1998; Schwob et al., 2002). For example, upon severe OE injury, D/G cells can
differentiate into Sus cells in order to facilitate tissue recovery (Huard et al., 1998)

1.2.B.4 Other Cell Types of the Olfactory Mucosa

The olfactory mucosa (OM) contains a variety of other cell types that serve a protective and/or supportive function for the OSNs, including olfactory ensheathing cells (OECs), olfactory mucosal derived mesenchymal stem cells (OM-MSCs), blood vessels, fibroblasts, neutrophils and macrophages (Vincent et al., 2005; Li et al., 2001, Mackay-Sim et al., 2011; Borders et al., 2007; LaMantia et al., 2000; Gilbert et al., 2015). OECs are morphologically similar to glial cells, and they wrap around the non-myelinated olfactory nerve bundles and support OSN regeneration during injury (Li et al., 2005). OECs and fibroblasts produce growth factors such as neuropeptide Y (NPY), p75 receptor, and fibroblast growth factors (FGFs) that support the growth and maturation of OSNs (Franceschini et al., 1996; Au et al., 2002; Bianco et al., 2004; Winsted et al., 2014; Holbrook et al., 2016). Macrophage and neutrophil cells also protect the OM against bacterial/viral infection, as well as clearing away cell debris and providing an inflammation-free niche to facilitate neuronal survival (Getchell et al., 1995; Smithson et al., 2010).
1.2.C  The Olfactory Sensory Neurons of the OE

Olfactory sensory neurons (OSNs) are bipolar, glutaminergic neurons whose ciliated dendrites project into the lumen of the nasal cavity for odorant detection, and whose single unbranched axon projects to the olfactory bulb (OB) via the olfactory nerve (Graziadei and Monti Graziadei, 1979; Moulton et al., 1970; Cuschieri and Bannister 1975; Firestein et al., 2001). Each mature OSN possess 10-30 non-motile cilia on the dendritic knob that function to increase the surface area of each OSN approximately 25-40-fold, thus enhancing odorant-binding capability (Lledo et al., 2005; Rawson and Yee, 2006). Odorant receptors (ORs) are expressed on the surface of the cilia where they bind to odorants to initiate a signal transduction into the lumen (Nagai et al., 1993; Godfrey et al., 2004; Young et al, 2002; Zhang and Firestein, 2001, Mombaerts et al., 2004). The OR-mediated signal transduction system will be described in detail in section 1.4 below.

As described previously, OSNs are continually produced during adult life (Graziadei and Monti Graziadei, 1979). The positions of neuronal subpopulations in the OE correspond to different stages of neuronal maturity: the most recently differentiated immature neurons are located more basally (proximal to the basal stem cell populations) and the most mature OSNs are located more apically (proximal to the outer surface of the OE) (Iwema et al., 2003; Brann et al., 2014, Schwob et al., 2017). Several molecular markers are expressed in both mature and immature OSNs, including neuron specific beta-III tubulin (Tuj1) and neural
cell adhesion molecule (NCAM) (Roskams et al., 1998; Calof et al., 1998; Key et al., 1990). Mature OSNs express the olfactory marker protein (OMP) (Margolis et al., 1972, 1999, Bozza et al., 2004), adenylyl cyclase-III (Adcy3), olfactory-specific G-protein (Golf) and βγ subunits of GPCR (Belluscio et al., 1998, Bakalyar et al., 1990; Pfeuffer et al., 1989). Immature OSNs express the growth-associated protein (GAP43), Tuj1, and NCAM. GAP43 is a nervous tissue specific cytoplasmic protein located in axons and growth cones during development and its synthesis is up-regulated during axonal regeneration (Rodriguez-Gill et al., 2015; Iwema et al., 2003; Tan et al., 2015). In general, genes associated with chromatin modification, intracellular transport, neurite growth, RNA processing, cholesterol biosynthesis, and apoptosis are overrepresented in immature OSNs, whereas, ion transporters/channels, presynaptic functions and cilia-specific genes are overrepresented in mature OSNs. (Lyons et al., 2013; Zhang et al., 2016; Krolewski et al., 2012). These expression profiles are an important consideration for Chapter 3 of my thesis, where I discuss the differentiation-dedifferentiation potential of immature OSNs (e.g. OP6 cells) with focus on molecular markers to stage developmental status.

In the following sections, I will describe our current understanding of the differentiation of OSNs from stem-like populations (section 1.3 below), including regenerative properties of the system, and the acquisition of OSN identity via the monogenic/monoallelic expression of one OR gene per neuron (section 1.4 below), including key regulatory attributes involved in OR regulation.
1.3 Development of the Olfactory Epithelium

1.3.A Embryonic Development

In order to understand regenerative capacity of the OE and OSN differentiation through adulthood, it is important to first understand embryonic developmental events of the olfactory system, such as the formation of olfactory placode (OP), generation of OE, development of the OB and glomerular formation. The OP is derived from rostro-lateral ectoderm of the head and gives rise to the OE, whereas, the OB is derived from the neural crest of the telencephalon. Formation of the peripheral OE precedes formation of the central OB (Hinds et al. 1972a).

1.3.A.1 Olfactory Placode (OP) Development

By embryonic day 9 (E9), two patches of thickened epithelial cells emerge bilaterally on the ectodermal plate (Cuschieri and Bannister, 1975b). By E10, these patches of specified ectoderm grow the rostral aspect on the head (Hinds et al., 1972 a,b). Progressive invagination of this thickened ectoderm forms nasal pits (Figure 4). By E11, nasal pits constrict to give rise to narrow slit-like openings that further deepen to generate nasal cavities. By E11.5, secondary recesses are formed, and their rims unite, creating nostrils (Treloar et al. 2010, Feinberg et al., 2013). At E11.5, a secondary invagination on the medial wall of the nasal cavity give rise to the vomeronasal organ (VNO) (Smith et al., 2000).

The molecular basis for OP formation includes coordinated activity of transcription factors and signaling molecules. Transcription factors, such as
Dlx5, Dlx3, Pax6, Sox2, and Oct1 induce OP development (Bhattacharyya et al. 2008, Cornesse et al., 2005; Donner et al. 2006). Mesenchymal cells (e.g. olfactory ensheathing cells) secret morphogens, such as retinoic acid (RA), fibroblast growth factor (fgf8), sonic hedgehog (shh), and bone morphogenic proteins (BMPs), which play important roles in olfactory lineage specification, OSN differentiation, tissue patterning, and OE morphogenesis, respectively (Lamantia et al. 2000, Kawauchi et al. 2004, 2005, Beites et al. 2005, Rawson and Lamantia 2006, Bhasin et al. 2003; Kawauchi et al., 2005)

**Figure 4: Embryonic Development of the Olfactory Placode**
The olfactory placode emerges from thickened neuro-ectoderm patches on each side of the neural plate (E9). Neural groove deepens, neural crest merge (E10), olfactory placodes grow rostrally creating nasal pit (E11), which eventually forms nasal cavity (E11.5). Neural crest contains mesenchymal cells that secrets morphogens (BMPs, SHH, RA and growth factors) required for OE development. (adopted from Feinberg et al., 2013)
1.3.A.2 Generation of the Olfactory Epithelia

The adult OE contains multiple cell types, including mature and immature OSNs, that reside in the middle layer of the OE, the Sus, Microvillar and duct cells that reside in the apical layer, and the GBC, HBC and gland cells that reside in the basal layer (Graziadei and Monti Graziadei, 1979; Cuschieri and Bannister, 1975b; Hinds et al. 1972a; Frisch et al., 1967; Farbman et al., 1992; reviewed in Schwob et al., 2002) (Figure 2). During early embryogenesis (E10-E12), the mitotically dividing olfactory stem cells are located more apically, whereas during late embryogenesis and in the adult, they are located basally in the OE (Farbman and Menco, 1986; Smart et al. 1971). Electron microscopy studies suggest that by E10, the OE contains pale and dark cell types, the former corresponding to differentiating olfactory receptor cells and latter likely comprising the stem cells (Cuschieri and Bannister, 1975b; Farbman et al., 1992). During development, the increase in olfactory receptor cells correlate well with a decrease in the number of apical stem cells (Graziadei and Monti Graziadei, 1979; Schwob et al., 1995).

It is hypothesized that early in the differentiation of OSNs, a handful of OR genes are selected, and as the OSN matures, the developing OSN is tuned to express a single OR allele (Tian et al., 2008; Shykind et al., 2004; Abdus). This hypothesis is supported by recent progress in several single cell transcriptome studies showing that multiple ORs are expressed in immature OSNs (Hanchate et al., 2015, Tan et al., 2015; Saraiva et al., 2015).
OSN axons are visible in the OP as early as E9.5; they migrate out of the OP by E10 (Thieler et al., 1972; Kaufman et al., 1992). The axons of OSNs penetrate the underlying mesenchymal layer by E11 (Cuschieri and Bannister, 1975b; Valverde et al., 1992), and around this time, microtubules and clusters of centrioles give rise to receptor dendrites on the apical side (Cuschieri and Bannister, 1975b). By E12-E16, OSNs becomes multiciliated, and fully mature receptor neurons capable of specific sensory activity are observed as early as E17 (Treloar et al., 1999). Non-neuronal cells, such as Sus, Microvillar cells, and D/G cells, emerge from remaining GBCs starting at E17, coincident with the migration of mitotic cells from the apical to the basal layer of the OE (Cuschieri and Bannister, 1975b; Holbrook et al., 1995). The last cell type generated in the OE is the HBC, generally not identified prior to E17.5 (Figure 5) (Packard et al. 2011b; Reviewed in Schwob, 2017).

**Figure 5: Timeline of the Olfactory System Development**
d.p.c= days of gestation, OE= olfactory epithelium, ON= olfactory nerve, pONL= presumptive olfactory nerve layer, DZ= dendritic zone (Figure adopted from Shay et al., 2008)
1.3.A.3 Development of the Olfactory Bulb

The development of olfactory bulb (OB) begins around E11.5, once the OSN axons reach the telencephalic vesicle (Treloar et al., 1999, 2002). Around E11.5, OSN axons first contact the OB and form the presumptive olfactory nerve layer (pONL); by E12, evagination of the telencephalic vesicle is stimulated by the formation of the presumptive olfactory bulb (pOB) (Treloar et al. 1996, Shay et al. 2008, Gong and Shipley et al. 1995). By E13, the olfactory nerve increases in size as more olfactory neurons are generated (Valverde et al. 1992), and by E15, the first axons grow into the dendritic zone (Treloar et al., 1999). A delay in axonal targeting between E13 and E15 is thought to be required for axonal sorting (St John et al., 1996; Treloar et al., 1997), and by E17, glomerulogenesis is induced by formation of protoglomeruli (Treloar at al. 1999, 2002). By E19 (P0), protoglomeruli have matured into microscopically distinct glomeruli structures (Figure 6) (Treloar et al. 1999).
**Figure 6: The Olfactory Bulb Development**

Sagittal sections of mouse embryo at E12 (A), E13 (B), and E15 (C). Scale bars=200 μm. GAP43 (green), NCAM (red), DAPI (blue). OE=olfactory epithelium, ON=olfactory nerve, pOB=presumptive olfactory bulb. (Figure adopted from Miller et al., 2010)

patterning (Bishop et al. 2000, Zhou et al. 2001). Underlying patterning of the telencephalon ensures proper maturation of mitral and tufted cells, as well as facilitates proper targeting of OSN axons during embryonic to postnatal development of the OB (Rubenstein et al., 1999; Hebert et al., 2002; Gong and Shipley, 1995; reviewed in Lopez-Mascaraque, 2002)

As previously mentioned, ORs play a role in guiding and targeting OSNs to their targets in the OB (Feinstein and Mombaerts 2004). Interestingly, substituting ORs with a generic 7TM GPCR, such as the beta-adrenergic receptor, is sufficient to target an OSN axon to a particular glomerulus, suggesting GPCR-dependent, but olfactory independent mechanism for sorting neuronal targets (Chesler et al. 2007, Feinstein and Mombaerts 2004). The axon guidance and glomeruli targeting of OSNs is described more extensively in 1.3.D below, including this fascinating topic of how GPCRs generally influence the wiring of the system.

1.3.B Basal Cells

Two types of basal cells are found in the OE that are capable of generating all the cell types in the OE: Horizontal Basal Cells (HBCs) and Globose Basal Cells (GBCs). They exhibit different functions (e.g., different responses to injury), morphological and immunological features, and separate developmental timepoints, as detailed below.
1.3.B.1  HBCs and Developmental Potential

Morphologically, HBCs are flattened, with round morphology (4-7 μm in diameter) and electron-lucent nuclei that are attached to the basal lamina as a monolayer (Cuschieri and Bannister, 1975a; Graziadei and Monti Graziadei, 1979). The close proximity between OSNs and HBCs enables direct signaling from the OSNs on the status of the population (Holbrook et al., 1995). Lineage tracing experiments suggest that HBCs are a “reserve” stem cell population that remain relatively quiescent, and proliferate at a very low rate during normal tissue homeostasis, or even with acute loss of mature OSNs that arise when the OB is removed entirely (Leung et al., 2007). However, HBCs respond to severe injuries that deplete the main stem cell population, the GBCs (Leung et al., 2007). HBCs are attached to the basal processes of Sus cells via Notch-jagged (receptor-ligand) interaction at the surface, and during a severe injury that causes depletion of Sus cells, this interaction is broken (Herrick and Schwob, 2015). As a consequence, the intracellular domains of Notch receptors translocate to the nuclei and repress the p63 transcription factor, which reverses quiescence and activates HBC proliferation (Herrick and Schwob, 2015) (Figure 7).
Figure 7: HBC Developmental Potential
HBCs are connected with foot processes of Sus cells via Notch signaling. Death of Sus cells diminishes this interaction and activates HBCs via down-regulation of NICD, Hes1, and p63 transcription factors. (Figure adopted from Schwob et al., 2017)

Active proliferation of HBCs depends on cytokines, such as TGFβ, EGF, FGF, and NGF (Calof and Chikaraishi et al., 1989; Goldstein et al., 1997; Shou et al., 1999). HBCs express stem marker genes, such as Sox2, P63, Pax6, EGFR, and FGFR (reviewed in Schwob et al., 2002). These cells also express ICAM-1, α and β integrins, isolectin binding sugars (α-galactose/α-N-acetyl galactosamine) proteins that provide structural integrity to the OE (Mahantappa and Schwarting, 1993; Calof et al., 2004). In general, HBC molecular profiles share common features with the basal cells of other tissues, such as the respiratory,
lung and intestinal organs (Rock et al., 2010; Randell et al., 1991; Hogan et al., 2014).

What is the biological significance of having both a quiescent and active stem cell pool in the OE niche? Due to constant environmental insult, the OE niche has evolved an efficient regenerative strategy. Active stem cells (GBCs) carry the brunt of the work, whereas quiescent stem cells (HBCs) ensure that the GBC population is robustly maintained. This complementary two state model of self-renewal and tissue repair has gained momentum in other rapidly regenerative niches (Greco et al., 2009; Tumbar et al., 2004; reviewed in Burclaff and Mills, 2018; reviewed in Li and Clevers, 2010). The molecular mechanism by which quiescent and active stem cells interconvert while residing in the same regenerative niche has been an active area of research for regenerative medicine and in cancer stem cell fields (Battle and Clevers, 2017; Dutta and Clevers, 2017).

1.3.B.2  **Globose Basal Cells and Developmental Potential**

Globose Basal Cells (GBCs) are small, round, multipotent, neuro-progenitors, containing electron-lucent nuclei with relatively less heterochromatin, that sit between the more basal HBCs and the more apical immature OSNs they generate (Graziadei and Monti Graziadei, 1979; Cuschieri and Bannister, 1975b). BrdU incorporation studies demonstrated that GBCs are the predominant proliferating cell type in the OE (~ 97.6% of the actively dividing cells) and proliferate at a
high rate in uninjured tissue to maintain tissue homeostasis (Moulton et al., 1974; Hinds et al., 1984; Schwob et al., 2017). GBCs are highly heterogenous with respect to transcription factor expression and progenitor capacity. There are three major types of transit amplifying GBCs (TA-GBCs): Sox2- and Pax6-positive TA-GBCs are totipotent and exhibit long-term self-renewal capacity (Schwob et al., 2012); Ascl1(Mash1) positive TA-GBCs have limited capacity of self-renewal (Gordon et al., 1995); and Neurog1/NeuroD1 positive TA-GBCs have restricted capacity to regenerate only the neuronal cell lineage (Figure 8) (Packard et al., 2011a).

**Figure 8: GBC Developmental Potential**

multipotent progenitor (GBCmpp, Mash1+) self-renew and differentiate into transit amplifying neuronal progenitor (GBCta-n, Ngn1+), to immature neuronal progenitor (GBCinp, Ngn1+), to immature olfactory neuron (OSNi, GAP-43+), to mature olfactory neuron (OSNm, OMP+) (Schwob et al., 2002)
In general, all GBCs label positive for Ki67 and PCNA mitotic markers of cell proliferation (Jang et al., 2014). However, a rare population of GBCs is P27\textsuperscript{kip1} positive (a cell cycle arrest marker) that will transform into Ki67+ neurons upon injury and later return to a quiescence state after recovery (Ladha et al., 1998). This observation suggests the existence of both quiescent and active GBC populations in the MOE (Jang et al., 2014; Mackay-Sim and Kittel, 1991; Duggan and Ngai, 2007; Fletcher et al., 2011). Recent studies from our laboratory demonstrated that lysine specific demethylase (LSD1), an epigenetic regulator of OR genes (discussed in detail below), is compartmentalized in P27\textsuperscript{kip1} positive early post-mitotic cells of the OSN lineage \textit{in vivo}, and the coincidence of OR activations in these subpopulations and the onset of OSN differentiation events in these cell subpopulations might be tied to this chromatin regulating protein (Kilinc et al., 2016). This critical stage of early differentiation – when transitioning from stem-like states to early post-mitotic states – is the core of my thesis work, which focuses on chromatin modulations (mediated by factors such as LSD1) associated with OR regulation and OSN differentiation within a cell line derived from an early OSN progenitor (the OP6 cell line, described in more detail in Section 1.5).

1.3.C OSN Maturation

Earlier studies in adult rodent models suggest that cells near the basal layer of the OE can incorporate thymidine analogues (e.g. BrdU/EdU) into their DNA
during active cell division (Jang et al., 2014). Lineage tracing experiments where GBCs are labeled with replication-incompetent retroviruses capable of generating neuronal progeny that also express the retrovirus, have permitted numerous insights about OSN maturation. First, lineage tracing following acute injury demonstrates that GBCs generate clones of immature neurons suggesting repopulation capability in the OE niche (Huard et al., 1998). Second, immunohistochemical analysis suggest intermediate cells generated post-injury co-express markers for GBCs and neurons suggesting transit amplifying cells originated from GBCs to give rise to the OSNs (Goldstein et al., 1996). Third, transplantation of genetically labeled GBCs from donor to the recipient injured OE give rise to the new neurons in the recipient OE suggesting that the OSN maturation is triggered via intrinsic and extrinsic factors (Goldstein et al., 1998; Chen et al., 2001; Caggiano et al., 1994; reviewed in Schwob et al., 2002).

During neurogenesis, a population of self-renewing GBC stem cells (GBC_{mpp}) differentiate into Mash1+ (a bHLH transcription factor) transit amplifying GBCs (GBC_{TA}) (Guillemot et al., 1993; Gordon et al., 1995). GBC_{TA} give rise to intermediate neuronal precursors (GBC_{INP}), which express another bHLH transcription factor, Neurogenin-1 (Ngn1) (Cau et al. 1997, 2002). GBC_{INP} cells generate immature OSNs (OSN_i) that express β3-tubulin (Tuj1) and GAP43 (Varhaagen et al., 1989; Schwob et al., 1991). Terminal differentiation of OSN_i generates mature OSN (OSN_m) that expresses neural cell adhesion molecule (NCAM) and the olfactory marker protein (OMP) (Nicolay et al. 2006). The entire
process to generate OSN$_m$ following mitotic exit takes approximately one week (reviewed in Schwob et al., 2017).

Along with aforementioned transcription factor-based regulation, neurogenesis is also supported by growth factor-based signaling. Fibroblast growth factor-2 (FGF2) and bone morphogenic proteins (BMPs) play an important role in stimulating and retarding neurogenesis, respectively. FGF2 stimulates proliferation and growth of the neurogenic GBCs progenitors, (Calof et al., 1991; Mumm et al., 1996), and suppresses the neuronal differentiation by maintaining larger colony sizes of neuronal progenitors (DeHamer et al., 1994; Goldstein et al., 1997). BMPs cause the rapid degradation of an important neurogenic transcription factor, Mash1 (Shou et al., 1999), as well as suppresses neurogenic transition from GBC$_{TA}$ (Mash1+) to GBC$_{IMP}$ (Ngn1+) (Shou et al., 2000). Two other factors, TGF-β1 and TGF-β2, have been shown to promote neuronal differentiation in primary cell culture derived from the OE (Mahanthappa et al., 1993; Newman et al., 2000). Other growth factors, such as BDNF, NT-3 and IFG-1, stimulate differentiation and survival of mature OSNs (Roskams et al., 1996; Pixley et al., 1998). Thus, OSN maturation is accompanied by both intrinsic factors (transcription factor network) and extrinsic factors (extracellular signaling pathways).

Lineage tracing has also provided insights on the OSN maturation timeline. Immature neuronal precursors are lineage traced using BrdU incorporation and chasing experiment that are found within 24 hours; after 3
days, post-mitotic GAP43+ neurons are evident that express olfactory receptors (ORs) (Graziadei and Monti Graziadei, 1979; Morrison and Costanzo, 1989; Schwob et al., 1995, 1999). Functional and morphological maturation of OSNs occurs in 5-7 days, at which time the GAP43+ immature marker is down-regulated and the OMP+ mature marker is up-regulated (Varhaagen et al., 1989; Schwob et al., 1991). Therefore, the transition from GAP43 to OMP expressing neurons is a stage that is correlated with OR choice (i.e. terminal differentiation) as well as represents an intermediate stage prior to full OSN maturation.

Immunohistochemical analysis suggests that upon environmental insult and with regular turnover kinetics, the immature neurons express apoptotic factors that ensure their normal lifespan is ~30-90 days (Deckner et al., 1997; Calof et al., 1996; Mackay-Sim and Kittle, 1991; Hinds et al., 1984). However, mature OSNs wired to the OB persist for more than one month, suggesting that glomerular contacts enable prolonged survival, presumably due to trophic dependence on OB-derived factors (Carr and Farbman, 1992, 1993; Schwob et al., 1992).

1.3.D OSN-Glomerulus Targeting

1.3.D.1 Axon Convergence as an Organizing Principle

OSNs form axon bundles as they exit the OE, which are insulated by olfactory ensheathing cells in the lamina propria (Vincent et al., 2005, Li et al., 1997, 1998). These axon bundles cross the cribriform plate and migrate along
glial/Schwan cells to form the olfactory nerve that innervates the OB (Au and Roskams et al., 2003; Mackay-Sim and St John, 2011). In other sensory systems, such as vision and hearing, the relative positioning of peripheral sensory neurons generates a spatially invariant topographic map that decodes the sensory stimuli (reviewed in McLaughlin et al., 2003; Fritzsch et al., 2003). Unlike other sensory systems, the spatial organization of OSNs is random in the OE, an organization that ensures robust sensitivity to odorant qualities that are likewise spatially unpredictable. Genetic lineage tracing experiment show that OSNs expressing the same ORs are randomly distributed into one of four overlapping zones in the OE, yet converge their axons to a stereotypic pair of glomeruli, each located on medial and lateral sides of the OB (Vassar et al., 1993; Ressler et al., 1993; Iwema et al., 2004). Thus, the zonal organization of OR expression in the OE is maintained in the OB, meaning, dorsally located OSNs form dorsal glomeruli, whereas, ventrally located OSNs form ventral glomeruli (Miyamichi et al., 2005). This organizing principle generates a topographic map on the OB: e.g., a given odor that binds to a subset of ORs would activate the corresponding group of OSNs in the OE, whose stereotypic projection on the OB would activate a corresponding subset of glomeruli. The internal representation of odor quality is therefore encoded by spatial pattern of glomerular activation.
1.3.D.2  Axon Guidance Molecules and the OSN Targeting Process

How do the OSNs create a topographic map at the OB? As mentioned previously, each OSN expresses a single OR protein (Buck and Axel, 1991; Chess et al., 1994; Serizawa et al., 2000), and the type of OR expressed influences axon guidance and targeting, thus providing a plausible explanation for why OSNs expressing a given OR, despite diverse origins in the OE, nevertheless are able to converge to a common glomerulus in the OB. This connection appears to be a direct influence from the GPCR protein itself: not only is the OR expressed at the axon termini (where no olfactory perception takes place; Firestein et al., 2004), but a non-OR GPCR is also able to predictably target an OSN to a specific glomerulus target (Strotmann et al., 2004; Omura et al., 2014). Deletion of OR coding sequence disrupts the glomerular targeting (Mombaert et al., 1996), and substitution of an OR coding sequence with a different OR redirects axons to different (novel) glomerulus (Wang et al., 1998). However, it should be noted that OR swapping experiments do not seem to precisely switch the OSN from its host to donor glomerulus target (reviewed in Sakano et al., 2010), suggesting OR protein dependent axonal guidance is necessary but not sufficient. For example, changes in the timing of outgrowth and/or the specific OSNs that co-fasciculate during outgrowth could cause the same OR to address in a slightly different way. Also, other non-OR intrinsic factors, such as Nrp-1 (Imai et al., 2009), Nrp-2 (Norlin et al., 2001), plexin-A1 (Sweeney et al., 2007), Slit1 (Cho et al., 2007), Sema3F (Takeuchi et al., 2010) and IGFR (Scolnick et al., 2008), have been shown to play
a role in processing dorsal-ventral, anterior-posterior, and medial-lateral cues during OSN axon projection (Miyamichi et al., 2005; reviewed in Imai, Sakano and Voshall, 2010; Imai and Sakano, 2011), and these interactions might also render context-dependent modification of OSN targeting.

Dorsal-ventral axon guidance is dependent on the position of OSNs in the OE, as well as sequential projection of OSN axons out of the OE (Tsuboi et al., 2006; Takeuchi et al., 2010; Nguyen-Ba-Charvet et al., 2008). D-V projection is determined by axon guidance receptors expressed in graded manner within the extending axons (Takeuchi et al., 2010). OSNs in the D-V zone mature earlier than those in the V-L zone of the OE (Sullivan et al., 1995). D-V OSNs express Robo-2 receptors, whereas V-L OSNs express Nrp-2 receptors on the axon termini (Cho et al., 2007; Nguyen-Ba-Charvet et al., 2008). A repulsive ligand Slit1 and Sema3F is produced on V-L and D-V side in the OB, which prevents invasion of D-V and V-L OSNs into each other's territory, respectively. Thus, D-V patterning is coordinated by guidance cues (e.g. Robo2, Nrp2) expressed on OSN axon termini as well as gradient of ligands (e.g. Sema3F and Slit1) in the developing OB (Figure 9A).

Medial-lateral projection is guided by the insulin growth factor receptor-1 (IGFR-1). Innervation of lateral OB is severely reduced in the absence of IGFR1 signaling. For example, P2 OR normally targets lateral glomerulus, however, deletion of Igf1r causes misrouting of P2 axons towards more medial glomeruli
Anterior-posterior projection is guided by OR-derived cAMP signals. cAMP activity determines the expression level of neuropilin-1 (Nrp1) and plexin-A1 (PlxnA1) at the OSN axon termini, which are expressed in a gradient at the OB (Imai et al., 2006; Lattemann et al., 2007; Sweeney et al, 2007; Komiyama et al., 2007). Mutant ORs in which the highly conserved “DRY” amino-acid motif, essential for GPCR signaling, is changed to “RDY”, completely abolished the odor-evoked calcium response in these OSNs (indicating the loss of GPCR signaling) and importantly, also disrupted A-P axonal projection patterns (Imai et al., 2006; 2009). Moreover, deletion of the major signaling component, adenyl cyclase-3 (AC3), disrupts the lateral axonal projection of M71 OR at the posterior OB (Zou et al., 2007) (Figure 9B). Thus, the A-P patterning in the OB via GPCR transduction mechanisms is at least one way in which the OR itself is contributing to glomeruli targeting.

To summarize, the olfactory topographic map is dependent on spatial cues established in the D-V, M-L, and A-P axes, as well as activity-dependent mechanisms, the former likely generating coarse innervation patterns and the latter likely refining the final OSN-glomeruli map.
1.3.D.3 Glomeruli Development

Each glomerulus receives input from axons of many OSN expressing the same ORs that synapse with thousands of dendrites of the second-order projection neurons, such as Mitral cells and Tufted cells (M/T cells) (Reviewed in Schoppa and Urban, 2003). Continuous neurogenesis happens in the OB. Neuroblasts are produced from the astrocytes in the lateral ventricular (LV) zone, where they migrate through rostral migratory stream (RMS) (Kaplan et al, 1985; Arturo

Figure 9: OB Axonal Targeting

(A) Axonal projection along Dorsal-ventral (D-V) axis is mediated by complementary gradient of repulsive ligands Slit 1 and Sema3F and their respective receptors Robo2 and Neuropilin1 (Nrp1). (B) Medial-lateral (M-L) axonal projection is mediated by IGF-1 receptor signaling. (C) Anterior-posterior (A-P) axonal projection is mediated by OR protein induced cAMP, which in turn regulates levels of guidance receptor, Neuropilin1 (Nrp1). (Figure adopted from DeMaria et al., 2010)
Alvarez-Buylla et al., 2002) and differentiate into inhibitory interneurons of the OB, such as granular cells (GCs) and periglomerular cells (PGCs) (reviewed in Lledo et al., 2006). PGCs and GCs are GABAergic neurons that synapse with excitatory M/T cells to modulate action potential of M/T cells via a lateral inhibition mechanism (Reviwed in Kosaka and Kosaka, 2005).

By E14.5, the first OSN axons are detected in the olfactory nerve layer, and by E15.5, axons begin to terminate into the presumptive glomerular layer (Potter et al., 2001). Over the next several days (E17-P7), axons may extend to inappropriate glomeruli deeper in the OB, but these connections are later refined to discrete glomeruli by adulthood (Royal and Key, 1999). Glomerular segregation (segregation of discrete neuronal networks in the OB) occurs by homophilic adhesive molecules (Kirrel 2 and 3) and contact-induced repulsive interactions (Eph and ephrin-A) that sort neighboring axons (St. John et al., 2000; Cutforth et al., 2003; Reviewed in Mori and Sakano, 2011). For example, OSNs expressing different ORs have different levels and different types of the Kirrel2 and Kirrel3 receptors, as well as Eph (ligand) and ephrin-A (receptors) (Figure 10). OSN axons expressing similar levels and same types of Kirrel receptors tend to coalesce together via homophilic adhesion (Zou et al., 2007). Therefore, OR-mediated refinement in the OSN-glomeruli map involves a network that links OR protein qualities (e.g., nascent activities) to discrete adhesion properties that influence OSN sorting.
Figure 10: OR Activity Dependent Axon Sorting for Glomerular Development
OR-A produces higher cAMP than OR-B. Levels of expression of homophilic receptors (Kirrel2 and 3) as well as contact-dependent repulsive molecules (Eph-A and ephrinA) varies due to cAMP levels (left panel). OR-A and OR-B axon targeting is sorted out via homophilic adhesion or contact-mediated repulsion as shown in (right panel) (Adopted from Imai and Sakano, 2010).

1.3.E Regeneration of the OE

1.3.E.1 Biological Problem (Environmental Insult)

As described previously, OSN lifespan is ~30-90 days (Graziadei and Monti Graziadei, 1979). OSN lifespan seems to be influenced by a number of intrinsic and environmental factors. For example, the proliferation rate of basal cells was
correlated with the longevity of the OSNs, suggesting a more programmed cycle (Graziadei and Metcalf, 1971). Yet, OSNs grown in a sterile (non-insulting) condition significantly increased OSN lifespan, suggesting that OSN lifespan is modulated by less predictable environmental factors (Hinds et al. 1972a). The health status and age of an animal are also factors in OSN lifespan (Brieipohl et al., 1986; Doty et al., 2012); e.g., OSNs die faster during hypothyroidism (Mackay-Sim et al., 1987). Another factor is OSN density (Mackay-Sim et al., 1988).

Neuronal activity and nurturing factors also dictate lifespan; e.g., immature neurons die at higher rate than mature OSNs because axons of the mature OSNs can get trophic support from the OB (Carr and Farbman, 1992; Schwob et al., 1992). OSNs reside in the OE where they come in direct contact with the air that contains airborne toxins, infectious agents, and viruses, as well as exposing these OSNs to physical and chemical insults. For example, exposure to acute injury (chemical or physical) can lead to widespread death of OSNs and loss of smell, a condition that would absolutely be fatal for any animal depending on its sense of smell for finding food and avoiding predators.

The latter condition – acute loss of the system due to injury – has been modeled and studied extensively by our collaborators in the Schwob lab (Schwob et al., 1991, 1993, 1995, 2002; Jang et al., 2003; reviewed in Schwob et al., 2017). All cell types within the OE are reconstituted after the injury. Most significantly, the basal cells within the OE proliferate rapidly and differentiate into new functional OSNs (Schwob, 1995). The olfactory nerve regenerates and
the OB is reinnervated (Costanzo et al., 1984). But even in the absence of acute assault, environment-induced OSN death requires a system of neurogenesis that persists throughout animal life, a capability supported by the stem cell populations described previously (Graziadei and Graziadei, 1979). In the following section, I will discuss the regeneration models used to study olfactory adult neurogenesis in the OE.

1.3.E.2 Regeneration Models

Regeneration of the OE allows the study of how the olfactory system responds to selective loss of neurons, as well as whole tissue repair in an injury paradigm. There are two injury models commonly used in the field: 1) olfactory bulbectomy (OBX), 2) methyl bromide (MeBr) lesion (Schwob et al., 1992).

OBX induces indirect but selective damage to OSNs by removal of the OB (Schwartz Levy et al., 1991). This surgical procedure destroys axons of mature OSNs (i.e., only those cells directly connected to the OB) caused by retrograde degeneration or apoptosis (Carr and Farbman, 1992). All other cell types in the OE are spared, including the immature OSNs. As a regenerative response, only the stem-like GBCs (not quiescence HBCs) proliferate and differentiate to replace these lost neurons. In the absence of the OB, the remaining immature OSNs, as well as newly differentiated OSNs, extend axons that attempt to form glomeruli but these cells die due to lack of trophic support from the OB (Schwob et al., 1992). Therefore, in the OBX model, GBCs proliferate for extended periods, as
immature neurons increase in number, yet fail to connect and continue to die (Christensen et al., 2001). Another regenerative model commonly applied is one in which the olfactory nerve is removed while leaving the OB intact. As with OBX, this injury activates GBCs to regenerate new OSNs, resulting in the recovery of the olfactory nerve and re-innervation of the OB with surprising precision (Costanzo and Farbman, 1984; Costanzo and Graziadei, 1983).

Methyl bromide (MeBr) is an olfactotoxic gas, which upon prolonged inhalation generates free radicals (Hurtt et al., 1987, 1988; Hallier et al., 1994). MeBr destroys all cell types in the OE except the basal stem cells (Schwob et al., 1994, 1995). MeBr lesion is most extensively used as an injury model because it represents a more natural form of environmental damage. MeBr treatment activates both GBCs and HBCs (Jang et al., 2014; Schnittke et al., 2015; Packard et al., 2016). New Sus cells appear at 2 days post-MeBr, new immature OSNs appear at approximately 4 days, and mature neurons are evident at 10 days (reviewed in Schwob et al., 2017). As noted previously it takes ~7 days for OSN maturation in the embryonic and adult (non MeBr) context. Notably, the complete function and structure of the OE is apparently fully restored in merely two weeks post treatment (Reviewed in Schwob et al., 2002).

1.3.E.3 Sequence of Events During Tissue Regeneration

After MeBr or OBX injury, the GBCs undergo massive proliferation and production of new OSNs. These events can be tracked by expression of
developmental stage markers discussed previously: from Sox2*/Pax6* multipotent progenitors (MPP), to Ascl1* transit amplifying- neuronally committed cells (GBCs-TA-N), to Neurog1* and NeuroD1* immediate neuronal progenitors (INP), to GAP43* immature OSNs (iOSNs), to OMP* mature OSNs, (Manglapus et al., 2004; Packard et al, 2011a, b; Reviewed in Schwob et al., 2002, 2016). Sus cell differentiation proceeds from Sox2*/Pax6* multipotent progenitors (MPP), to Hes1* transit amplifying committed cells (GBCs-TA-N), to Hes1* differentiated Sus cells, (Cau et al., 2000). The other cell types restored by GBCs (e.g., D/G cells) can be similarly tracked from Sox2*/Pax6* multipotent progenitors (MPP) through the differentiation process (Packard et al., 2016).

The quiescent HBC stem cells are attached to Sus cells via a Notch-Jagged receptor-ligand interaction, and as discussed previously, Sus cell death triggers a Notch-mediated pathway (Hes1 and ΔN-p63 down-regulation) to activate these HBCs (Herrick and Schwob et al., 2015). Activated HBCs divide and differentiate to replenish the pool of GBCs that have been spent during the regenerative process (Figure 11) (Herrick and Schwob et al., 2015; Schnittke et al., 2015, Reviewed in Schwob, 2016).
Figure 11: Process and Timeline of the Epitheliogenesis Post-MeBr Injury
Shaded color indicates the timeline for basal cells to give rise to various cell types within the olfactory lineage after the injury. The gradient color within activated GBCs is meant for heterogenous transcriptional states. Straight arrows represent lineage progression, and circular arrow represents self-renewal property. Notch signalling plays an important role in activating HBCs and determining lineage for active GBCs. Transcription factor activity that determine the lineage progression is depicted along the forward arrows. Recent evidence suggests that neuronal differentiation is reversible and immature neuronal progenitors dedifferentiate to multipotency progression is reversible. via de-differentiation (Figure adopted from Schwob et al., 2017)
1.3.F Epigenetic Regulation of OSN Specification and Differentiation

1.3.F.1 Epigenetic Transitions During Neuronal Development

A pool of highly specified neuroepithelial GBCs are responsible for generating the cellular diversity during olfactory placode development, as well as into adulthood (Klein and Graziadei et al., 1983, Graziadei and Monti Graziadei, 1985). To accomplish the ordered production of various cell types in the OE, GBCs transition between lineages via the expression of key epigenetic and transcriptional regulators (reviewed in Schwob, 2002). The epigenetic mechanisms that regulate transcriptional networks to give rise to various cell types in the olfactory lineage remain unclear.

How are neuronal stem cells (NSCs) in the other tissues (e.g. cortex, hippocampus, subventricular zone) specified to become neuronal (or non-neuronal cells) during early embryogenesis? The lineage specification in other neuronal tissues is a coordinated process involving extracellular signaling, transcription factors and epigenetics (reviewed in Goncalves and Gage, 2016; Yoon and Song, 2018; Rubenstein et al., 2005, 2017; Lehtinen and Walsh, 2011). “Epigenetic” modifications are defined as nonpermanent yet heritable changes that regulate gene expression and accomplish phenotypic changes in an organism without altering DNA sequence (Waddington, 1942; Russo et al., 1996; Ledford, 2008; Berger et al., 2009). Well-documented examples of epigenetic
forms of regulation include DNA methylation, histone modifications, chromatin remodeling, and changes in 3D nuclear architecture.

DNA methylation at a promoter causes gene repression (Guo et al., 2011a, reviewed in Li and Zhang, 2014). Genome-wide DNA methylome profiling suggests that during the neurogenic stage of NSCs, the promoters of glia-specific genes (e.g. GFAP, S100β are hypermethylated by de novo DNA methyltransferases (e.g., DNMT3b); through this mechanism of lineage-gene silencing, NSCs are directed to differentiate into neurons and not glia (Takizawa et al., 2001; Nimihira et al., 2004; Smallwood et al., 2014; Sanosaka et al., 2017). Later, during the gliogenic stage, glial gene-specific promoters are demethylated, thereby promoting glial differentiation (Nakashima et al., 1999; He et al., 2005).

DNA demethylation can be both active or passive. For example, mature neurons signal to NPCs via Notch signaling, which activates nuclear receptor IA that in turn binds to glial gene specific promoters to prevent DNMT1 activity, a protein responsible for passive maintenance of DNA methylation through the cell cycle. This loss of passive DNA methylation ultimately leads to neurogenic-to-gliogenic transition of NPCs (Namihira et al., 2009). Alternatively, active DNA demethylation is exemplified by TET family enzymes (Guo et al., 2011a, b; Wu and Zhang, 2014), which convert 5-methyl cytosine (5mC) to 5-hydroximethyl cytosine (5hmC). In the case of the Dlx2 neurogenic gene, deposition of 5hmC in the coding region prevents binding of the polycomb repressor complex (PRC2) facilitating neurogenic gene activation in NSCs (Wu et al., 2010; Hirabayashi et
al., 2009). Accordingly, loss of TET1 in mice (Zhang et al., 2013) and TET3 in *Xenopus laevis* (Xu et al., 2012) decreases neurogenesis.

In the OE, DNA methylation enzymes are differentially expressed in a cell stage-specific manner (MacDonald et al., 2005; Colquitt et al., 2014). DNMT3b is expressed in mitotically dividing GBCs, where it is proposed to methylate developmental genes involved in glial or neuronal lineage commitment, in order to prevent differentiation (MacDonald et al., 2005). Later in post-mitotic immature neuronal progenitors, DNMT3a is co-expressed with HDAC2 (co-repressor) and silences immature neuronal markers (e.g. GAP43, cadherins) and other non-neuronal genes to promote OSN maturation (MacDonald et al., 2005; Colquitt et al., 2014; Krolewasky et al., 2013). Also, the Neurog1, Dlx4/5, and Emx1/2 neurodevelopmental gene clusters are regulated by DNMT3a/TET activity (Colquitt et al., 2013, 2014). DNMT1 is expressed in cycling GBCs to preserve methylation states around genes that must be silenced to maintain self-renewal properties (MacDonald et al., 2005). DNMT1 activity may also participate in differentiation in GBCs that transition along the OSN lineage by inhibiting targets that promote re-entry into the cell cycle (Sen et al., 2010; Georgia et al., 2013). Thus, the hierarchical activity of DNMTs (via DNA methylation) and TETs (via DNA demethylation) are coupled with the respective developmental stages to specify various cell lineages originating from common progenitors (reviewed in Namihira and Nakashima, 2013).
Histone modifications (methylation, acetylation) are catalyzed by proteins that function as “writers” (histone methyltransferases, acetyltransferases), recognized by “readers” (e.g., HP1 and other chromatin binding proteins) and reversed by “erasers” (histone demethylases, deacetylases) (Kornberg, 1999; Helin et al., 2013, Schuettengruber et al., 2007). Histone methylations at H3K4, H3K36, H3K79 residues are correlated with transcription activation, whereas methylations at H3K27, H3K9, and H3K20 are correlated with transcriptional repression (Vakoc et al., 2006). Histone acetylation neutralizes the positive charge on lysine residues of the histone tail, decondensing chromatin, and facilitating transcription activation (Kouzarides, 2007; reviewed in Jaenisch and Bird, 2003; reviewed in Strahl and Allis 2000). Therefore, the presence of numerous histone-modifying proteins, each of which can form complexes with other transcription factors, enables a broad and dynamic means for altering the epigenetic code in a genome.

In embryonic stem cells (ESCs), neurogenic genes (e.g., Dlx2/3 and Mash1) are “poised” for activation, yet are transcriptionally inactive, due to the presence of bivalent domains consisting of both activating H3K4me3 and repressing H3K27me3 marks (Bernstein et al., 2006; Hochedlinger and Jaenisch, 2015). During the transition from ESC to an NSC, bivalent domains at neurogenic genes are resolved by a component of the trithorax group, MLL1 (Lim et al., 2009). The MLL1 complex recruits the KDM6 histone demethylase, jmd3, to selectively remove H3K27me3 in order to de-repress the neurogenic gene, Dlx2.
Simultaneously, MLL1 prevents the PRC2 repressor from binding at the Dlx2 promoter (Lim et al., 2009; Burgold et al., 2008; reviewed in Lim and Alvarez-Buylla, 2016). Repressor inhibition is a key aspect of this regulatory switch, since deletion of $\text{ezh2}$ and $\text{Ring1}$ (components of PRC2) from gliogenic NSCs is sufficient to trigger neurogenic NSC differentiation (Hirabayashi et al., Pereira et al., 2010). The histone deacetylase, HDAC1, is expressed in NSCs, while the histone deacetylase HDAC2 is expressed in neurons, suggesting delineated roles of HDACs in establishing lineage transcriptional networks (MacDonald and Roskams, 2008); accordingly, the conditional deletion or cell stage specific inhibition of HDAC1 impairs neuronal differentiation (Montgomery et al., 2009; Yuniarti et al., 2013). HDACs form complexes with other regulators to execute targeted gene repression. For example, HDAC association with LSD1 and the REST co-repressor targets genes in the neurogenic program in non-neuronal cells, because these target genes contain the Repressive Element 1 (RE-1) binding site for the REST co-factor (Shi et al., 2004; Ballas et al., 2005). LSD1 is another chromatin modifying protein that mediates H3K4 and H3K9 demethylation (Shi et al., 2004; Metzger et al., 2005). LSD1 is expressed in NSCs, where it binds with the co-repressor Rcor2 to suppress Dlx2 and the sonic hedgehog (Shh) pathway (Wang et al., 2016). LSD1 promotes neurogenesis by removing H3K4me3 from HEYL (a Notch target gene), keeping HEYL inactive, which facilitates neuronal differentiation (Hirano et al., 2016). Accordingly, depletion of LSD1/Rcor2 decreases neurogenesis (Wang et al., 2016). In 1.4, I
will return to a discussion of LSD1-mediated epigenetic regulation of OR genes during OSN development.

Another mechanism for modifying epigenetic states of target genes involves chromatin remodeling complexes. These proteins (e.g. SWI/SNF, Brg/BAF) hydrolyze ATP to relax condensed chromatin and actively move nucleosomes, which functions in some contexts to increase/decrease accessibility to transcription factors (De la Serna et al., 2006; Crabtree, 2011; reviewed in Hargreaves and Crabtree, 2011). During neurogenic transitions, the Brg/BAF complex undergoes combinatorial subunit switching – e.g., switching to NSC-specific BAF subunits within the complex – in order to remodel chromatin around key neurogenic genes (Ho et al., 2009; Takebayashi et al., 2013; Bachmann et al., 2016; Lessard et al., 2007; reviewed in Staahl and Crabtree, 2013). To date, chromatin remodeling complexes have not been implicated in OSN differentiation.

3D organization of the genome within the nuclear space is yet another important means for regulating transcription of NSC target genes (Cremer et al., 2006, 2010; Gilbert et al., 2004). Chromatin is highly organized within cell nuclei, and there are several examples that illustrate the importance of this organization in gene activation/silencing in the context of neuronal differentiation. Generally, chromosomes are partitioned into euchromatic (transcriptionally open) and heterochromatic (transcriptionally inactive) regions and these regions are embodied within long loops, in which topologically
associated domains (TADs) are formed (Dixon et al., 2012). Long-range 3D interactions between enhancers and genes are prevented/enabled within TADs (Nora et al., 2012, Sexton et al., 2012). For example, heterochromatin domains labeled with H3K9 methylation are tethered to the nuclear lamina at nuclear periphery; these domains are called lamina-associated domains (LADs) (Steensel and Henikoff, 2000; Pickersgill et al., 2006; Kind et al., 2013; Thomson et al., 2004), whereas euchromatin domains exist in TADs that facilitate access to transcription factor hubs and RNA polymerase factories (Cremer et al., 2006; Hakim et al., 2011). The 3D architecture within interphase nuclei changes as the cell differentiates (Gilbert et al., 2004; Williams et al., 2002, 2006). For example, in the progenitors of the rod photoreceptor cells in nocturnal animals, the chromocenters are associated with lamina at the nuclear periphery, but as cells differentiate, the chromocenters merge to form a single centrally located chromocenter in the nuclear interior (Solovei et al., 2009). This inverted pattern of nuclear architecture is critical for proper photoreceptor activities in these animals (Solovei et al., 2009). I will revisit this particular example of remodeled nuclear architecture in the context of OR regulation (Section 1.4), as it appears that a very similar chromocenter phenomena occurs during the maturation of OSNs and the establishment of OSN identity via OR gene silencing.
1.3.F.2 Elaboration on Two Key Epigenetic Regulators

My thesis research focused on two specific chromatin regulators, the lysine-specific demethylase-1 (LSD1) and the eukaryotic histone methyl transferase-2 (EHMT2, referred hereafter to its other name, G9a). LSD1 and G9a are factors that remove and add methyl groups, respectively, to specific histone residues, with the former demethylating H3K4 or H3K9, and the latter methylating H3K9 (Shi et al., 2004; Metzger et al., 2005; Tachibana et al., 2001). This section elaborates further on our current state of knowledge about these two chromatin regulators in particular.

**LSD1.** Histone methylation was believed by some groups to be irreversible. The landmark discoveries of histone demethylases, such as lysine specific demethylase (LSD1), suggested that the methylation marks are, indeed, reversible (Shi et al., 2004). LSD1 was initially identified as a component of REST-CoREST repressor complex, where LSD1 selectively removes the methyl group from H3K4me2/1, a histone mark correlated with active gene expression (Hakimi et al., 2002; Shi et al., 2004, 2005). Thus, the first description of LSD1 was as a transcriptional repressor. Recently, LSD1 has also been shown to demethylate H3K9me1/2, a histone mark correlated with transcriptional inactivation, and therefore, in some contexts LSD1 functions as a transcriptional activator (Metzger et al., 2005).
LSD1 as a repressor. There are many examples of LSD1-mediated repression (via H3K4 demethylation) in various developmental contexts. For example, in the pituitary gland, LSD1 is recruited to the CtBP-CoREST complex to repress the expression of growth hormones (Wang et al., 2007). In the hematopoietic system, the LSD1-CoREST complex is recruited by Gfi1/Gfi1b (growth factor independent-1 transcription repressors) at differentiation gene specific promoters to repress differentiation of hematopoietic progenitors (Saleque et al., 2007). During ES cell differentiation, LSD1 partners with nucleosome remodeling and histone deacetylase (NuRD), where it silences pluripotency-specific enhancers thereby allowing ES cells to transition towards differentiation (Heintzman et al., 2009; Whyte et al., 2012). LSD1 associates with CoREST complex containing SIRT1 (histone deacetylase) to repress Notch target genes during development in Drosophila (Mulligan et al., 2011). In fission yeast, LSD1 binds at heterochromatin-euchromatin boundary elements, where it prevents euchromatin spreading in order to maintain the integrity of chromatin domains (Nakayama et al., 2001; Lan et al., 2007). These examples illustrate a variety of regulatory contexts, as well as potential co-factors, that are instrumental in promoting epigenetic re-programming, and memory (Liang et al., 2017). Chapter 2 of this thesis explores a novel repressor role for LSD1, in which perturbation of this protein in immature cells of the OSN lineage result in an excess of OR activations per cell (Vyas et al., 2017).
LSD1 as an activator. The mechanism for LSD1-mediated activation is the removal of repressive H3K9 methylation marks (de-repression), and this functionality is now evident in numerous developmental contexts (reviewed in Cloos et al. 2008; Nottke et al., 2009). LSD1 functions as a co-activator of estrogen receptor α (ERα) target genes involved in cell proliferation programs (Metzger et al., 2005; Garcia-Bassets et al., 2007). During ERα dependent gene activation, LSD1 forms interchromosomal granules thought to facilitate trans interactions among different chromosomal loci (Hu et al., 2008). LSD1 partners with the androgen receptor to de-repress androgen target genes, and this is of clinical importance in the context of prostate cancer (Metzger et al., 2005; Wissmann et al., 2007). Recently, a neuron-specific isoform of LSD1, LSD1-8a, was shown to complex with supervillain to demethylate H3K9 differentiation specific target genes to promote neuronal differentiation (Zibetti et al., 2010; Laurent et al., 2015). In plants, FLD (an LSD1 homolog) cooperates with RNA processing proteins (e.g. RNAi machinery) to regulate polyadenylation, a function that facilitates a key developmental transition in flower development (Liu et al., 2007). As I will discuss further in Section 1.4, LSD1 is also known to function in the regulation of OR gene expression; e.g., to remove H3K9 methylation on the one OR gene chosen to express in each OSN (Lyons et al., 2013).

Non-histone demethylation by LSD1. In development, the activation/repression of transcriptional programs is the principle mechanism by
which cells differentiate, and therefore, the histone-modifying roles of LSD1 are most critical in this context. However, LSD1 regulation can also be post-transcriptional via demethylation of non-histone target proteins that alter cellular behavior of these targets. For example, LSD1 demethylates the p53 (tumor suppressor protein), and p53 inhibition promotes apoptosis (Huang et al., 2007). In vertebrates, LSD1 demethylates DNMT1, which leads to DNMT1 degradation, and this loss has been shown to promote developmental processes, such as imprinting in germ cells (Wang et al., 2009). In fission yeast, LSD1 regulates replication fork pausing and imprinting (Holmes et al., 2012).

In summary, LSD1 plays both a transcriptional activator or repressor role depending on the cellular and developmental context. It exerts these activities by partnering with various co-activator/co-repressor complexes. The dynamic functions of LSD1 are also influenced by the type of isoforms expressed in specific tissues, e.g. LSD1-E8a isoform mediates H3K9 demethylation. LSD1 plays a broad role in disease conditions, development and differentiation. In this thesis work, I investigate a role for LSD1 in OR gene regulation, as well as the OSN development, as discussed in Chapter 2 and 4, respectively.

**G9a.** G9a is a histone methyltransferase (HMT) that catalyzes mono and dimethylation at Histone 3 (H3) lysine 9 (K9) histone marks (Tachibana et al., 2001; Ogawa et al., 2002). H3K9 methylation is associated with heterochromatin, and therefore, G9a can function as a gene repressor. G9a
repressor function is dependent on its SET enzymatic domain that adds methyl
group at H3K9 (Tachibana et al., 2001). There are several examples of G9a-
mediated repression in developmental contexts. For example, G9a, along with its
corepressor complex containing HDAC1, Suv39h1, and Jarid1, is recruited by
the NF-E2/p45 transcription factors at the embryonic globin locus (Ey) to
repress embryonic globin expression in adults (Chaturvedi et al., 2012; reviewed
in Sankaran et al., 2013). G9a represses pluripotency genes (e.g. Oct3/4 and
Nanog) in ES cells to promote differentiation (Tachibana et al., 2002; Yamamizu
et al., 2012). G9a plays a repressive role in genomic imprinting by silencing
imprinted alleles of Osbp15, Cd81, Ascl2, and Kcnq1 (Wen et al., 2009; Zhang et
al., 2016; Wagschal et al., 2008). G9a is involved in pathogenesis of Prader-Willi
syndrome by disrupting allelic exclusion of the Prader-Willi imprinting control
center (Xin et al., 2003). G9a has also been implicated in the silencing of
enhancers during early embryogenesis (Zylicz et al., 2015). G9a methylates
pericentromeric and telomeric histones to maintain structural integrity of
chromosomes; loss of G9a function leads to depletion of H3K9me and
chromosome instability (Yokochi et al., 2009; reviewed in Verdaasdonk et al.,
2011). G9a repression of developmental genes is critical in the terminal
differentiation of blood, retina, cardiac, neuronal, muscle and germline lineages
(Olsen et al., 2016; Rao et al., 2010; Inagawa et al., 2013; Chen et al., 2012;
Fiszbein et al., 2016) (Figure 12).
Even though G9a adds repressive marks on histone proteins, and is therefore built for gene silencing functions, G9a has also been shown to collaborate with other proteins to activate gene targets independent of methyltransferase activity and dependent on protein interactions with ankyrin and other partners (Milner et al., 1993; Zhang et al., 2002; Wu et al., 2010). For example, G9a partners with the non-enzymatic p300 co-activator complex to activate adult globin genes (Lee et al., 2006; Bittencourt et al., 2012). G9a

**Figure 12: Transcriptional Activator and Repressor Roles of G9a**

G9a is recruited by transcription factors. The repressor or activator role of G9a is determined by co-repressor or co-activator complexes shown in the figure. G9a repressor role is SET domain dependent whereas its activator role is SET domain independent (published in Shankar et al., 2013)
functions as a scaffolding protein recruited by the Runx2/Nfe2 transcription factor to facilitate gene activation of Runx2 target genes, such as MPP9, CSF2, SDF1, and CST7 in prostate cancer (Purcell et al., 2012). Like LSD1, G9a also regulates proteins post-transcriptionally. For example, G9a is required to repress M-phase specific cyclins (e.g. Cyclin D), thus regulating cell cycle exit, a pivotal step in neurogenesis (Olsen et al., 2016). G9a inactivates p53 protein via methylation (Chen et al., 2010), as well as activates p53 protein via recruiting a p300/CBP histone acetyltransferase complex (Rada et al., 2017). G9a-mediated methylation impacts the stability of the Sox2 protein (Lee et al., 2015), hypoxia inducing factor 1 (HIF1α) (Oh et al., 2015) and MyoD (Kontaki et al., 2010; Jung et al., 2015), as well as many other targets in a diverse set of developmental and disease contexts (reviewed in Deimling et al., 2017; Shankar et al., 2013).

1.4 Odorant Receptors and OSN Development

1.4.A.1 The Odorant Receptor

The Odorant Receptor (ORs) gene family is the largest gene family in the mammalian genome, first discovered by Linda Buck and Richard Axel in 1991 (Buck and Axel, 1991). In rodents there are ~1400 OR genes clustered at >50 chromosomal loci (Godfrey et al., 2001; Young et al., 2002; Zhang et al., 2004, 2007); the human genome encodes ~ 900 OR genes, including numerous
pseudogenes that have lost function during primate evolution (Glusman et al., 2001; Malnic et al., 2004; Keller et al., 2007; Mainland et al., 2014).

OR genes encode G-protein coupled receptors (GPCRs). These are proteins with 7 transmembrane helical domains that are important for ligand (odorant) interactions, along with contributions from extracellular loop domains (Buck and Axel, 1991; Mombaerts et al., 1999). The intracellular side of the receptor contains several conserved residues that function in signal transduction, including the LHTPMY motif within the first intracellular loop, MAYDRYVAIC motif at the end of TM domain 3 (TM3), a short SY motif at the end of TM5, a FSTCSSH motif at the beginning of TM6, and an PMLNPF motif in the TM7 (Buck and Axel, 1991; Kobilka et al., 1992, Michaloski et al., 2006).

1.4.A.2 Odor Stimulation and Coding

In order to be perceptible by our noses, odorants need to be lipophilic, small (molecular weight < 300 Da) and volatile organic compounds. Odorants dissolve in the mucous and then couple to the odorant receptors that are expressed on the plasma membrane of the OSNs. Odor stimulation occurs when an odorant molecule with specific chemical structure fits into a binding pocket generated by the extracellular loops of the OR protein. A change in a single amino acid can change the odor binding pocket, thus altering the chemical binding specificity. For example, the mouse I7 OR protein, unlike the rat I7 OR protein, produces a greater response to heptanol (seven carbon atoms containing alcohol) than
octanol (eight carbon atoms containing alcohol), simply due to a single valine-to-isoleucine substitution (Krautwurst et al., 1998).

OR proteins undergo conformational changes upon binding of an odorant into an extracellular loop of the receptor protein. This conformational change is sensed by the intracellular domain of the OR protein, which in turn activates a cascade of signal transduction via secondary messengers (section 1.4.A3). This leads to the opening of the ion channels, change in the electrical potential across the plasma membrane, and generation of an action potential, which will convey the odorant presence to the brain.

OR-ligand matching was a difficult task partly due to OSNs expressing a given OR were only represented in ~0.1% of the OSNs in the OE. The Firestein group overcame this challenge by overexpressing adenovirus that carries an I7 OR gene cassette, such that the I7 receptor was now ubiquitously expressed in the OE (Zhao et al., 1998). Electro-olfactograms (EOG) recorded from OE slices suggested that I7 OR responded to more than one odorant and it could discriminate between structurally similar odorants (Zhao et al., 1998). Later, the Buck group conducted experiments in which OSNs were loaded with calcium responsive dye, and following exposure to specific odorants, they isolated activated OSNs by calcium imaging and identified the ORs in the responsive neurons (Malnic et al., 1999). Three conclusions were drawn from these experiments: 1) each OR protein can respond to multiple odorants; 2) each odorant molecule binds to multiple different ORs with varying affinities; 3) a
mixture of different odorants is detected by a distinct combination of receptors. This later became a principle and widely accepted as a “combinatorial code” model for decoding smell perception in the brain (Malnic et al., 1999; Araneda et al., 2000; Kajiya et al., 2001; Saito et al., 2009).

Most of the odorant receptors remain “orphaned” (i.e. their ligands are unknown). Deorphanizing a receptor is carried out by expressing a receptor in heterologous cells (i.e. a system in which OR proteins are typically not expressed) (Krautwurst et al., 1998). It has been difficult to achieve functional expression of cloned ORs in heterologous cells because most OR proteins are retained in the endoplasmic reticulum and do not transport to the plasma membrane (Krautwurst et al., 1998). However, this challenge was later overcome by fusing the first 20 amino acid residues of rhodopsin to the N-terminal of an OR cassette, or co-expressing ectopic ORs with the olfactory-specific chaperones, RTP1 and REEP, which allows robust membrane expression of ORs in heterologous systems (Krautwurst et al., 1998; Matsunami et al., 2009).

1.4.A.3 OR-specific Signal Transduction

As discussed above, odorants from the external environment bind to the large extracellular loop (variable region) of the OR protein, which undergoes a conformational change to activate intracellular G-protein (Gαolf, an olfactory specific GPCR) (Pace et al., 1985; Jones and Reed, 1989). The odorant activated Gαolf subunit exchanges GDP to GTP, which causes dissociation from Gβγ (Levy et
al., 1991; Ronnett and Moon, 2002). GTP-bound $G_{olf}$ activates adenylyl cyclase III (ACIII), an olfactory-specific enzyme that hydrolysis ATP to produce increases in the secondary messenger cAMP (Figure 13) (Jones and Reed, 1989; Firestein et al., 2001; Wong et al., 2000). As cAMP levels increase, the cyclic nucleotide-gated channels (CNGC) are activated, which allows Na+ and Ca2+ cations to enter the OSN cilia, causing depolarization (Belluscio et al., 1998; Brunet et al., 1996; Firestein et al., 1996). High intracellular Ca2+ levels activate chloride channels, and the Cl- efflux induces electrical excitation of the OSN membrane and a resulting action potential (Kleene et al., 1991; Kurahashi et al., 1993; Lowe et al., 1993; Firestein et al., 1987). Once the action potential reaches the axon terminal, glutamate neurotransmitter is released (Berkowicz et al., 1994), which activates glutamate receptors on the post-synaptic mitral and tufted cells of the glomeruli, to trigger further action potentials in relaying information up to higher brain centers (Reviewed in Reed et al., 2004).

![Figure 13: The Olfactory Receptor Signal Transduction Pathway](image_url)

**Figure 13: The Olfactory Receptor Signal Transduction Pathway**

ORs binds to odorants from environment, undergo conformational change, activates intracellular signaling components that activates ACIII that hydrolyses ATP to make cAMP, which binds to cyclic nucleotide gated channels (CNGCs) that allows influx of Na and Ca, generates action potential (Kaupp et al., 2010)
1.4.A.4 OR Genome Organization and Evolution

The ORs are categorized into several gene families based on the gene structure and sequence similarity. Most OR genes have an unusual gene structure with 1-kb intronless coding region terminated by the polyadenylation signal, and short 5’ and 3’ untranslated noncoding regions (Glusman et al., 1996; Sosinsky et al., 2000; Young and Trask, 2002). ORs sharing >40% sequence similarity are categorized as being members of a common family and those with >60% sequence similarity are categorized as members of a common subfamily (Lancet et al., 1993; Lane et al., 2001; Malnic et al., 2004). Olfactory sensitivity to particular odorants is a function of the number of OSNs expressing OR proteins that detect an odorant at high affinity, as well as the sheer size of the olfactory epithelium; the discriminatory power and range of odorant detection will be a function of the number of intact ORs encoded in a genome. For example, dogs have larger epithelia than mice and are known for excellent olfactory capability, yet, the number of OR genes encoded in the dog genome is smaller than many rodents (Quignon et al., 2003; Godfrey et al., 2004; Steiger et al., 2008). Cross-species comparisons indicate that terrestrial mammals share numerous OR subfamilies; e.g., there are numerous orthologous OR genes shared between mouse and human (Young et al., 2002; Lane et al., 2005; Godfrey et al., 2004), and therefore, these two species are able to recognize many of the same odorant structural motifs (Malnic et al., 2004; Zhang and Firestein et al., 2002). However, the sheer size of the OR repertoire in mice as compared to humans indicates that
mice have higher discriminatory power (and probably greater sensitivity) than humans (Godfrey et al., 2004).

The human genome encodes approximately 339 intact OR genes (Malnic et al., 2004; Olender et al., 2008). These ORs are distributed in 17 OR gene clusters found on almost all chromosomes except chromosome 20 and Y (Zhang and Firestein, 2002; Glusman et al., 2001). There are about 172 OR gene subfamilies, distributed among 51 distinct genomic loci (Young et al., 2002; Malnic et al., 2004). Many ORs from any single chromosomal location may recognize closely related odorant structures given the fact that subfamilies with similar sequences are clustered together (Malnic et al., 1999; Sullivan et al., 1996). The relative decline in human OR repertoires as compared to other mammals may reflect the increase in visual dominance as a primary sensory system with the advent of upright posture, color vision, and expansion of the prefrontal cortex (Gilad et al., 2005, 2007; Shepherd et al., 2004; Quignon et al., 2003; McGunn et al., 2017). In general, ORs are rapidly evolving by a gene birth-death process, facilitated by duplication and recombination events accelerated by repeated homologous segments that characterize OR clusters (Young et al., 2002; Kambere et al., 2007; Zhang et al., 2003; Olender et al., 2008; Issel-Tarver and Jasper Rine, 1997). Species-specific differences and other adaptive evolutionary forces were likely important for establishing olfactory capabilities for the lifestyles of individual species and niches (Zhang et al., 2003; Gilad et al., 2003; Malnic et al., 2004; Dong et al., 2009; Brand et al., 2017).
Mammalian OR genes are classified into two distinct groups: Class-I and Class-II (Glusman et al., 2000; Zhang and Firestein, 2002). Most ORs (~90%) are Class II, and are expressed throughout the dorsal and the ventral OE (Glusman et al., 2001; Zhang and Firestein et al., 2002; Niimura et al., 2007). Class I ORs (~10%) are referred to as “fish-like”, since they are the only ORs found in fish (Ngai et al., 1993), and are expressed exclusively in the dorsal OE (Young et al., 2002; Godfrey et al., 2004). It has been postulated that Class-II ORs evolved specifically for terrestrial life and a new set of volatile (air-born) chemistry that terrestrial animals must process (Freitag et al., 1998; Niimura et al., 2009; Hayden et al., 2010; Hoover et al., 2013; Bear et al., 2016).

1.4.B  OR Gene Regulation is Critical to Olfactory Function

1.4.B.1  Monogenic OR Expression as an Organizing Principle

There were early clues in the study of OR expression that suggested OR genes are tightly regulated within the OE. First, any given OR probe used in RNA FISH on the OE recognizes only ~0.1% of non-overlapping OSN population in the mouse OE (Ressler et al., 1993). Given that the mouse genome encodes on the order of ~1,000 ORs, the expected average number of OSNs per OR would be on the order of ~0.1% (1/1000) if each OSN only expressed one OR. Second, single cell RT-PCR yields only a single type of OR transcript per cell (Malnic et al., 1999); these early observations have recently been confirmed with deep mRNA sequencing on single mature OSNs, and confirm that each OSN expresses only
one OR above noise levels (Hanchate et al; Tan et al., 2015). Third, Chess et al. discovered that OR genes replicate asynchronously and are subject to allelic inactivation (Chess et al., 1994), suggesting that even two highly related alleles of the same gene are mutually exclusively expressed. Finally, a series of transgene experiments provided strong evidence for mutually exclusive OR expression, since even absolutely identical OR transgenes labeled with different fluorescence markers nevertheless do not co-express in the same OSN (Serizawa et al., 2000; Ishii et al., 2001; Shykind et al., 2004). Interestingly, even when an endogenous OR locus is replaced with a non-OR transgene (e.g. β-2 adrenergic receptor or OMP), the non-OR mimics the monogenic and monoallelic expression pattern, suggesting that monogenic regulation is governed by local chromatin and sequence environment (Firestein et al., 2004; Pyrski et al., 2001). I will return to discuss the genetic and epigenetic mechanisms that regulate OR transcription in Section 1.4C and 1.4.D below.

One interesting question concerning the establishment of monogenic OR expression is, how does the cell prevent more than one OR being transcribed? Or a related question is, how does the cell know it has successfully activated its one OR gene, so it knows not to select another? Important insights came from two similar experiments conducted independently by two different groups. In both cases, an OR coding region was replaced with a reporter gene (e.g., LacZ), which means, when that particular gene is activated, those OSNs will be visible yet assuredly will not produce any OR protein (Serizawa et al., 2003; Lewcock et al.,
Interestingly, OSNs that select the “vacated” OR gene locus initiates a new OR selection, indicating that the cell somehow recognized the absence of a successful selection after the locus had been transcribed. A subsequent experiment clarified that it was not the absence of OR DNA or RNA that drove this result, rather, it was the absence of OR protein: by introducing a single nucleotide change (e.g., an early frame-shift insertion/deletion) such that essentially the entire gene was present, the entire mRNA transcript would be present, yet the mutation would prevent productive translation into an OR protein, the OSN would still re-select another OR gene (Shykind et al., 2004; Serizawa et al., 2003; Lewcock and Reed, 2004). The OR protein dependent negative feedback loop is mediated through the presence of signaling components (e.g. Adcy3) that are activated through unfolded protein response (UPR) pathway described in section 1.4D (Dalton et al., 2013). So, a functional OR protein on the surface of the OSN is required in order for the cell to establish the completion of a successful choice, which triggers a negative feedback signal to prevent further OR activation (reviewed in Serizawa and Sakano, 2005; Imai and Sakano, 2007). Such a feedback loop is useful in two ways: first, if the feedback loop down-regulates proteins involved with OR activations, it can help prevent multiple ORs from becoming active in a given cell; and second, if one of the numerous OR pseudogenes is haphazardly selected, it provides a mechanism for saving that cell’s fate by driving a process of re-selection.
The biological significance of monogenic OR expression is that it underlies neuronal specificity. By expressing only one OR per OSN, it functionally specifies that OSN to the odorant-binding capabilities of that OR. Thus, the problem of odorant discrimination is reduced to the question of which specific OSNs are active at a given moment. Of course, this functionality is only possible if the brain can decipher which neurons are active at a given moment, and the solution to this problem lies in the connections made by specific OSN types in the bulb. Interestingly, the OR protein serves a critical role in the wiring of OSNs, as discussed next.

1.4.B.2 OR protein Functions in Glomeruli Targeting

As just discussed, monogenic OR regulation serves a fundamental role in giving each OSN its unique identity. Here, we discuss how monogenic OR regulation serves a fundamental role in establishing appropriate OSN connections in the brain. As mentioned previously, OR protein is expressed both in the ciliated portions of the OSN where odorant-binding occurs in the nose, as well as at axon termini that synapse in the OB (Mombaerts et al., 1996; Wang et al., 1998). Single amino acid changes in an OR sequence can change the odorant response profile and axonal targeting preference (Firestein et al., 2004; Bozza et al., 2009; Zhang et al, 2013). Presumably, the re-routing of OSNs that contain a slightly different OR protein is a critical feature that permits precise OSN sorting irrespective of the specific OR repertoire a given animal or individual possesses.
How does the OR protein function in glomerulus targeting? OSNs expressing an OR-GFP fusion protein label not only their dendrites but also label axons, suggesting OR proteins play a non-odorant binding role in the extending axon (Firestein et al., 2004). Replacement of an OR coding region with the coding region of another OR affects the position of axonal coalescence, as well as switches the glomeruli target, suggesting that the OR is important in axon guidance (Mombaerts et al., 1996). OR coding gene swap or deletion experiments perturb the canonical axon convergence and result in OSNs that target a variety of different glomeruli, presumably due to re-selection of variety of other ORs (Wang et al., 1998; Serizawa et al., 2003; Lewcock and Reed, 2004). Ultimately, the final OSN map is formed by the collaboration of numerous factors, of which the OR is one contributor. As discussed previously (Section 1.3D), the gradient expression of axonal guidance molecules such as Slit1, Semaphorin3F, Robo2, Neuropilin2, IGF, are important in OSN targeting (Cho et al., 2007, 2010; Takeuchi et al., 2010; reviewed in Mori and Sakano, 2013). Moreover, OSN axonal coalescence is dependent on gradient and mosaic expression of signaling and cell adhesion molecules such as kirrels, ephrinA and protocadherins (Serizawa et al., 2006; Cutforth et al., 2003; Mountoufaris et al., 2017). Finally, odorant stimulated activity of the OSNs play a role in pruning of mistargeted axons and refinement of glomeruli by influencing the expression of guidance receptors and adhesion molecules, including kirrels and neuropilin2 (Zou et al., 2004; Serizawa et al., 2006; Nakashima et al., 2013).
1.4.C The Genetics of OR Gene Regulation

1.4.C.1 Spatial Patterning of OR Expression

Initial studies suggested that ORs are expressed in spatially segregated, non-overlapping zones (Zones I-IV) along the dorsal-ventral axes, forming stripes along the anterior-posterior axis of the OE (Ressler et al., 1993; Vassar et al., 1993). More detailed studies later suggested that ORs are expressed continuously in multiple overlapping domains as opposed to co-expressing within well-defined and restrictive zones (Iwema et al., 2004, Strotmann et al., 1994; Miyamichi et al., 2005). Thus, the zonal boundaries do not exist in the OE, rather graded OR expression is evident along the dorsal-ventral axes in the OE.

It was initially proposed that the 5’ cis-regulatory sequences played a role in OR patterning (Qasba and Reed, 1998; Asai et al., 1996; Vassalli et al., 2002). Six conserved O/E-like promoter motifs were identified in genes of the Olfr37 subfamily, whose expression is restricted to the central region of the OE (Strotmann et al., 1994, 2006; Hoppe et al., 2003, Iwema et al., 2004). Promoters of dorsally expressing Class I ORs contains multiple O/E-like binding sites as compared to ventrally expressing Class II ORs, suggesting these cis regulatory sequences might play a role in OR patterning (Hope et al., 2006). Promoters of Class II ORs (e.g. the M171 subfamily) contain a homeodomain (HD) (LHX2 transcription factor binding motif) and O/E-like binding motifs (Michaloski et al., 2006; Clowney et al., 2011). Mutations in these motifs reduces the
representation of Class II ORs and alters their expression towards more ventral regions of the OE (Hirota and Mombaerts, 2004; Rothman et al., 2005).

Substituting the OR coding region with a β2 adrenergic receptor gene or a reporter gene (e.g. LacZ) mimics the spatial patterning of ORs, indicating that cis promoter/enhancer sequences govern this layer of regulation (Firestein et al., 2004). Deletion of the 5’ intron region in the MOR23 minigene construct abolishes the zonal expression pattern, suggesting that the OR intronic sequence may play a regulatory role in determining spatial patterning of OR expression (Vassalli et al., 2002). Finally, the transgene copy number and integration site can influence OR patterning (Lewcock and Reed, 2004, Zhang et al. 2007), suggesting that genomic context and local chromatin structure might also play a role in establishing exactly which portions of the OE a given OR gene is most likely to be selected for expression (Rothman et al., 2005; Zhang et al., 2007).

OR patterning established during embryonic stages persist into adulthood, as well as being maintained during the regenerative process after OE injury (Sullivan et al., 1995; Rodriguez-Gil et al., 2010; Schwob et al., 2002). Genetic ablation of particular receptor neurons regenerated the original OR patterning (Gogos et al., 2000). More globally, MeBr based unilateral OE injury approximately recapitulated the OR patterning of all ORs, indicating that spatial cues that establish developmental OR patterning persist into adulthood (Iwema et al., 2004).
1.4.C.2 Temporal Patterning of OR Expression

In situ hybridization experiments on embryonic OE suggested that OR mRNAs were first detected by E11.5 (Sullivan et al., 1995). More sensitive PCR approaches established an earlier timeframe for OR expression by E9 (Rodriguez-Gil et al., 2010). The onset of OR mRNA expression is concurrent with observing OSN soma in the developing OE but prior to the emergence of OSN axons (Rodriguez-Gil et al., 2010; Whitesides and LaMantia, 1996). Moreover, ORs are expressed in mice lacking the OB suggesting that OR expression precedes OB formation and is not dependent on the presence of the target tissue (Iwema et al., 2003; Sullivan et al., 1995).

The nATF5 protein, a transcription factor within the olfactory UPR pathway responsible for promoting Adcy3 transcription during the feedback process, is not expressed until E11 (Wang et al., 2012; Hansen et al., 2002), whereas, ORs expresses well before E11 (Rodriguez-Gil et al., 2010) suggesting that OR feedback is not necessary for initial OR gene choice. The progenitors that give rise to OSNs appear within the olfactory placode at E10.5 (Cuschieri and Bannister, 1975b; Farbman et al., 1992). At this time, the progenitor cells are located more apically, pointing to a possibility that these cells might give rise to mature OSNs at early embryonic stages and then migrate towards the basal layer at a later developmental stage (E14.5-E17.5) (Klein and Graziadei, 1983). Therefore, cells derived from olfactory placode at E10.5, such as OP6 and OP27
(Illings et al, 2002), may represent a developmental stage where OR genes are first being expressed and the first mature OSNs are being born.

The adult animal has an ongoing gradient of developmental stages present within the immature and mature OSN populations (see Section 1.3). Expression studies across this developmental gradient indicate that OR mRNAs, as well as OR proteins, are expressed in mature OSNs (OMP+) and immature OSNs (GAP43+) (Iwema et al., 2003; Magklara et al., 2011). However, some of the immature OSNs appear to express multiple ORs per cell at low levels, whereas all mature OSNs express only one OR at robust levels (Hanchate et al. 2015; Tan et al., 2015). This observation is further supported by electroolfactogram (EOG) studies indicating that immature OSNs can be responsive to broader odorant stimuli as compared to mature OSNs (Gesteland et al., 1982). These observations are mimicked in early versus late embryonic stages: at earlier embryonic stages (E16-E17), OSNs are generally not selective with responsiveness, whereas at later stages (E18), the OSNs become more selectively and robustly responsive (Gesteland et al., 1982). Therefore, the mechanisms that establish and/or maintain monogenic OR expression are not absolute during early OSN stages.

1.4.C.3 **OR Promoter Structures and OR Enhancer/LCR like Functions**

Transgene experiments indicate that as few as 300 bp upstream of transcription start sites is sufficient to capture an OR’s entire promoter (Vassalli et al., 2002). OR promoters are AT rich sequences that usually lack a TATA box (Sosinsky et
al., 2000; Lane et al., 2002; Michaloski et al., 2006) and CpG islands (Clowney et al., 2011; Plessy et al., 2011). OR promoters exhibit low levels of conservation (orthologs or paralogs) and exhibit only a few characteristic motifs (Clowney et al., 2011). OR promoters contain transcription factor binding motifs for homeodomain (e.g. LHX2, EMX2) and olfactory/early B transcription (O/E) transcription factors (Wang et al., 1997; Vassalli et al., 2002; Hope et al., 2003; Plessy et al., 2011). Deletion/mutation in LHX2 binding sites reduces the expression of nearby ORs, whereas knocking in multiple copies of these motifs increases the probability as well as levels of nearby ORs (Rothman et al., 2005; Hirota and Mombaerts et al., 2004; Vassalli et al., 2011).

The first OR enhancer discovered was called the “H region” and is located upstream of the MOR28 OR cluster (Lane et al., 2002). Deletion of this sequence abolished the expression of three downstream OR genes, suggesting that this was a shared enhancer (Serizawa et al., 2003). Addition of this sequence to a position immediately adjacent to an OR gene causes the proximal OR to be significantly over-represented in OSN populations (Nishizumi et al., 2007), substantiating its positive regulatory role. A second enhancer, called the “P element”, was identified at a different OR cluster, and like the “H region”, it is required for expression of a set of cis ORs (Bozza et al., 2009). To date, there are now >50 putative OR enhancers that have been identified and characterized across the OR regions of the mouse genome (Markenscoff-Papadimitriuo et al., 2014; Farber and Lane, submitted). These putative OR enhancers are identified
by the characteristic epigenetic marks, such as H3K27Ac, H3K4me1, and DNase hypersensitivity, as well as a H3K79 signature mark that differs from non-OR enhancers (Markenscoff-Papadimitriou et al., 2014). OR enhancers exhibit interchromosomal interactions with the expressed OR locus, and disruption of this intrachromosomal enhancer interaction resulted in mutigenic OR expression, suggesting a role for these enhancers in the establishment or maintenance of monogenic OR expression (Markenscoff-Papadimitriou et al., 2014). OR enhancers are enriched with motifs for putative transcription factors, such as Lhx2, Ebf, Foxj2, Cdx, C/EBPα, Bptf factors, whose roles are largely unclear (Markenscoff-Papadimitriou et al., 2014). However, the transcription factors Lhx2 and Ebf are shown to bind OR enhancers in a functionally cooperative manner that adds an additional layer of specificity in enhancer recognition (Monahan et al., 2017).

1.4.D Mutually Exclusive OR Expression as an Epigenetic Problem

Up to this point, I have discussed the differentiation of sensory neurons in the olfactory epithelium (1.3.C), and how the odorant receptor itself is important for the targeting of each OSN to the olfactory bulb (1.3.D), in addition to its odorant-binding function (1.4.A). In this way, specialization of each OSN accomplished by monogenic OR expression is fundamental to the proper function of the olfactory system (1.4.B). In the previous section (1.4.C), I have discussed genetic contributions to proper OR regulation, including promoter-proximal sequences
that govern spatial/temporal expression patterns and more distal cis enhancer sequences required for expression of a subset of clustered OR genes.

In this section, I turn to epigenetic contributions to proper OR regulation, and in particular, frame the monogenic OR expression problem as one dependent on chromatin modifications and nuclear organization. In the following paragraphs, I will first discuss insights gained about key factors and mechanisms involved with mutually exclusive transcription of alleles ("allelic exclusion"; 1.4.D1) and genes ("monogenic expression"; 1.4.D2) in other systems that leads to the premise that DNA methylation, histone modifications, and chromatin organization are core attributes of these regulatory phenomena. Next I will focus on the role for key regulatory factors that establish and maintain OR silencing and contribute to the establishment/maintenance of monogenic OR transcription (1.4.D3); my particular focus will be on two chromatin-modifying factors, LSD1 and G9a, that appear to be at the core of OR regulation and OSN differentiation. Finally, I will summarize the current working model in the field for mutually exclusive OR expression, including discussing weaknesses with this model that my thesis work challenges (1.4.D4).

1.4.D.1 Allelic Exclusion

Allelic exclusion is a process by which one of the two homologous alleles of the same gene is differentially expressed (Chess et al., 1994). There are two categories of allelic exclusion: non-random (programmed/deterministic) allelic
exclusion and random allelic exclusion (Chess et al., 2013). Non-random allelic exclusion (imprinting) occurs at imprinted genes (e.g. Igf2, Igf2r, H19, Dlk1, PWS-AS) (Zemel et al., 1992; Brandeis et al., 1993; Ferron et al., 2011; Nicholls et al., 1989). Random allelic exclusion occurs on the X chromosome (X inactivation), as well as for several genes on autosomal genes (e.g. immunoglobulin genes, interleukin genes, protocadherin genes, pheromone receptor genes, and odorant receptor genes) (Pernis et al., 1965; Held et al., 1995; Mostoslavsky et al., 2001; Gimelbrant et al., 2007; Chess et al., 1994).

Monoallelic expression provides a mechanism for creating greater cellular diversity and phenotypic heterogeneity (Chess et al., 2012; Gimelbrant et al., 2007; Jeffries et al., 2012). For example, in the olfactory system, monoallelic expression essentially doubles the diversity of sensory neuronal cell identities if the two OR alleles differ (even subtly) in odorant-binding capabilities. Monoallelic expression may also increase the combinatorial possibilities for proteins that form protein complexes inside the cell (Chess et al., 2012). Most importantly, monoallelic expression is a mechanism to control gene dosages, and the importance of dosage control for some genes is evident by mutations that disrupt allelic regulation that give rise to human diseases, such as Prader-Willi syndrome (Buiting et al., 1995), Angelman syndrome (Nicholls et al., 1989), and Wilm's tumor (Jirtle et al., 1999).

In addition to monoallelic expression, there are common epigenetic features of allelic exclusion shared among autosomal, X-linked, and imprinted
genes, including asynchronous DNA replication, differential DNA methylation, asymmetric distribution of post-translational histone modifications, non-coding RNA expression, and differential nuclear compartmentalization (reviewed in Eckersley-Maslin and Spector, 2014).

Asynchronous replication timing allows one allele to enter into S-phase earlier than the other allele (Chess et al., 1994; Singh et al., 2003; Zhou et al., 2004). Asynchronous replication is associated with differential epigenetic marking of the two homologous alleles (Lomvardas et al., 2006; Alexander et al., 2007). Asynchronous replication is coordinated across the entire length of autosomes, even though monoallelically expressed genes are interrupted by large segments of chromosomes containing bi-allelically expressing genes (Chess et al., 2012). Based on this observation, it has been proposed that allelic regulation on autosomes may be regulated by a similar/common mechanism as X-inactivation, only the former has much larger portions of the chromosome expressing biallelically (i.e., that “escape” allelic inactivation) than the latter.

Differential DNA methylation is most well characterized for imprinted genes. For example, germline imprinting of the \( H19 \) non-coding RNA and the \( Igf2 \) growth factor gene arises from DNA methylation of the Imprinting Control Region (ICR) at this locus in the male germline (Zemel et al., 1992); in the female germline, the regulatory protein CTCF protects the ICR from DNA methylation (Brandeis et al., 1993). The consequence of DNA methylation of the paternal allele is the repression of the nearby \( H19 \) gene, whereas the more distal \( Igf2 \)
gene is able to express normally (Stoger et al., 1993; reviewed in Li et al., 1993); in contrast, the absence of DNA methylation on the maternal allele permits H19 gene expression, while CTCF binding interferes with the \( Igf2 \) promoter from interacting with its enhancer (Sasaki et al., 1992; reviewed in Reik and Murrell, 2000).

Differential non-coding RNA expression is most well characterized for X-inactivation. X inactivation is a dosage compensation mechanism for X chromosome specific genes (Brockdorff and Turner, 2015). A random choice is made between the maternal or paternal X chromosome in each cell (Avner and Heard, 2001). X-inactivation is initiated at the X Inactivation Center (XIC) (Jeon et al., 2012; Avner and Heard, 2001), which is analogous to the ICR for imprinted genes. In undifferentiated ES cells, the \( Oct4 \) pluripotency factor represses the expression of the Xist non-coding RNA (ncRNA), the key RNA that will define the eventual inactive X in a given cell, thus preventing X-inactivation until differentiated cells are produced (reviewed in Jeon and Lee, 2011). At the onset of differentiation, the two homologous XICs pair, a process thought to asymmetrically distribute regulatory factors between the two X chromosomes (Navarro et al., 2010). The presumptive inactive X (Xi) is associated with expression of \( Rnf12 \) (Xist activator protein), as well as several ncRNAs. The \( Jpx \) and \( Ftx \) ncRNAs positively regulate \( Xist \) expression on the Xi, whereas, \( Tsix, Xite, Tsx \) ncRNAs negatively regulate \( Xist \) expression on the Xi (Jonkers et al., 2009; Donohoe et al., 2009; Spencer et al., 2011; reviewed in Tukiainen et al., 2017).
contrast, the Tsix ncRNA is expressed on the active X (Xa), which recruits DNMT3a to methylate a CpG island located at the promoter of Xist, thereby silencing it (Nesterova et al., 2008). Ultimately, the presence of Xist on just the Xi leads to allele-specific inactivation by recruitment of additional repressive factors, including the transcriptional repressor, YY1, that facilitates tethering of Xist transcripts in order to mediate spreading of Xist in cis along the entire Xi (Zhao et al., 2008; reviewed in Jeon and Lee, 2011). Eventually, Xist coats the entire Xi and recruits a number of repressive complexes, such as the polycomb repressive complex (PRC2), histone deacetylases, DNA and histone methyl transferases, and the histone variant macroH2A1.2 (reviewed in Galupa and Heard, 2018; Panning, 2008; Zlotorynski et al., 2015).

OR genes, like X-linked genes, exhibit random allelic exclusion. Although almost nothing is known about OR allelic regulation, there exist several parallels that suggest the potential for common mechanisms. As mentioned previously, OR genes (like X-linked genes) are subject to asynchronous replication, and the early- or late-replication timing is coordinated along the entire length of the chromosome (Chess et al., 2013). OR gene clusters are predicted to express long non-coding RNAs (Plessy et al., 2012; Clowney et al., 2011) and these loci, like the X chromosome, are densely populated with LINE repeats (Lyon et al., 2000; Young et al., 2003; Kambere and Lane, 2009). Both ncRNAs and LINE repeats have been shown to be important for spreading of heterochromatin along the chromosome (Lyon et al., 2006; Pinter et al., 2012; reviewed in Lee et al., 2011).
The inactive X chromosome condenses to forms Barr bodies to maintain efficient silencing of Xi specific genes (reviewed in Galupa and Heard, 2018), and similarly, inactive OR genes aggregate to forms a single heterochromatin chromocenter in the nucleus (Clowney et al., 2012).

Nevertheless, it is possible that allelic exclusion of OR genes is inconsequential for proper OR regulation, given that OR regulation is also subject to mutually exclusive expression involving all OR loci, not just the two alleles of a given OR locus. That is, if mechanisms are in place to ensure only 1 of ~1,000 OR gene loci are expressed per cell, presumably, this same mechanism would ensure that only one of the two alleles are expressed as well. It is possible that an X-inactivation-like process is important in the context of OR regulation mostly for its “heterochromatic spreading” mechanisms, since initial silencing of all OR loci occurs prior to the selection of OR gene/allele for expression (Kembere et al., 2009; Magklara et al., 2011).

1.4.D.2 Monogenic Expression

The key OR regulatory paradigm is the expression of one and only one OR gene (and allele) per sensory neuron, which gives that neuron a unique functional identity. The first evidence of monogenic regulation came from the Sakano group, which showed that two OR transgenes with identical coding and promoter sequences never co-express in the same OSN even though they integrate into a common genome location (Serizawa et al. 2000). These
transgenes are targeted to neighboring but distinct glomeruli suggesting that these transgenes act like two distinct ORs genes. Interestingly, when beta-adrenergic receptor is driven under the control of OR promoter, it also expresses monogenically with respect to the other OR genes in the genome (Firestein et al., 2004). Subsequent RNA FISH experiments showed that probes designed against different OR gene sequences never co-express in the same cell (Firestein et al., 2004; Clowney et al., 2012), and single-cell RNA-seq experiments demonstrated that each mature OSN contains only one OR transcript above background noise levels (Hanchate et al., 2015; Tan et al., 2015; Saraiva et al., 2015; Scholz et al., 2016; Gadye et al., 2017).

Monogenic regulation is also a property of immunoglobulin (Ig) gene expression, accomplished by programmed DNA rearrangements that results in only one Ig gene positioned so it can be transcribed (Hozumi et al., 1976). This type of mechanism was tested directly for OR loci by cloning a mouse from DNA isolated from a single mature OSN, however, the cloned mouse expressed the entire OR repertoire, a result not predicted by this hypothesis assuming that the presumptive DNA rearrangement leading to one OR expression cassette would be irreversible (Eggan et al., 2004; Li et al., 2004).

A second hypothesis gain momentum involving an iterative feedback loop. When an OR coding sequence is replaced with a GFP/LacZ reporter, OSNs that selected the reporter locus “switched” to another OR gene, suggesting that the expression of the OR coding sequence is required for monoallelic OR
expression (Serizawa et al., 2003; Lewcock and Reed, 2004). Moreover, when the OR coding sequence is altered by introducing merely a single base-pair frame-shift insertion/deletion, such that the OR coding sequence is essentially intact but no OR protein can be produced, the same OR “switching” phenomenon was observed (Lewcock and Reed, 2004). From these data, it was proposed that the OR protein itself initiated a negative feedback loop that prevents an additional OR selection. Formally, such a feedback loop might be sufficient to account for monogenic OR expression if the process of selection is slow enough so that the gradual appearance of super-threshold OR protein can interfere with a second selection event (Tan et al., 2013). Recently, a model for this negative feedback loop was elucidated by the Lomvardas group: the presence of OR protein activates the unfolded protein response (UPR) pathway, resulting in a kinase cascade that, among other targets, activates adenyl cyclase-3, which down-regulates OR transcriptional regulators, such as LSD1 (Dalton et al., 2014). However, four problems for this model are: (1) no OR transcriptional regulator, including LSD1, has been shown to be necessary and sufficient for OR activation events (I will return to a fuller discussion of LSD1 in section 1.4.D3 below); (2) monogenic OR expression is nevertheless observed in a cell line that does not appear to express OR protein (Kilinc et al., 2015); (3) mature OSNs lacking the Adcy3 component in the UPR feedback loop nevertheless express a single OR gene (Hanchate et al., 2015); (4) some of the immature OSNs in the mouse olfactory epithelium express only one OR gene per cell even though components
of the UPR pathway are not present in these cells. At this time, it is not clear whether the feedback loop is necessary and sufficient to account for monogenic OR expression, whether it serves more of a post-selection maintenance/commitment function, or whether it merely functions to ensure OSNs “switch” (or continue the selection process) in the event an OR pseudogene is initially selected.

The fact that single cell transcriptome studies indicate that immature OSNs express low levels of multiple ORs, and mature OSNs express a single dominant OR (Hanchate et al., 2015; Tan et al., 2015; Saraiva et al., 2015; Scholz et al., 2016; Gadye et al., 2017) suggests the presence of an iterative process whereby monogenic OR expression results from a refinement process. That is, these studies indicate that initially not just a single OR allele is activated in immature OSNs rather a subset of ORs from multiple different chromosomal locations are activated, and monogenic OR expression is achieved by reducing the number of competitor ORs during the differentiation of the OSNs (Hanchate et al., 2015; Nagai et al., 2016). A recent genetic study demonstrated a post selection refinement (PSR) process that ensures monogenic OR expression (Abdus-Saboor et al., 2016). According to this study, initially immature OSNs express multiple ORs at low levels, which activates the PSR process, the outcome of which is dependent on the relative levels of competing OR transcripts. The implication of this result is that a subset of ORs might compete with each other for limited factors in order to ultimately achieve monogenic OR expression, and
that the OR transcripts themselves may be involved in the competitive process (Abdus-Saboor et al., 2016). An example of such a model (although there is no evidence for it) might be an RNA-mediated trans-silencing effect, whereby the RNA transcripts from one OR gene locus silence competing OR loci in trans, and thus, once a particular OR locus starts to “win” the competing loci increasingly “lose”, and this initial skewing would continue until the process is completed.

In the previous section (1.3), I had introduced the concept of a locus control region, or LCR. LCRs are enhancers that co-regulate sets of target genes in a mutually exclusive manner. For example, the fetal-to-adult globin genetic switch is regulated by the upstream LCR (Paterson et al., 2012; reviewed in Sankaran and Orkin, 2013). The globin LCR contains a cluster of DNase hypersensitivity sites (HSs) that loop over to the promoter of one of the four globin genes in a mutually exclusive and developmental stage-specific manner (Deng et al., 2014). The exact mechanism by which this LCR functions is not known, however, it is hypothesized that these HSs function to exclusively deliver the key transcription activators to one globin promoter (Bender et al., 2012). During adulthood, the promoter of the γ-globin locus (fetal globin) is bound by Bcl11A (repressor) and NuRD (repressive chromatin remodeling complex) (Sankaran et al., 2008, 2010). As a result, the upstream globin LCR cannot loop over the γ-globin gene promoter; it instead loops over to the promoter of the adult β-globin gene (reviewed in Vinjamur, Bauer and Orkin, 2018). At OR loci, LCR-like sequences (e.g. H and P elements) have been identified that are
necessary for the selection of proximal OR genes (Serizawa et al., 2003; Bozza et al., 2009; Khan and Mombaerts, 2011). Moreover, placement of an OR LCR immediately upstream of an OR gene results in the hyper-selection of that OR in the olfactory epithelium (Nishizumi et al., 2007). Could mutually exclusive OR expression be due to LCR competition, as is the case for mutually exclusive expression of globin genes? To investigate this hypothesis, the H-region LCR was deleted, however, this deletion had no effect on the expression of OR genes residing on other chromosomes (Fuss et al., 2007; Nishizumi et al., 2007; Khan et al., 2011), indicating that monogenic OR expression is not accomplished by competition for a single enhancer element. Nevertheless, it has been shown that the LCRs from various OR loci aggregate in the nucleus (Markenscoff-Papadimitriou et al., 2014), so it remains possible that the “winning” OR gene is the one that has recruited to its promoter the full collection of available LCRs; such a model would predict that the deletion of any one LCR might render the cis OR genes incompetent to “win” such a competition, nevertheless, would not prevent another locus from “winning” through recruitment of all other/any LCRs present. A test of this hypothesis would require the deletion of numerous or all LCRs and/or engineering of a cell in which LCRs failed to aggregate.

In our lab, we have considered another iterative and competitive model for how monogenic OR expression might be achieved: competition for a unique chromatin/epigenetic state. Chromatin-based regulation is at the core of all monogenic-monoallelic expression paradigms (e.g. immunoglobulins and T-cell
receptor genes, Hox genes, and protocadherins). For example, in B cell progenitors, both alleles of immunoglobulin gene segments (heavy chain-IgH and light chain-Igκ and Igγ) and T cell receptor gene segments (TCR-α, -β, -γ, -δ) contain high levels of DNA methylation (reviewed in Cedar and Bergman, 2008). Later in B cell development, allelic exclusion is initiated by asynchronous replication, histone acetylation followed by DNA demethylation at the eventual active allele, while the other non-chosen allele is recruited to heterochromatin via lamins at the nuclear periphery (reviewed in Krangel et al., 2009). At later stages, the rearranged (active) allele remains in a euchromatin compartment, and is marked by histone acetylation (H3K27Ac) and histone methylation (H3K4me3), with corresponding high rates of transcription (Mostoslavsky et al., 2004; Alt et al., 2007, Subramanyam and Sen, 2012). The unarranged (repressed) allele remains associated with pericentromeric heterochromatin, and is marked by DNA hypermethylation and repressive histone modifications (H3K9me3), adopting a condensed chromatin structure (reviewed in Cedar and Bergman, 2008).

Another example is the monogenic regulation of protocadherin (Pcdh) genes that provide unique identity to each neuron for self-avoidance (Esumi et al., 2015; reviewed in Chen and Maniatis, 2013). The promoters of variable exons are stochastically chosen for activation via cis-regulatory enhancers (HS5-1 and HS-7), resulting into alternative splicing of these exons (Ribich et al., 2006; Zipursky and Grueber, 2013). During activation, the HS5-1 and HS-7 enhancers
loop to the chosen promoter, stabilized by CTCF/cohesion, and mark the selected gene with active H3K4me3 marks that facilitates recruitment of this allele into a transcriptional hub (reviewed in Chen and Maniatis, 2013). The remaining promoters are marked by DNA methylation, repressive histone modifications (H3K9me3 and H4K20me3), lose CTCF/cohesin complexes, and segregate away from the transcriptional hub (Kawaguchi et al., 2008; Guo et al., 2012; Monahan et al., 2012; Dekker and Misteli, 2014; Kahayova et al., 2011; reviewed in Chen and Maniatis, 2013; Lomvardas and Maniatis, 2016).

Hox clusters contain monogenically expressing developmental genes, which are critical for normal body axes development (Noor-dermeer et al., 2011; reviewed in Andrey and Duboule, 2014). In ES cells, Hox genes are marked with bivalent domains containing H3K4me3 and H3K27me3 marks (Bernstein et al., 2006; Rada-Iglesias et al., 2011). Later in development, these bivalent domains are resolved by EZH2 (a polycomb group protein), as well as by spatial and temporal activity of the flanking enhancers, resulting in monogenic expression of Hox genes in a cell type dependent manner (Montavon et al., 2011; Ferraiuolo et al., 2010; Wyngaarden et al., 2011). For example, cells responsible for limb development activate 3’ Hox genes via telomeric enhancers that segregate into euchromatic domains, while suppressing 5’ Hox genes responsible for digit development within heterochromatic domains (reviewed in Lyons et al., 2014).

In each of these examples, the establishment of monogenic/monoallelic expression depends on differential epigenetic structures between the one active
and the competing silenced alleles. If some of the chromatin modifying factors that give rise to these differential chromatin features are limiting, it is feasible that the winning locus might be the first one to achieve a stably active chromatin structure. Chapter 2 of this thesis more deeply explores such an iterative model for monogenic OR regulation that involves a competitive balance between two histone marks, repressive H3K9me3 and activating H3K4me3 that might establish the inactive and active OR alleles, respectively. Two key chromatin factors are involved with the regulation of these two marks in particular, LSD1 and G9a. I next turn attention to discuss evidence for these two factors being involved with OR regulation and OSN differentiation more generally.

1.4.D.3 Chromatin Factors: The Role of LSD1 and G9a in Singular Odorant Receptor Gene Regulation

As discussed in section 1.4.D1 and D2, various epigenetic factors (e.g. DNMTs, polycomb repressive complex, non-coding RNAs, histone methyl transferases) are involved in establishing monogenic-monoallelic expression in random (X inactivation) and non-random (imprinting) allelic exclusion. In all monoallelically expressing biological systems described in section 1.4.D1 and D2, activating H3K27 histone acetyl transferases and H3K4 histone methyl transferases function on the chosen allele, whereas repressing H3K27, H3K9 and H4K20 methyl transferases function to silence the remaining alleles (Rada-Iglesias et al., 2011; Chen and Maniatis, 2013; Cedar and Bergman, 2008).
G9a is a histone methyl transferase involved with the silencing of non-selected OR genes via H3K9 methylation (Lyons et al., 2014); LSD1 is a demethylase involved with the removal of H3K9 methylation on the chosen OR (Lyons et al., 2013). Immunohistochemical and transcriptome experiments suggest that G9a and LSD1 expression is high in neurogenin-1 (Ngn1) positive immature neuronal progenitors and low in OMP positive mature OSNs, an expression pattern that correlates with the timing of OR selection during OSN differentiation (Lyons et al., 2013; Krolewaski et al., 2013; Kilinc et al., 2016). Our group found that LSD1 is compartmentalized in early post-mitotic cells (both in vitro and in vivo), and this LSD1 compartment co-localized with ~1 OR locus per nucleus, providing further momentum to the hypothesis that LSD1 might be limiting in OR regulation (Kilinc et al., 2016).

To further investigate LSD1 function, the Lomvardas group knocked out LSD1 at three distinct developmental time-points in the developing OE: prior to OR choice in the GBC cell population, during OR choice in immature neurons, and after OR choice in mature OSNs (Lyons et al., 2013). LSD1 depletion prior to OR choice resulted in significant reduction of OR expression levels, aborted terminal OSN differentiation, and produced a glomerular targeting deficit of OSN axons (Lyons et al., 2013). Importantly however, ORs were still expressed, albeit at much lower levels, suggesting that LSD1 might play a role in attaining robust OR expression, but does not appear to be required for initial OR activations (Lyons et al., 2013). LSD1 depletion during or after OR choice (e.g. immature or mature
OSNs) had no significant phenotypes, suggesting that LSD1 activity is dispensable after an OR is chosen (Lyons et al., 2013); this conclusion is consistent with the endogenous down-regulation of LSD1 following OR activations during OSN differentiation (Dalton et al., 2013). Moreover, ectopic expression of an OR transgene in mature OSNs is sustained in the absence of LSD1, suggesting that H3K9 demethylation activity of LSD1 is not required to maintain robust OR expression once established (Lyons et al., 2013). One possible interpretation of these results is that LSD1 might demethylate the repressive H3K9 mark following an initial OR selection in order to eventually enable robust expression levels during a post-selection commitment phase.

In addition to H3K9 demethylation, LSD1 is also responsible for H3K4 demethylation in several other systems, and therefore, LSD1 can function as a repressor (H3K4 demethylation) as well as an activator (H3K9 demethylation) depending on context (Wang et al., 2007; Whyte et al., 2012; reviewed in Shi et al., 2013; Maiques-Diaz and Somervaille, 2016). Interestingly, overexpression of LSD1 in mature OSNs inhibits OR expression suggesting that LSD1 might also target H3K4 demethylation at OR loci (Magklara et al., 2011; Lyons et al., 2013). There was no previously published experimental data demonstrating the importance of H3K4 demethylation in OR regulation, however, the results I report in Chapter 2 of this thesis provide momentum for this hypothesis. In addition, a computational model was recently published from the Xie group that demonstrated a theoretical framework for how LSD1 might be involved with
monogenic OR regulation without necessarily being required for OR activation per se (Tan et al., 2013). According to these theoretical models, OR singularity (expression of a single OR allele of a single OR gene) is achieved by equilibrium between opposing histone methyltransferase activities that add repressive H3K9 methylation and histone demethylase activities that remove repressive marks. Specifically, enzymes like G9a, which methylates H3K9 (to silence), in conjunction with LSD1, which demethylates H3K4 (to deactivate), might be in delicate balance with the opposing activity of LSD1, which demethylates H3K9 (to de-repress). I will return to this model in light of my results in which perturbation of LSD1 resulted in a subtle disruption of monogenic regulation in which some cells exhibit polygenic OR expression, suggesting a shift in a putative chromatin state equilibrium among competing OR loci (Chapter 2).

As mentioned previously, G9a and GLP (G9a-like protein) have been previously characterized as H3K9 methyltransferases important in the initial silencing of OR loci in early OSN progenitors (Lyons et al., 2014). Surprisingly, double knockouts of G9a and GLP substantially decreased expression of most ORs, a result not anticipated if G9a/GLP functions in global OR repression (Lyons et al., 2014). These double-knockout mice contained some OSNs that expressed multiple ORs per cell, suggesting that G9a/GLP nevertheless is important for establishing monogenic OR expression (Lyons et al., 2014). However, G9a/GLP-depleted OSNs appeared to express a modest number of ORs, and not the entire OR repertoire, as might be predicted with a global loss of repression (Lyons et
al., 2014). A similar multigenic OR phenotype is observed with pharmacologic inhibition of G9a in the fish (Ferreira et al., 2014). A simple hypothesis to account for these data is that G9a/GLP probably contributes to the global silencing of OR genes since the absence of these enzymes results in the expression of additional ORs per cell (i.e., additional de-repression), however, the loss of G9a/GLP seems to have a subtle impact (e.g., shift of an equilibrium state), as opposed to an absolute and global disruption of OR silencing. Moreover, the overall decreased OR expression levels in these mice might be due to the failure to fully acquire a monogenic state – robust OR expression might depend on singularity if there is a limiting resource that must stably associate with a single, selected OR gene, as opposed to being competed away by other ORs.

To summarize this section, I have described both genetic (e.g., LCRs) and epigenetic (e.g., H3K9 and H3K4 methylation) factors that are involved with the establishment of monogenic OR expression. I have also introduced key regulators, LSD1 and G9a, that likely play roles in establishing and/or maintaining monogenic OR expression. I have provided momentum for models that include iterative feedback loops, enhancer competition, and mutually exclusive epigenetic structures that might collaborate in this process. In the following section (1.4.D4), I elaborate on the favored model in the field to account for mutually exclusive OR expression, as well as critically evaluate the
weaknesses of this model that will frame questions addressed by my own thesis work.

1.4.D.4 Establishment of Mutually Exclusive OR Expression

The Lomvardas group has been at the forefront of most of the published work that has framed our understanding of OR regulation – this group has identified and characterized OR LCRs (Markenscoff-Papadimitriou et al., 2014), elucidated the molecular mechanisms for the OR-mediated negative feedback loop (Dalton et al., 2013), described the epigenetic properties at OR loci, including H3K9 and H3K4 methylation states (Magklara et al., 2011), and produced several knockout mice, including G9a and LSD1, that clarify roles for these regulatory proteins (Lyons et al., 2013, 2014). Figure 14 is the Lomvardas interpretation of all of these data into coherent model (Lomvardas and Maniatis, 2016). According to this model, all OR genes are decorated with H3K9me2 in the basal stem cells of the OE, and later acquire H3K9me3 (as well as H4K20me3, also a repressive mark) as these cells transition to immature neurons in order to achieve a fully repressive state (Magklara et al., 2011). These repressed ORs from multiple different chromosomes aggregate to form a centrally located heterochromatin foci, which excludes ORs from accessing Pol factories (Clowney et al., 2012). During OSN differentiation, the single OR allele chosen for expression is liberated from heterochromatin foci by LSD1-mediated H3K9 demethylation, thus liberating the selected OR from the heterochromatic compartment (Lyons
et al., 2013). This OR allele acquires H3K4me3 marks via an unknown H3K4 methyltransferase and consequently expresses at a high level (Magklara et al., 2011). OR transcripts from the expressed locus activate UPR pathway that induces Adcy3 expression and activates the negative feedback loop (Dalton et al., 2013). The negative feedback loop commits the cell to its chosen OR by downregulating LSD1, which prevents additional OR activations (Dalton et al., 2013; reviewed in Dalton and Lomvardas; Monahan and Lomvardas, 2015).

There are several problems with this interpretation. In the above model, only one OR allele is activated by LSD1 in the immature OSNs (Reviewed in Monahan and Lomvardas, 2015). A prediction of this model would be the complete loss of OR expression in LSD1 knockout (KO) mice, since LSD1 serves a gatekeeping function by selectively removing the repressive mark, and therefore, LSD1 depletion should result in no ORs being liberated. However as discussed in the previous section, ORs are still expressed in LSD1 knockout mice (Lyons et al., 2013). The authors account for this result by suggesting that not all OR genes might be dependent on LSD1 for liberation, however if this were the case, then these same ORs would co-express with the LSD1-selected OR in wild-type cells according to the model. A second prediction of this model is that LSD1 over-expression, if anything, might result in additional OR expressions per cell; instead, LSD1 over-expression seems to have a repressive effect in mature OSNs (Lyons et al., 2013).
Figure 14: The “Lomvardas Model” for OR Gene Regulation
From top to bottom, the DNA containing hypothetical OR cluster OR1, OR2, OR3 is shown. From ES to OE specific stem cell development, the OR clusters are marked with H3K9me2, a facultative heterochromatin. Transitioning from the OE specific basal stem cells to immature OSNs, the OR clusters acquire the constitutive heterochromatin marks, H3K9me3 and H4K20me3. OR choice is made somewhere along the terminal differentiation of the immature OSNs to mature OSNs. LSD1 removes H3K9 specific methylation marks. The expressed OR2 locus acquire H3K4me3, transcriptionally active chromatin marks. OR protein mediated Adcy3 signal to downregulate LSD1 and stabilizes the OR choice (figure adopted from Lomvardas and Maniatis, 2016).
A third prediction of this model is that monogenic OR expression should be absolute – the model cannot easily account for some immature neurons expressing multiple ORs, and eventually refining to a single OR during maturation (Hanchate et al., 2015; Tan et al., 2015; Saraiva et al., 2015; Scholz et al., 2016; Gadye et al., 2017). A fourth prediction of this model is that G9a/GLP knockout mice would fail to position any OR into the heterochromatic compartment due to loss of H3K9 methylation, and therefore, should no longer depend on LSD1 for de-repression, and yet, these mice do not express all ORs in every cell (Lyons et al., 2014). A fifth prediction of this model is that failure to recruit OR genes into the heterochromatic compartment would similarly result in multi-genic OR expression, and while this appears to be the case in mature OSNs (Clowney et al., 2012), our lab has published that chromocenter recruitment is not a prerequisite for monogenic OR expression in an olfactory cell line that models an immature neuron (Kilinc et al, 2014). This model also does not seem to account for differing levels of OR expression during differentiation (immature versus mature OSNs) or in different genetic backgrounds (e.g., in LSD1 or G9a knockout mice). The model also does not account for any differences in allelic regulation (i.e., both parental alleles are described as silenced by H3K9me3 and compartmentalized together); however, the Malnic group reported that the two homologous alleles of a given OR are segregated into different spaces, one near constitutive heterochromatin.
(H3K9me3) and the other near facultative heterochromatin (H3K27me3) (Armelin-Correa et al., 2014a, b).

Perhaps most importantly, the Lomvardas model does not distinguish between “establishment” and “maintenance” of monogenic OR expression, nor between “specification” and “commitment” to a given OR gene, both of which are likely important aspects of this regulatory paradigm. For example, the “establishment” of mutually exclusive OR expression might not depend on repressive heterochromatization or compartmentalization of competing OR genes, whereas long-term maintenance of mutually exclusive OR expression through the changing gene-expression landscape during OSN differentiation might obligate more robust ongoing silencing of these competing ORs. Similarly, the initial specification of one or more OR expression events per cell at a stage when enhancer activity is low, expression levels are low, and/or the incidence of activation is slow, might have very different epigenetic requirements (histone marks, nuclear compartmentalization) than the post-refinement/commitment stage, when enhancers might be fully active and the mature OSN is now driving high levels of OR expression. As I now discuss in the last section of this Introductory chapter (1.5), my thesis work takes advantage of a model system, the OP6 cell line, that we believe isolates a pre-commitment, pre-maintenance state that might be difficult to achieve in vivo. Using this system, my research has focused on the investigation of the two key regulatory proteins described above: LSD1 (Chapter 2) and G9a (Chapter 3).
1.5 The OP6 Model System for Study of Odorant Receptor Gene Regulation and Olfactory Sensory Neuronal Development

Two clonal cell lines, OP6 and OP27, were derived from E10 mouse olfactory placode (OP) (Illings et al., 2002). These lines were immortalized using a temperature-sensitive allele of the large T antigen. Based on transcriptional profiling, OP6 cells are staged at the intermediate-late OSN developmental lineage (Illings et al., 2002; Pathak et al., 2009). At the permissive temperature (33°C), OP6 cells expand in culture and remain undifferentiated. Upon switching to the non-permissive temperature (39°C), the large T antigen deactivates, and with further treatment with all-trans retinoic acid, OP6 cells differentiate into bipolar odorant receptor neurons expressing mature OSN markers such as Golf, Adcy3, OCNC1, and OMP (Illings et al., 2002; Pathak et al., 2009).

The onset of OR expression in vivo occurs by E9.5 (Rodriguez-Gil et al., 2010), which is prior to the embryonic stage (E10) from which the OP6 cells are derived (Illings et al., 2002). Therefore, it is not surprising that OP6 cells express OR genes. Using degenerate PCR followed by sequencing, as well as RNA FISH, the Lane group established that single OP6 cells express OR genes monogenically and monoallelically just like the singular OSNs in vivo (Pathak et al., 2009; Kilinc et al., 2015; Vyas et al., 2017). This OR singularity is established in the undifferentiated OP cells and persists after OP6 cell differentiation (Pathak et al., 2009). Initially, the Roskam group suggested that OP6 cells express only ~one (or a very small number) of OR genes in the population, suggesting that OR
choice might be stably inherited through the cell cycle (Illings et al., 2002). Later, the Lane group showed that OP cells express at least 100 different ORs from multiple different chromosomal locations, indicating that OR re-selection is occurring (Pathak et al., 2009; Vyas et al., 2017; Noble et al., 2018). Our lab also recently showed that OR choice is not stable through the cell cycle, suggesting that the epigenetic structure associated with OR activation is not a heritable property (Noble et al., 2018). When OP6 cells re-select an OR, there is a clear, intrinsic bias for subsets of OR genes, which are consistently over-represented in OP6 cell cultures (Pathak et al., 2009; Nobel et al., 2018).

The expressed OR in each OP6 cell expresses at a significantly lower level than OR expression within mature OSNs in vivo (Pathak et al., 2009; Kilinc et al., 2015). As noted previously, robust OR expression in vivo appears to occur at later commitment stages of OSN differentiation (Clowney et al., 2012; Lyons et al., 2013), and therefore, the lower OR expression in OP6 cells might indicate a pre-commitment/immature status. Moreover, unlike mature OSNs in vivo (Clowney et al., 2012), the inactive ORs do not aggregate into one or small number of chromocenters within the nucleus; chromocenter aggregation is evident upon further differentiation of OP6 cells (Kilinc et al., 2015). We do not know whether these differences between OP6 cells and mature OSNs simply indicate that OP6 cells capture a more premature (perhaps even transient) stage of OSN differentiation, or whether these differences are a consequence of immortalization (e.g., rapid re-entry into the cell cycle might prevent significant
accumulation of OR transcripts prior to re-selection, or it might interfere with progressive chromocenter aggregation).

It is also unclear whether OR protein is produced in OP6 cells and/or whether OP6 cells are able to trigger the UPR negative feedback loop. First, the level of OR expression might be below the threshold level that is required to activate ER stress-induced UPR feedback (Dalton et al., 2014). Second, the initial onset of nATF5 activity that is dependent on the translation of OR transcripts, an important component of UPR pathway, is not active until E11.5, whereas OP6 cells are derived from an earlier stage of development when OR transcripts express at subthreshold levels, perhaps unable to activate UPR pathway (Wang et al., 2012; Hansen et al., 2002; Pathak et al., 2009). Third, undifferentiated OP6 cells seem to retain OR mRNA within the cell nucleus, suggestive of a missing mRNA-export signal/chaperone in these cells, and therefore, we expect little or no translational product (Pathak et al., 2009). Fourth, OP6 cells express LSD1, which is down-regulated in vivo following the UPR-mediated feedback loop (Dalton et al., 2013), suggesting that OP6 cells at least do not execute the complete feedback loop. Therefore, the frequent OR “switching” in OP6 cells might be due to a non-operational negative feedback loop that normally prevents “switching” in vivo (Magklara et al., 2011; Lyons et al., 2013; Shykind et al., 2005).

The OP6 cell line represents an opportunity to study monogenic OR regulation in an unusual cellular context. OP6 cells are isolated from in vivo
signals that would normally drive differentiation/developmental progression, and therefore, the cell line captures a particular immature state that might be transient *in vivo*. Perhaps due to a broken feedback loop, OP6 cells are constantly “switching” from one OR gene to another (yet always monogenic), which provides a unique opportunity to investigate OR selection mechanisms isolated away from downstream maintenance and/or commitment events. And due to continuous re-entry into the cell cycle, OP6 cells represent a unique opportunity to study the stability of the OP6 differentiated state; i.e., in terms of OSN identity (as defined by OR choice), as well as differentiation/de-differentiation/trans-differentiation properties of cell line.

This thesis uses the OP6 cell line to investigate the epigenetic regulation of mutually exclusive OR expression and cellular plasticity within the OSN lineage. My work is specifically focused on the role of LSD1 in monogenic-monoallelic OR gene expression and the role of G9a in OSN lineage plasticity. Much of the focus in the field has been on an activating role for LSD1 in OR regulation via removal of repressive H3K9 methylation marks (Lyons et al., 2013). In this thesis work, I discovered a repressive role for LSD1, presumably accomplished by removal of activating H3K4 methylation marks on OR genes (Chapter 2; Vyas et al., 2017). G9a-mediated OR heterochromatinization has been proposed to facilitate singular OR expression by silencing non-expressed ORs through addition of H3K9 methylation marks (Lyons et al., 2014). In this thesis work, I perturbed G9a expression in OP6 cells in order to further
investigate its role in OR regulation, however, I discovered a completely unexpected phenotype: G9a perturbation caused global chromatin reorganization and apparent reprogramming of OP6 cells. In Chapter 3, I preliminarily characterize this OP6 cell de-differentiation. Finally, in Chapter 4, I describe unpublished data with emphasis on future experimental questions raised by my work.

References


absence of normal projection neurons or GABAergic interneurons. *Neuron, 21*(6), 1273-1282.


Noble, J. C., Meredith, D., & Lane, R. P. (2018). Frequent and biased odorant receptor (OR) re-selection in an olfactory placode-derived cell line. PLoS One, 13(9), e0204604. doi: 10.1371/journal.pone.0204604


Wu, H., Coskun, V., Tao, J., Xie, W., Ge, W., Yoshikawa, K., . . . Sun, Y. E. (2010). Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. Science, 329(5990), 444-448. doi: 10.1126/science.1190485


This portion of my thesis work aimed to characterize a role for LSD1 in OR gene co-regulation. The function of the mammalian olfactory system depends on specialized olfactory sensory neurons (OSNs) that each express only one allele (“monoallelic”) of one odorant receptor (OR) gene (“monogenic”) from a large repertoire of OR genes encoded in the genome (Buck and Axel, 1991). Initially, all OR genes are repressed by H3K9me3 and H3K20me3 repressive histone marks (Magklara et al., 2011). During activation of a single OR allele, the H3K9me3 mark is removed and the H3K4me3 mark is added (Magklara et al., 2011). The high level of OR expression triggers a negative feedback loop that down-regulates LSD1, which is thought to prevent selection of another OR (Lyons et al., 2013; Dalton et al., 2013). OR switching is evident at a low incidence in immature OSNs (Shykind et al., 2004). During switching, the previously expressed OR allele is demethylated at H3K4 and a newly chosen OR allele is demethylated at H3K9 (Magklara et al., 2011). A candidate demethylase, LSD1, has been shown to have substrate specificity for both H3K9 and H3K4 demethylation, and this protein is expressed in the OSN lineage in immature OSNs prior to stabilization of OR choice (Shi et al., 2004; Metzger et al., 2005; Lyons et al., 2013; Krolewaski et al., 2013; Kilinc et al., 2016). Ectopic expression
of LSD1 in mOSNs delays terminal differentiation and induces OR gene switching (Lyons et al., 2013). Therefore, down-regulation of LSD1 is required for the stability of OR choice, at least in mature OSNs (reviewed in Monahan and Lomvardas, 2015). A study from our laboratory suggests that LSD1 is compartmentalized in a cell-cycle dependent manner, and this restricted spatial and temporal organization in the nucleus could contribute to restricted LSD1 activity, for example, restrict LSD1 activity to a single OR locus at a time (Kilinc et al., 2016). Recently, RNA-seq studies suggest that multiple ORs initially express per neuron, requiring a later refinement process to ensure singularity (Hanchate et al., 2015; Tan et al., 2013, 2015). Therefore, we expect that LSD1 is not functioning in an absolute manner; for example, LSD1 alone serving a “gate-keeping” function as proposed in the Lomvardas model (Figure 14).

LSD1 has been characterized as both a putative activator and repressor of target gene activity depending on the context and associated complexes (Shi et al., 2004; Metzger et al., 2005; Wang et al., 2007). Interestingly, LSD1 is required for the stability of DNMT1 and maintenance of global DNA methylation (Clements et al., 2012), suggesting that the LSD1 pathway is linked to long-term epigenetic memory. In this chapter, I investigated the function of LSD1 in the context of pre-committed immature OSNs. Towards this goal, I knocked down LSD1 using a lentivirus-based RNAi approach in the OP6 cell line. I demonstrated that the process of monogenic-monoallelic OR gene expression is partially dependent on the LSD1 protein. Using RNA FISH, I showed that the depletion of
LSD1 in OP6 cells resulted in multigenic and multiallelic OR gene expression, a phenotype never observed in wild-type OP6 cells. Using degenerate PCR and sequencing, I show that LSD1-depleted, clonally expanded OP6 colonies are able to activate ORs de-novo (i.e., they switch ORs normally). Using chromatin immunoprecipitation (ChIP), I show that the prolonged culturing of the LSD1-depleted OP6 cells exhibit gradual and systematic increases in H3K4 methylation at OR loci, consistent with the hypothesis that LSD1 regulates H3K4me methylation levels, presumably as a mechanism to ensure efficient silencing of competing ORs. In conclusion, my work suggests that LSD1 is involved in ensuring OR singularity by silencing competing ORs in the absence of feedback inhibition within the context of OP6 cells.

The following is from my publication of these results in Molecular and Cellular Neuroscience (Vyas et al., 2017).

2.1 Introduction

The proper function of the mouse olfactory system depends on the development of specialized sensory neurons that each respond to a narrow range of odorant chemistry. This specialization is accomplished by the expression of one allele (“monoallelic”) of one odorant receptor (OR) gene (“monogenic”) from a repertoire of ~1400 OR genes encoded in the mouse genome (Buck and Axel, 1991; reviewed in McClintock, 2010; Rodriguez, 2013). A very active area of
research and the focus of this study concerns the mechanism whereby each olfactory sensory neuron (OSN) transcribes OR genes/alleles in a mutually exclusive manner.

Recent evidence points to epigenetic mechanisms contributing to mutually exclusive OR transcription (reviewed in Monahan and Lomvardas, 2015). Three compelling lines of evidence are summarized here. First, identical OR transgene cassettes incorporated into the same mouse genome are expressed in a mutually exclusive manner (Serizawa et al., 2000, 2003), indicating that this phenomenon must be governed at a level beyond the DNA sequence per se. Second, disruption of epigenetic states, including nuclear chromocenter structure (Clowney et al., 2012) or deletion of proteins that establish heterochromatic marks (Lyons et al., 2014), results in multiple OR genes transcribing per cell. And third, two heterochromatic histone marks, H3K9me3 and H4K20me3, are deposited at all OR gene loci prior to choice, and are subsequently removed exclusively on the selected OR gene allele, and a histone mark associated with activation, H3K4me3 is exclusively acquired on the selected OR gene allele (Magklara et al., 2011), suggesting that OR activation involves selective de-repression. The simplest model to account for these observations is one in which OR genes are initially sequestered to a silencing chromatin compartment, with one allele only being liberated from this heterochromatin so that it can access transcriptional machinery.
A prediction of this model is that there exists a de-repressive protein complex to which competing OR gene loci have restricted access. The lysine-specific demethylase-1 (LSD1) is a plausible candidate constituent of such a putative complex. LSD1 is a versatile protein utilized in a broad range of activating and repressing functions in various developmental contexts (Cai et al., 2014; Foster et al., 2010; Laurent et al., 2015; Maes et al., 2015; Mosammaparast and Shi, 2010; Ray et al., 2010; Rusconi et al., 2015; Shin et al., 2015; Su et al., 2009; Toffolo et al., 2014; Tsai et al., 2008; Wang et al., 2007). The protein functions in both the H3K4 and H3K9 histone demethylation pathways, the latter important for gene de-repression; the H3K9 demethylation observed exclusively at the selected OR gene locus (Magklara et al., 2011) suggests a possible role for LSD1 in OR activation. LSD1 is expressed in the early cells of the OSN lineage, but it is down-regulated later in the lineage after OR choice has occurred (Kilinc et al., 2016; Krolewski et al., 2013), and the deletion of mouse LSD1 in vivo results in a dramatic decrease in OR expression within the mature olfactory epithelium (Lyons et al., 2013). Finally, we recently reported that mouse LSD1 protein is compartmentalized in the nucleus during a narrow developmental window within the earliest post-mitotic cells of the OSN lineage when OR selections are thought to be occurring; these LSD1 compartments appear to interact with one or a small number of OR loci at a time, consistent with possibly playing a role in mutually exclusive selection (Kilinc et al., 2016). Together, these findings provide momentum for this study, where we directly investigate the
consequence of LSD1 depletion in OR regulation within an immortalized olfactory-placode derived cell line (OP6).

The OP6 cell line is derived from an immortalized post-progenitor receptor neuron from E10 mouse olfactory placode (Illing et al., 2002). The onset of OR expression is thought to occur by E9.5 (Rodriguez-Gil et al., 2010), or before the embryonic state from which OP6 cells were derived (E10). Individual OP6 cells in culture appear to express OR genes monoallelically and monogenically (Kilinc et al., 2014; Pathak et al., 2009). However, full OSN maturation including up-regulation and stabilization of the selected OR gene, presumably via the unfolded protein response (UPR)-mediated feedback loop (Dalton et al., 2013), is not realized in OP6 cultures. Therefore, we suggest that OP6 cells might represent a stage between “OR choice” and “OR commitment”, and thus an opportunity to distinguish between these two regulatory events not easily separable in vivo. Consistent with this perspective, we observe that OR expression levels in OP6 cells are much lower than the levels found in more mature OSNs of the olfactory epithelium, thus probably below a threshold required to trigger the UPR feedback pathway; perhaps consequently, we also observe frequent OR switching during OP6 cell growth indicating a lack of post-feedback commitment (Pathak et al., 2009). Moreover, OP6 cells do not express the olfactory marker protein, OMP, which is up-regulated in mature OSNs after OR commitment; moreover, OP6 cells express both GAP43 and LSD1 (Kilinc et al., 2014; Pathak et al., 2009), which are down-regulated in mature OSNs after
OR commitment (Dalton et al., 2013; Lyons et al., 2013). The latter observation is noteworthy, since it has been previously argued that down-regulation of LSD1 might be necessary to stabilize OR selection by preventing additional OR activation events (Lyons et al., 2013). Thus, one of the motivations of our study was to investigate whether OR choice indeed becomes stabilized when LSD1 is artificially down-regulated as normally occurs post-feedback in vivo. We used RNAi to knockdown LSD1 expression to undetectable levels in a high percentage (~90%) of OP6 cells in culture. We find that LSD1 depletion does not appear to interfere with the ability of OP6 cells to switch from one OR gene to another during culturing, a result not predicted if LSD1 was required for de novo OR activation events. This result also seems consistent with results obtained in LSD1 knockout mice, where OR genes are able to express in this context as well, albeit at reduced levels (Lyons et al., 2013). Thus, neither study supports a model in which LSD1 activity is necessary for initial OR selection. Surprisingly however, we observe that LSD1 depletion in OP6 cells disrupts both monoallelic and monogenic OR transcription, suggesting that LSD1 is normally required to prevent multiple OR activation events per cell. This result suggests that LSD1 plays a role in suppressing competing OR genes/alleles.
2.2 Methods

2.2.A Mouse OP6 Cell Preparation and Immunofluorescence

The OP6 cell line was cultured at 33° C in Dulbecco’s Modified Eagle’s Medium (DMEM, Life Technologies) supplemented with 10% fetal bovine serum (Gibco), as described previously (Illing et al., 2002). In some studies, we generated clonal colonies of various sizes from single OP6 cells. For subsequent immunofluorescence and FISH analyses, cells were seeded on 22 cm² coverslips coated with 0.1% gelatin (Sigma) in a 6 well plate at about 50% confluency and expanded for one day to near confluency. Briefly, cells were fixed with 3% paraformaldehyde in PBS for 10 minutes, permeabilized in 0.5% Triton-X (Sigma) for 10 minutes, and blocked in 4xSSC, 0.20% Tween, 4.0% BSA for 20 minutes at 37° C. The primary and secondary antibody incubations were performed at 37° C for 45 minutes in a humidified chamber. The primary antibodies used in this study were rabbit anti-Lsd1 (Abcam, ab129195, 1:100), mouse anti-CoREST (Millipore, MABN486, 1:250), rabbit anti-LBR (Abcam, 122919, 1:50), and mouse anti-methylcytosine (Millipore, MABE345, 1:200). The secondary antibodies used in this study were donkey anti-mouse-Cy3 (Jackson Immunoresearch, 715-165-150, 1:100), donkey anti-rabbit Alexa 488 (Jackson Immunoresearch, 711-545-152, 1:100), and goat anti-rabbit-Cy3 (Millipore, AP132C, 1:800). Images were acquired using a Deltavision RT imaging system (Applied Precision) adapted to an Olympus (IX71) microscope equipped with
XYZ motorized stage. Each image was sectioned with 0.5 μm intervals to ensure complete coverage of the nucleus. Images were processed using Softworx (Applied Precision).

2.2.B LSD1 Knockdown by RNAi

We adopted an RNAi protocol with slight modifications described at (http://www.broadinstitute.org/rnai/public/resources/protocols, 2012).

Briefly, lentiviral constructs containing mouse LSD1 shRNA (RMM3981-201787935) and scramble (non-targeted) shRNA (RHS6848, GE-Dharmacon Life Science) were transiently transfected along with packaging plasmids pCMV-dR8.2 dvpr (Plasmid#8455, Addgene) and pCMV-VSV-G (Plasmid#8454, Addgene) (Sarbassov et al., 2005) into HEK-293T cells (ATCC) using standard calcium phosphate reagents according to manufacturer protocols (Invitrogen). Virus-containing supernatants were collected at 36 and 60 hours after transfection and concentrated by ultracentrifugation. Approximately $10^5$ host OP6 cells were subjected to spin-infection at RT for 90 minutes in the presence of 8 μg/ml polybrene (Stewart et al., 2003). Lentivirus-transduced OP6 cells (LSD1-knockdown and scramble controls) were selected using puromycin and clonally passaged.
2.2.C cDNA Analyses

Total RNA was extracted from transduced OP6 cells using RNeasy Mini kit (Qiagen), and cDNA generated using superscript reverse transcriptase per manufacturer’s protocols (Invitrogen). In some studies, we produced cDNA from single cells using the SMARTer PCR cDNA synthesis kit (Clontech). All cDNA experiments included negative control samples in which reverse transcriptase was omitted (“no-RT controls”). Quantitative PCR (qPCR) was conducted on an ABI 7300 real-time PCR machine (Applied Biosystems). qPCR primer sequences used in this study are available upon request. For colony studies, we seeded individual OP6 cells (confirmed by microscopy) and grew ~200-cell clones. cDNA produced from these colonies were subjected to degenerate OR PCR using the following primers (Clowney et al., 2012; Michaloski et al., 2006):

ATGGCITAYGAYMGITAYGTIGCIATHTG (135R), RTTICKIARISWRTAIATRAAIGGRTT (P8), GCITAYGAYCGITAYGTIGCIATITG (P26), and ACIACIGAIAGRTGIGAISCRCAIGT (P27). Restriction digests were conducted on degenerate PCR products using Msel, MluCl, or HinfI (New England Biolabs). Degenerate PCR products were cloned into pCR-TOPO4.0 (Invitrogen), and several clones per sample were selected for plasmid preparation and sequencing.
2.2.D Random Sampling Simulations to Estimate OR Complexity

We simulated random sampling from 200-cell colonies across various levels of OR complexity to identify a complexity level in which the average simulation yields the observed OR redundancy in that random sample. To exemplify, we obtained 41 distinct OR genes from a random sample of 60 sequenced templates from the Scr-12 colony; this degree of redundancy (41 distinct ORs, 19 redundant ORs in a random sample size of 60 sequences) is the average outcome in a simulation where the overall complexity in the product is 73 OR genes. The results of the simulations for each colony are shown in Table 1. In addition, we tested the null hypothesis that OR representation is not significantly different between KD and Scr colonies by simulating random sampling from a pair of colonies and estimating the expected extent of OR-gene overlap from the two samples assuming an equal representation (73 distinct genes all equally likely to be sampled). To exemplify, a random sampling of 41 ORs from Scr-12 and 19 ORs from KD1 produces an average of 10.2 overlapping ORs between the two samples in the simulation; the observed 9 overlapping ORs between these two samples is not significantly different than the simulation (p=~0.35 to observe 9 or fewer overlaps). The results of these simulations for pairs of colonies are shown in Table 2.
2.2.E Western Blot Analyses

Total protein was extracted from \( \sim 10^6 \) OP6 cells using 1X RIPA buffer (Cell-Signaling) containing 1X protease inhibitor cocktail (Roche). Protein levels were measured using the Pierce BCA Protein assay kit (Thermo Scientific). Western blot experiments were conducted on 8-10% SDS-PAGE gels transferred to polyvinylidene membranes (Thermo Scientific), blocked in 5% BSA in TBST, followed by overnight incubation with primary antibodies (see below). Following washes, membranes were incubated for 1 hour using anti-rabbit AP (Abcam), washed, and antigen visualized using NBT/BCIP reagent (Roche). Blots were imaged using G-Box F3 (Syngene) and densitometry analyses conducted using ImageJ software (NIH). The following primary antibodies were used for Western blots: anti-LSD1 (ab129195), anti-Beta-actin (ab8227), and anti-acetylated histone H4 (06-866, Millipore). For the global histone modification analysis, we followed the protocol described in (Rambaugh and Miller, 2011).

2.2.F Chromatin Immunoprecipitation (ChIP).

Native chromatin was prepared from \( \sim 10^7 \) OP6 cells as described previously (Brand et al., 2008). Briefly, nuclei were extracted, digested with micrococcal nuclease (Thermo Scientific) and purified using micro-spin hydroxyapatite chromatography columns (The Nest Group) to yield mono-nucleosomes. Purified chromatin was incubated with pre-mixed Protein-A Dynabeads (Invitrogen) plus either anti-H3K4me3 (04-745, Millipore), anti-H3K4me2 (07-030, Millipore),
anti-H3K9me3 (ab8898, Abcam) or anti-H3K9me2 (ab1220, Abcam).

Immunoprecipitated DNA was isolated by phenol-chloroform extraction and ethanol precipitation. Enrichment relative to input fractions was determined using qPCR; primer sequences for various control and OR genes used in various ChIP experiments are available upon request. All enrichment levels were noise-subtracted using a negative control presumed to not contain any histones with the modification in question, and normalized using a positive control presumed to not vary between samples in the histone modification in question.

2.2.G  RNA FISH

Gene-specific probe templates were generated by PCR (primer sequences available upon request) and cloned into pCRII-TOPO vector (Invitrogen) containing flanking SP6 and T7 promoters to enable in vitro production of sense and antisense probes. All cloned products and their orientations within the vector were verified by sequencing. Following linearization, in vitro transcription was carried out by either SP6 or T7 (Roche) RNA polymerases using a digoxigenin (DIG) RNA labeling mix (Sigma; 11277073910) or biotin RNA labeling mix (Sigma; 11685597910). RNA FISH was performed as described previously (Kilinc et al., 2014; Schaeren-Wiemers and Gerfin-Moser, 1993). Briefly, cells were permeabilized in CSK buffer with 2 mM vanadyl ribonucleoside complex (NEB) and 0.5% Triton-X100 (Sigma), then fixed in 4% paraformaldehyde, followed by dehydration in a 70%-80%-95%-100% ethanol
series prior to incubation in 50% formamide/2X-SSC. Approximately 50-100 ng of probe was added to the hybridization buffer (50% formamide, 2X-SSC, 10% dextran sulfate, 5 μg Cot1-DNA, 500 μg/ml salmon sperm DNA) and incubated overnight at 37°C in a humidified chamber. Following washes (maximum stringency = 50% formamide, 0.5X-SSC at 37°C), samples were blocked for subsequent antibody incubations (4% BSA, 4X-SSC, 0.2% Tween-20). DIG signals were visualized using sheep anti-DIG FITC (11207741910, Roche) and donkey anti-sheep FITC (sc-2476, Santa Cruz Biotech) antibodies. For RNA FISH experiments conducted with biotinylated probes, we amplified signals by including a third antibody incubation as follows: primary incubation using avidin-rhodamine (A-2012, Vector Laboratories), secondary incubation using biotinylated anti-avidin (ab7235, Abcam), and a third incubation using avidin-rhodamine (A-2012, Vector Laboratories). For two-color RNA FISH experiments, we combined primary antibodies and secondary antibodies in respective incubations.

2.3 Results

We have utilized a cell line derived from E10 mouse olfactory placode, the OP6 cell line, to investigate the molecular mechanisms underlying mutually exclusive OR transcription. The founder cell to the OP6 line has been staged as a post-mitotic immature neuronal cell (Illing et al., 2002). We have previously shown that OP6 cells express OR genes in a mutually exclusive manner (both
monallelically and monogenically), although at much lower transcript level per cell than is the case in mature olfactory sensory neurons (OSNs) (Kilinc et al., 2014). We have also shown that OP6 cells "switch" their OR choice during culturing (Pathak et al., 2009). Together, the reduced expression level and the lack of stability in OR choice suggest that OP6 cells represent an intermediate state between OR selection and commitment.

OR commitment is thought to involve a feedback system designed to prevent the activation of a second gene once an initial selection has been successfully realized, and presumably to amplify the expression level of the selected OR so the neuron has robust odorant-binding function. The selected OR protein generates this feedback via activating the unfolded protein response (UPR) pathway, with one downstream consequence being the down-regulation of the LSD1 protein (Dalton et al., 2013), which might prevent the cell making an additional OR selection (Lyons et al., 2013). In the absence of OR protein and this feedback process, the cell will "switch" to another OR (Lewcock and Reed, 2004; Serizawa et al., 2003; Shykind et al., 2004). The fact that OP6 cells “switch”, as well as express LSD1 at pre-feedback levels, is further indication that OP6 cells represent an intermediate state between OR selection and commitment.

2.3.A LSD1 Knockdown in OP6 Cells.

A simple model has been proposed that links the down-regulation of LSD1 to the commitment of the neuron to its initial OR choice, in which OR selection is dependent on LSD1-mediated H3K9 demethylation activity (Lyons et al., 2013;
Tan et al., 2013; Tian et al., 2016). Thus, this model predicts that the down-regulation of LSD1 would be sufficient to prevent additional OR activation events. A problem with this model is that LSD1 knockout mice appear to make OR selections, albeit at reduced expression levels (Lyons et al., 2013), suggesting that LSD1-mediated H3K9-demethylation might instead be important with OR enhancement or some other downstream function, as opposed to being required for initial OR selection per se.

To further investigate a role for LSD1 in OR selection and commitment pathways, we used RNAi to deplete LSD1 from OP6 cell cultures. Following selection of transfected cells, resulting OP6 populations exhibit stable ~90% knockdown of LSD1 mRNA (through Day 21), as compared to control OP6 populations that were transfected with a “scrambled” cassette that should not target any known gene in the mouse genome (Fig. 1A). We also analyzed LSD1 protein immunofluorescence levels and Western blot levels in knockdown (KD) versus scrambled control (Scr) OP6 cell populations, and observe robust depletion of LSD1 in KD populations as compared to Scr populations (Fig. 1B). The immunofluorescence experiments indicate that the majority of KD cells (51/59) exhibit no detectable LSD1 protein, whereas a small subset of these cells (8/59) exhibit approximately the same level of LSD1 immunofluorescence as the control populations (Fig. 1B). This observation seems to indicate that the residual ~10% LSD1 mRNA measured in KD populations (~90% knockdown level, Fig. 1A) is probably due to a bimodal population (“all or nothing”) as
opposed to a homogeneous partial knockdown. We also observe several other interesting phenotypes in the LSD1-depleted cells. These preliminary observations will be discussed further in Chapter 4 (future direction).
2.3.6 LSD1 Depletion does not Prevent OR Switching during OP6 Clonal Expansion.

Our specific interest was to investigate the impact of LSD1-depletion on OR gene regulation. A prediction of a model in which LSD1-mediated demethylation is required for de novo OR activations during OP6 cell culturing (i.e., “switching”), is that no new ORs would get selected during expansion of an OP6 cell clone. We tested this prediction by growing ~200-cell clones from isolated single OP6 cells. We conducted RT-PCR on LSD1-depleted and control clones using degenerate PCR primers designed against conserved OR sequences. The specific degenerate
PCR primers utilized have been shown previously to amplify a large fraction of OR genes from a complex template mixture (Clowney et al., 2012; Michaloski et al., 2006). We then digested these degenerate PCR products using a variety of frequent-cutting (4-base) restriction enzymes, where the complexity of the digested fragments will be proportional to the complexity of the original PCR product. For example, we observe only two bands when digesting a gene-specific OR PCR product (whose fragment sizes exactly add up to the size of the original PCR product), whereas we observe hundreds of bands (whose fragments sizes add up to much more than the size of the original PCR product) when digesting a degenerate RT-PCR product from a large OP6 culture expressing numerous OR genes (Fig. 2). When applying this assay on LSD1-depleted and control cell clones, we observe no obvious reduction in the complexity of KD samples, as might be predicted if these clones were unable to switch and activate new ORs during colony expansion (Fig. 2, Table 1). We also sequenced several OR templates generated from LSD1 knockdown colonies, as well as from scrambled control colonies to confirm that multiple OR genes are expressed in all colonies (Table 1). Consistent with the OR heterogeneity observed in restriction digest patterns shown in Figure 2, we identified multiple OR genes present in degenerate PCR templates derived from KD samples (15, 19, and 24 distinct OR genes from the KD7, KD1, and KD6 colonies, respectively; Table 1). Assuming that our sequenced sample is unbiased and representative (see Methods), we used the degree of redundancy observed at various depths of sequencing of
these products to arrive at an approximation of the total OR complexity represented in the Scr and KD degenerate PCR products (Table 1). While these estimates are likely quite error-prone, given that the sample is probably not random due to gene-specific bias in the amplicon (see Methods), we note that these estimates (~76 and ~65 OR genes, respectively) are similar to previous estimates of the overall OR complexity within a typical OP6 culture as measured by cDNA microarrays (~80 OR genes; not shown). Together, our results suggest that LSD1 depletion does not dramatically impair the ability to select new OR genes during colony expansion: KD clonal colonies appear to exhibit a comparably rich array of OR genes. While “OR switching” per se is evident in KD populations (OR complexity is >1 in KD clones), we remain cautious about possible subtle differences in switching tendencies that might arise as a consequence of LSD1 absence.
Figure 2. Colony switching assays indicate comparable levels of OR complexity in knockdown versus control colonies.

Upper Panel: Degenerate OR RT-PCR products (using P135-P8 degenerate primers) from wild-type OP6 whole populations (WP), two scrambled control clonal colonies (Scr, 200 cell colonies each), and two LSD1-knockdown clonal colonies (KD, 200 cell colonies each) are shown undigested (u, 531-bp product) and digested with Mse1 restriction enzyme (M). A single OR gene product (OR31, 320 bp) is digested with Mse1 to control for complete digestion and false-positive complexity.

Lower Panel: Additional degenerate OR RT-PCR products from two knockdown (KD) and two scramble control (Scr) colonies each digested with three restriction enzymes (Ml=MluCI, Ms=MseI, H=HinfI).
2.3.C  LSD1 Depletion does not seem to Alter OR Representation in OP6 Cell Populations.

We were next curious to know whether LSD1 depletion impacted the representation of OR genes commonly and reliably represented in OP6 cultures. We used RNA FISH to determine the percentage of cells in LSD1 knockdown and control cultures for four specific OR genes (Olfr 58, Olfr 70, Olfr860, and Olfr1414) that are commonly represented in OP6 populations, as well as with several BAC-sized RNA FISH probes. Anecdotally, we were unable to identify any gains or losses in OR representation in LSD1-depleted populations using this non-exhaustive method. We also analyzed the OR sequences obtained from degenerate OR RT-PCR from LSD1-depleted and control populations. While the sampling is again non-exhaustive, we do observe numerous common ORs represented in both the KD and Scr degenerate PCR products. We simulated a random sampling from each of the clones to predict an expected degree of OR overlap, assuming both KD and Scr populations contained the same 76 OR genes (estimated complexity from Table 1), and found that the observed OR overlap was within statistical probability for all pairwise overlaps (except the Scr12-KD6 overlap, which was more divergent than expected; Table 2). That is, the observed overlap in OR representation between Scr and KD sequence samples is not significantly less than the null hypothesis might predict (i.e., if there was no change in OR representation). Therefore, the sampling methods utilized thus far (degenerate RT-PCR and RNA FISH) seem to indicate that LSD1 depletion does
not dramatically influence OR representation in OP6 cell populations. A more rigorous analysis of OR representation/switching tendencies in KD versus Scr OP6 cultures would depend on an exhaustive RNA-seq approach.

Table 1. Degenerate-OR PCR Product Complexity in Knockdown Versus Control Colonies.

The number of sequenced OR templates from degenerate RT-PCR products for two scrambled control colonies (Scr-11, Scr-12) and three LSD1 knockdown colonies (KD-7, KD-1, KD-6) is shown, as well as the number of distinct OR genes identified in the sequenced sample. The degree of redundancy observed in the sequencing sample was used to estimate the total OR complexity in the degenerate OR PCR product (right column; see Methods). Similarly, the total number of distinct ORs identified across both scrambled control (Scr-total) and all three knockdown (KD-total) was used to estimate the total complexity in these respective colonies.

<table>
<thead>
<tr>
<th>Clone</th>
<th>Sequenced</th>
<th>Distinct ORs</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scr-11</td>
<td>13</td>
<td>11</td>
<td>35</td>
</tr>
<tr>
<td>Scr-12</td>
<td>60</td>
<td>41</td>
<td>73</td>
</tr>
<tr>
<td>Scr (total)</td>
<td>73</td>
<td>47</td>
<td>76</td>
</tr>
<tr>
<td>KD-7</td>
<td>17</td>
<td>15</td>
<td>63</td>
</tr>
<tr>
<td>KD-1</td>
<td>26</td>
<td>19</td>
<td>38</td>
</tr>
<tr>
<td>KD-6</td>
<td>40</td>
<td>24</td>
<td>35</td>
</tr>
<tr>
<td>KD (total)</td>
<td>83</td>
<td>47</td>
<td>65</td>
</tr>
</tbody>
</table>
Table 2. Degenerate-OR PCR Product Overlap between Knockdown and Control Colonies.

The number of distinct ORs observed in various scramble control (Scr) and knockdown (KD) clones, and the expected (e) number of overlapping OR genes in the pairwise comparison assuming a simulated random selection from a total complexity of 76 ORs estimated in degenerate RT-PCR products (see Table 1). The observed number of overlapping OR genes (o) is not significantly different than predicted by the hypothesis (p-values from simulation shown, see Methods), with the exception of the Scr12-KD6 pair, where there are significantly fewer overlaps observed than predicted (p=0.04).

<table>
<thead>
<tr>
<th>Pair</th>
<th>ORs</th>
<th>e</th>
<th>o</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scr11-Scr12</td>
<td>11,41</td>
<td>5.9</td>
<td>5</td>
<td>0.42</td>
</tr>
<tr>
<td>KD1-KD6</td>
<td>19,24</td>
<td>6.1</td>
<td>5</td>
<td>0.43</td>
</tr>
<tr>
<td>KD1-KD7</td>
<td>19,15</td>
<td>3.7</td>
<td>4</td>
<td>0.68</td>
</tr>
<tr>
<td>KD6-KD7</td>
<td>24,15</td>
<td>4.8</td>
<td>5</td>
<td>0.70</td>
</tr>
<tr>
<td>Scr11-KD1</td>
<td>11,19</td>
<td>2.8</td>
<td>2</td>
<td>0.45</td>
</tr>
<tr>
<td>Scr11-KD6</td>
<td>11,24</td>
<td>3.5</td>
<td>2</td>
<td>0.26</td>
</tr>
<tr>
<td>Scr11-KD7</td>
<td>11,15</td>
<td>2.2</td>
<td>2</td>
<td>0.60</td>
</tr>
<tr>
<td>Scr12-KD1</td>
<td>41,19</td>
<td>10.2</td>
<td>9</td>
<td>0.35</td>
</tr>
<tr>
<td>Scr12-KD6</td>
<td>41,24</td>
<td>13.0</td>
<td>9</td>
<td>0.04</td>
</tr>
<tr>
<td>Scr12-KD7</td>
<td>41,15</td>
<td>8.1</td>
<td>8</td>
<td>0.57</td>
</tr>
</tbody>
</table>

2.3.D LSD1 Depletion Disrupts Monoallelic and Monogenic OR Expression in OP6 Cells.

In our RNA FISH experiments conducted on the LSD1-depleted OP6 populations at passage-8 (35 days post-infection), we were surprised to observe that all four OR genes exhibited an approximately 2-fold higher percentage of positive cells in knockdown cultures (Fig. 3). Even more surprisingly, we observed several of the positive cells for a given probe exhibited two positive signals per nucleus. In
numerous RNA FISH experiments over the previous several years, we have never before observed a multi-spot signal in wild-type OP6 cells. Therefore, LSD1 depletion in OP6 cells appears to disrupt allelic exclusion: while 100% of OP6 cells exhibit one-spot RNA FISH signals consistent with monoallelic expression, an average >10% of the positive cells for a given OR probe in knockdown populations exhibit two-spot RNA FISH signals consistent with a modest incidence of biallelic expression. We note that sense control probes were cleanly negative in all experiments for, n=200 cells (not shown).

FISH data using a single gene probe, demonstrating a disruption of allelic exclusion, led us to wonder if LSD1-depleted cells might also express multiple genes per cell. Such an outcome would predict a systematic increase in the percentage of positive cells for each representative OR; for example, a systematic ~2-fold increase in positive cells per OR tested in knockdown versus control populations (as evident in Fig. 3) would be consistent with each cell now expressing an average of two ORs per cell, assuming the overall OR representation in the population is unchanged. To seek evidence for multigenic OR expression in LSD1-depleted OP6 cells, we conducted RNA FISH experiments using pooled OR probes (Fig. 3). We make three observations from these experiments. First, we observe that RNA FISH experiments conducted with pools of probes suffer from a substantial increase in the false-negative rate as compared to experiments conducted with single probes. We conclude this because the percentage of positive cells in the pool of three probes is less than
the sum of percentages observed with each of the probes when tested individually. In one pooled experiment, we estimate that the additional false-negative rate due to pooling is ~38% (Fig. 3B, top histogram); in a second pooled experiment, we estimate that the additional false-negative rate due to pooling is ~56% (Fig. 3B, bottom histogram). Second, we observe an even higher percentage of multi-spot nuclei than is observed in experiments conducted with individual probes. That is, in the first pooled experiment, we observe 3.5% multi-spot nuclei in the pool, whereas the incidence of multi-spot nuclei in the three individual experiments would only predict a total of 2% incidence in the pool (Fig. 3A); similarly, in the second pooled experiment, we observe 1.7% multi-spot nuclei in the pool, whereas the incidence of multi-spot nuclei in the three individual experiments would only predict a total of 1% incidence in the pool (Fig. 3B). While perhaps within variation expected by chance alone, this increase in multi-spot incidence in the pool as compared to the sum of the parts suggests some incidence of multigenic OR expression above and beyond the incidence of multiallelic OR expression. And third, we observe an incidence of some nuclei with three or more RNA FISH signals when pooling probes that was not observed in any of the experiments conducted with individual probes. This observation also suggests some incidence of multigenic OR expression above and beyond the incidence of multi-allelic OR expression.
Figure 3. RNA FISH with Individual and Pooled OR Probes.
A. Representative RNA FISH signals (green) in scrambled control (Scr; solid bars) and knockdown (KD; open bars) cell nuclei for three OR antisense probes (58, 70, 860), as well as a pool of all three probes. Rare multi-spot RNA FISH signals are observed only in knockdown populations, including some cells in the pooled probe experiment that exhibit more than two spots. Sense control probes were all negative (not shown). B. Histograms showing the percentage of cell nuclei (n=number of cell nuclei scored for each probe) exhibiting a positive RNA FISH signal for each individual probe, as well as the pool, in two independent experiments. Hatched bars indicate the percentage of these positive cells that exhibit a multi-spot signal. Cell nuclei are labeled with DAPI.

To more rigorously distinguish between multiallelic and multigenic OR expression in LSD1-depleted OP6 cells, we conducted two-color RNA FISH
experiments with the Olfr288, Olfr860, and Olfr58 probes (Fig. 4). As observed in the previous one-color experiments reported in Figure 3, there is an approximate doubling of positive cells in KD versus Scr populations. The average percentage of positive cells in the three experiments conducted with individual probes (Olfr288 only, Olfr860 only, and Olfr58 only) increased from 2.2% (Scr) to 4.7% (KD), and the average percentage of positive cells in the two experiments conducted with pairs of probes (Olfr288+Olfr860 and Olfr288+Olfr58) increased from 3.7% (Scr) to 8.0% (KD) (Fig. 4B). As reported in Figure 3, we also observed an incidence of multi-allelic expression in only KD samples for these single-probe experiments (3 of a total of 21 positive cells exhibited a two-spot signal) (Fig. 4). In the two-color experiments conducted with pooled pairs of OR probes, we observed examples of multiallelic OR expression (e.g., top left panel in Fig. 4A) and multigenic OR expression (top right panel in Fig. 4A). We also observed an example of a 3-spot signal (in 24 total positive cells) that shows both multiallelic and multigenic OR expression in a single cell (lower left panel in Fig. 4A). Overall, the incidence of multiallelic cells was equal to the incidence of multigenic cells in these experiments (~1.7% incidence each; Fig. 4B), suggesting that the probability of selecting the other allele is approximately the same as the probability of selecting any other random gene. Of the 7 total cells exhibiting multi-colored RNA FISH signals (i.e., with at least one red and one green spot), we find one example of a yellow spot that would suggest a shared RNA polymerase (shown in the lower right panel in Fig.
4A); the other 6 cells are exemplified by the image shown in the upper right panel in Figure 4A, where the red and green spot are clearly separated in space suggesting transcription from a distinct RNA polymerase location in the nucleus. This finding seems to dismiss the notion of a dedicated OR-specific RNA polymerase factory for which ORs might compete as part of a mechanism to ensure mutual exclusivity, at least during this initial selection (pre-feedback) phase. Of note, one gene pair used in this experiment (Olfr860 and Olfr58) are ~70kb apart within the same OR cluster on chromosome 9, whereas the other gene pair (Olfr860 and Olfr288) are from different OR clusters on chromosome 9 and 15, respectively. We observed a roughly equal tendency for multigenic expression in the two experiments (~22% of positive cells co-express the gene pair from the same chromosomal cluster; ~20% of positive cells co-express the two genes from two different chromosomal clusters). The one observed yellow spot is from the OR pair from the same chromosomal locus (Olfr860 and Olfr58; lower right panel in Fig. 4A), so this co-localized transcriptional event may very well arise from cis co-activation as opposed to two different chromosomes looping into the same RNA polymerase factory. At this point, we do not have evidence with our very limited sampling to suggest gene proximity bias for OR co-expression in LSD1 depleted OP6 cells.
Figure 4. Two-color RNA FISH on LSD1-depleted Cells Using Pairs of OR Probes.

A. Representative two-color RNA FISH images conducted with a pair of OR probes (Olfr860=red, Olfr58=green). **Upper left:** an example of a multiallelic cell showing two positive RNA FISH signals from the same probe/color. **Upper right:** an example of a multigenic cell showing two positive RNA FISH signals, one from each OR probe/color. **Lower left:** an example of a cell with three positive RNA FISH signals, with both Olfr860 alleles expressing (red) along with the Olfr58 gene (green). **Lower right:** an example of a cell with Olfr860 (red) and Olfr58 (green) expressing from the same physical location in the nucleus, evident by a yellow spot (inset panels: same cell shown in the individual red and green channels).

B. Histogram summarizing results from three experiments conducted with each of the three OR probes individually (left) and the two experiments conducted with pooled pairs of OR probes (right). The incidences of monoallelic/monogenic OR expression (open portions of bars), multiallelic OR expression (gray portions of bars), and multigenic OR expression (black portion of bars) are shown for LSD1-depleted cells (KD) and scrambled controls (SCR). All sense probe controls were negative (not shown). Cell nuclei are labeled with DAPI.
2.3.E  H3K4 and H3K9 Methylation States at OR Loci in LSD1 Depleted cells.

We conducted chromatin immunoprecipitation (ChIP) experiments using antibodies against H3K4me2, H3K4me3, H3K9me2, and H3K9me3, in order to investigate the impact of LSD1 depletion for these marks at various OR promoter regions (Fig. 5). The H3K4me2/me3 levels at OR loci were predictably near background levels as established by the zfp560 negative control; the maximum enrichment level observed was for Olfr944, where the noise-adjusted H3K4me3 signal strength was ~7% that observed for GAPDH in KD populations (Fig. 5A, lower panel). Therefore, our general conclusion from H3K4 ChIP is that most OR loci continue to exhibit a paucity of this mark, as would be predicted if H3K4 is only rarely acquired at OR loci (e.g., most of the templates for a particular OR gene are not transcribed in the population at a given point in time).

To more rigorously investigate the long-term impact of LSD1 depletion at OR loci, we compared H3K4me2 levels in late- versus early-passage LSD1-depleted cell populations. We observe a systematic increase in H3K4me2 levels at higher passage numbers in LSD1-depleted cultures (Fig. 5C). The H3K4me2 enrichment levels even after prolonged absence of LSD1 for several generations remain very near background levels. Since we are measuring values close to zero in all cases, we presume that even small amounts of experimental error/variation might confound interpretation. Nevertheless, we observe a highly reproducible trend in which all 21 OR promoters tested (across four
independent ChIPs for two independent chromatin preps) exhibit a small, but statistically significant elevation in this mark (pairwise T-test, p-value $\sim 10^{-5}$). We make two arguments that this systematic trend is probably not an experimental artifact (e.g., due to systematic differences in the chromatin preparations or ChIP efficiencies). First, as noted, this trend is reproducible across independent preparations, and so it seems unlikely that a systematic bias within the experiment would be reproduced in multiple attempts. Second, we observe the opposite directionality (i.e., marginally higher H3K4me2 levels in the early- versus late-passage) in three of the four experimental replicas for the positive control GAPDH gene assumed to maintain consistently high, LSD1-independent levels of H3K4me2. We also observe the opposite trend for several developmental gene promoters (MyoD, Gap43, G-olf, Ngn1, and AC3) tested in one of the experiments (not shown); these gene promoters, like OR promoters, exhibit low levels of H3K4me2, but probably are not regulated by LSD1 and therefore might be expected to retain constant H3K4me2 levels between KD and Scr populations. Together, our data suggest that LSD1 depletion has a gradual and subtle impact on H3K4 methylation in a systematic way across all tested OR loci. Importantly, the gradual yet systematic increase in H3K4 methylation at OR loci is predicted if LSD1 normally functions to ensure these active histone marks are only rarely acquired at OR gene promoters.

As anticipated, the H3K9 methylation levels at OR gene promoters are uniformly high and approximate or exceed those observed at the zfp560 positive
control (Fig. 5B). These data therefore recapitulate previous reports indicating that the H3K9 methylation is strongly associated with OR promoters (Magklara et al., 2011), presumably as part of a mechanism to ensure the silencing of most ORs per cell. We do not observe a significant or systematic difference in H3K9me3 at OR loci in LSD1-depleted cells (Fig. 5B, lower panel). In contrast, we do observe a slight but apparently systematic increase in H3K9me2 levels (Fig. 5B, top panel; paired T-test \( p<0.02 \)), albeit across a small experimental sample. An accumulation of H3K9me2 at activated OR loci might be predicted if the first LSD1-independent step in the H3K9me3 demethylation pathway at a selected OR allele is transpiring normally (H3K9me3 to H3K9me2) but the second LSD1-dependent step in this pathway (H3K9me2 to H3K9me1/0) is perturbed.

In summary, we observe an apparent systematic accumulation of H3K4me2 (and possibly H3K9me2) in LSD1-depleted cell populations, consistent with the interpretation that LSD1 normally functions to remove these marks at OR gene promoters. These increases are modest on a gene-by-gene basis across the cell population; in addition, we observe no obvious trend with H3K9me3, a mark previously shown to be associated with OR silencing. We emphasize that these ChIP experiments were conducted on whole populations, whereas we anticipate that the regulation of H3K9 (and H3K4) methylation is likely occurring at a given OR locus within a very small fraction (<5%) of the population; i.e., in only those rare cells where a given OR gene has been activated. Therefore, even if the number of expressing cells for a given OR gene
has doubled (e.g., from ~1% to ~2% of the population), the anticipated decrease in H3K9 methylation from ~99% to ~98% of templates in the population (or the anticipated increase in H3K4 methylation from ~1% to ~2% of templates in the population) is likely to be within the margin of experimental error and noise parameters for this methodology. A more informative investigation of LSD1’s role in regulating H3K4 and H3K9 methylation at OR loci would depend on a single-cell ChIP methodology and/or application on more homogeneously expressing cell populations.
**Figure 5. Chromatin Immunoprecipitation (ChIP) to Measure Changes in H3K4 and H3K9 Levels at OR Loci in LSD1-depleted OP6 Cells.**

**A. and B.** Enrichment levels in all chromatin ChIP experiments are relative to the input fraction, noise-subtracted using a negative control locus presumed to lack the mark, and normalized to a positive control locus to correct for experimental variation (see text). **A.** Relative H3K4me2 (top) and H3K4me3 (bottom) enrichment levels in LSD1-knockdown OP6 cell populations (open bars) and scrambled control OP6 cell populations (solid bars) for various odorant receptor promoter regions ("Olfr" numbers indicated) using Zfp560 for noise subtraction (negative control) and GAPDH for normalization (positive control). Only Olfr944 and Olfr995 seem to exhibit above-threshold signal for these histone marks. **B.** Relative H3K9me2 (top) and H3K9me3 (bottom) enrichment levels in LSD1-knockdown OP6 cell populations (open bars) and scrambled control OP6 cell populations (solid bars) for various odorant receptor promoter regions ("Olfr" numbers indicated) using GAPDH for noise subtraction (negative control) and Zfp560 for normalization (positive control). Two independent H3K9me3 experiments were conducted due to higher noise levels (experiment#1: 350, 944, 995, 385, 1497; experiment#2: 350, 944, 995, 389, 860, 58 as noted above histograms). **C.** Relative H3K4me2 enrichment levels between early passages (solid bars) and late passages (open bars) of LSD1-depleted OP6 cell populations at 21 OR gene promoters. All levels are normalized to input fraction to correct for PCR bias and noise-subtracted using a satellite DNA negative control to establish background thresholds. Standard deviation bars are shown for data collected from two independent chromatin preparations used to produce four independent ChIP experiments (except for Olfr389, Olfr177, and Olfr39 where only a single experiment is shown for each).

### 2.4 Discussion

We have depleted the LSD1 protein in a cell line clonally derived from olfactory placode founders that were previously staged as immature neurons in the developing olfactory lineage (Illing et al., 2002). The “immature neurons” of the olfactory lineage remain a subject of active research and characterization. In (Hanchate et al., 2015), these cell populations were subdivided into “early” and
“late” immature stages based on developmental gene profiles and the probability of multiple transcribed ORs per cell, with the latter subpopulation being enriched for cells expressing a dominant OR (like OP6 cells). Both subpopulations exhibit OR expression levels that are significantly lower than in mature OSN populations (like OP6 cells). We previously characterized the OP6 cell line as a likely intermediate stage along the OR activation pathway, having already selected one and only one OR gene and allele to be transcribed per cell at premature levels of transcript, yet uncommitted to this choice, likely due to the non-execution of the feedback inhibition process (Pathak et al., 2009) resulting in persistence of the LSD1 protein (Kilinc et al., 2016) and frequent switching of OR genes during culturing (Pathak et al., 2009). Until the olfactory sensory neuronal lineage is more precisely described, we are reluctant to map the OP6 founder cell as to an exact developmental stage within the OE. Moreover, we note that much of the characterization of the olfactory neuronal lineage, and especially attributes of OR gene regulation within this lineage, stem from studies involving late embryonic, post-natal or adult mice, whereas little is known about OSN staging, OR expression profiles, and mechanistic details of OR regulation within an early embryonic developmental context, including the nascent E10 olfactory placode. Nevertheless, we propose that the OP6 cell line was derived from a developmental stage when OR genes are first expressing at low levels, yet incapable of the UPR-mediated feedback loop associated with OSN maturation, as well as robust OR expression and commitment.
We have worked with the OP6 cell line for more than a decade, and have never observed multigenic or multiallelic OR expression using a variety of methods. Herein we show that depletion of the LSD1 protein in OP6 cells disrupts both monoallelic and monogenic OR expression in this cell line. This result is consistent with LSD1 functioning in a “repressive” pathway otherwise important for maintaining the silencing of competing OR alleles/genes, which contrasts previous published studies that focus on possible “de-repressive” functions for LSD1.

We speculate on two possible “repressive” mechanisms in the context of maintaining the silencing of competing OR alleles/genes. First, LSD1 might be important not just for H3K9 regulation at OR loci, but also H3K4 regulation prior to the UPR-mediated OR commitment and further differentiation into a mature OSN. Several papers in the past year have reported single-cell transcriptome data within the OSN lineage, finding that multiple OR genes per cell are detected in less mature subpopulations (Hanchate et al., 2015; Tan et al., 2015). This observation raises the possibility of an iterative process, whereby an initial small subset of ORs might transcribe at low levels, but competitors are eventually silenced as the process ensues (Abdus-Saboor et al., 2016; Tan et al., 2013; Tian and Ma, 2008). LSD1 might be critical in such an iterative silencing process, e.g., by removing H3K4 methylation at competing loci (as discussed in Tian et al., 2016).
This paradigm predicts that the loss of LSD1 might result in multiple expressed OR loci per cell due to the accumulation of H3K4 methylation marks at more than one locus per cell. To seek evidence in support of this hypothesis, we conducted ChIP at OR loci, where the ~doubling in the incidence of OR activations in KD populations (e.g., from ~1% to ~2%) might correlate with a corresponding ~doubling in H3K4 methylation at expressed OR loci. The impact on H3K4 methylation in KD populations appears to be more subtle than this hypothesis might predict. At early passages, most ORs tested exhibit H3K4 methylation near the background noise levels; nevertheless, we cannot dismiss the possibility of a systematic doubling of H3K4 methylation at OR loci, since doubling the incidence might still be at or near detection thresholds (e.g., from ~1% of templates to ~2% of templates for a given OR within a heterogeneous cell population). At later passages, we observe a striking systematic increase in H3K4 methylation across a broad panel of ORs tested. This observation affirms that LSD1 would seem to play a role in maintaining lower levels of H3K4 methylation at OR promoters at the pre-commitment stage. However, we do not observe a systematic increase in OR expression levels at later passages as compared to early passages (not shown), as would be predicted if further accumulation of this histone mark was alone sufficient to drive more frequent OR expression.

An alternative hypothesis attributes the perturbation of singular OR expression per cell in LSD1-depleted populations as an indirect consequence of
interactions with other regulatory factors as opposed to a direct consequence of H3K4 elevation. For example, a critical activation factor might normally be sequestered within an LSD1 protein complex, thus limiting its access to competing OR promoters; the depletion of LSD1 in turn might lead to liberation of this putative activator so that it can interact with additional OR promoters (i.e., more than one OR). The LSD1 compartments described previously (Kilinc et al., 2016) might represent a scaffold for such an activator to prevent multiple/widespread OR associations. Unlike the H3K4-mediated activation hypothesis described above, this hypothesis does not predict a cumulative effect on OR expression within an LSD1-depleted environment, since once the LSD1 complex is disrupted, presumably the same access to a putative activator would be granted to competing ORs irrespective of how long the culture has grown in the absence of LSD1.

The two above hypotheses are not mutually exclusive - for example, acquisition of H3K4 methylation at multiple OR genes per cell might increase the probability of multiple activations per cell, and in addition, a putative activator normally sequestered by/complexed with LSD1 might also permit only a limited number of activation events per cell. A hallmark of both models is the notion of a limiting factor enabling OR activation (i.e., either H3K4 methylation being limiting and/or a putative activator protein being limiting). Thus, both models might predict that even when the system is perturbed, there might be an upper limit on the number of OR genes that can be activated per cell and/or the total
OR RNA output that can be generated when more than one OR gene is active per cell. Such a consideration would account for why both LBR overexpression and G9a/GLP double-knockout mice that perturb monogenic OR expression result in a finite number of ORs expressed per cell (Clowney et al., 2012; Lyons et al., 2014), as well as a compensatory reduction in OR expression levels per cell when multiple ORs are activate (Clowney et al., 2012; Hanchate et al., 2015).

An important observation in this study is the fact that LSD1 depletion did not perturb de novo OR activation events that occur during “switching” (Fig. 2). The LSD1 knockout mouse (Lyons et al., 2013) similarly suggests that LSD1 is dispensable for de novo OR activations, since these mice nevertheless express a broad set of OR genes, albeit at reduced levels. If LSD1-mediated removal of H3K9 marks is not required for OR activation per se, it could mean that this chromatin modification might occur downstream of selection (as opposed to a pre-requisite for selection). For example, perhaps LSD1-mediated H3K9 demethylation at the selected OR is important to ensure robust expression levels once an OR is chosen, thus explaining why LSD1 knockout mice exhibit lower OR expression levels, but not abolished OR expression (Lyons et al., 2013), and/or to ensure that this chosen OR escapes recruitment to nuclear chromocenters (where the other ORs are sequestered, possibly as a consequence of retaining this mark). It is also important to note that LSD1 targets the H3K9me2 mark, not the H3K9me3 mark, so it remains possible that the previous demethylation step from H3K9me3 to H3K9me2 (by an as yet uncharacterized protein) might be a
prerequisite to OR activation, even if the subsequent LSD1-mediated
demethylation step to follow is unnecessary at this stage.

We attempt to reconcile a model that is consistent with all the published
data, including observations made in the OP6 cell line. Rather than proposing a
single, rate-limiting step for OR selection (e.g., K9 demethylation), we instead
propose a model where both K9 and K4 methylation states are contributing to
initial OR selection. Proteins such as G9a (Lyons et al., 2014), along with LSD1
(Ooi and Wood, 2007; Roopra et al., 2004; herein), might collaborate to ensure
that K9 methylation (added/maintained by G9a) dominates across OR loci and
acquisition of K4 methylation is rare (aided by LSD1-mediated removal).
Perturbation of either enzyme (G9a, (Lyons et al., 2014); LSD1, herein) might
shift the equilibrium state such that K9 is less predominant (and K4 is less rare),
thereby increasing the probability of multiple OR activations per cell. The
disruption to monogenic OR expression observed here is reminiscent to the G9a
knockout mouse – neither results in a dramatic perturbation of K9/K4
methylation states and both seem to exhibit a modest multigenic phenotype –
suggestive of a more subtle role for both proteins in maintaining an appropriate
balance of the two marks. Such a model might also predict stochastic imbalances
arising in vivo, leading to rare and/or transient multi-OR cells during the initial
selection stages, an observation that has recently been made by two
independent studies (Hanchate et al., 2015; Tan et al., 2015). Importantly, we
suggest that initial choice might be more of a dynamic process in which
management of K9 and K4 marks leads to one (or a small number) of preliminary candidates, that will subsequently be refined, amplified, and stabilized during the feedback-mediated commitment process. There are certainly other models that could account for all current observations, but a key aspect of the described model is a shift away from regarding LSD1 as the essential and primary activating factor, as well as H3K9 demethylation as the direct causative molecular event in the establishment of monogenic OR expression, while raising new perspective about LSD1’s role in the process of excluding competitor OR loci.

Finally, our results might inspire two interesting questions within the field of epigenetic transcriptional regulation that warrant further study. First, the LSD1 requirement for maintaining OR monoallelic expression is perhaps surprising, given it would be reasonable to assume that allelic regulation is established in a non-lineage-specific manner in the early embryo, as it seems to be for other gene loci exhibiting allelic exclusion (Chess, 2012, 2013; Eckersley-Maslin and Spector, 2014; Ensminger and Chess, 2004; Savova et al., 2013). We note that in our two-color RNA FISH experiments, we observed approximately the same incidence of multiallelic and multigenic OR expression (Fig. 4B). This finding seems to support a hypothesis that the second OR allele in a given cell is treated equivalently with respect to selection as any other OR gene locus. Second, the simultaneous duality of LSD1 as possibly both a positive (Lyons et al., 2013) and negative (herein) regulator of OR transcription (i.e., within the
same cellular and developmental context), has not yet been established for other LSD1 regulatory targets, which broadens our perspective on the versatility of this important regulatory factor. This latter consideration might be especially critical in the context of OR gene switching, where presumably both activation and deactivation must be precisely coupled.

2.5 Concluding Remarks on Chapter Two

The “Repress all, de-repress one” model for OR singularity proposed by the Lomvardas group does not account for a finite failure rate of singular OR de-repression, as evident in single cell RNA-seq studies and OR refinement observations (Hanchate et al., 2015; Nagai et al., 2016). In the above studies, I showed that an assumption in this model, that LSD1 is a “gate-keeper” that ensures singular OR activations via an exclusive H3K9 demethylation event is incorrect, at least in the context of OP6 cells.

As noted previously, computational simulations from the Xing lab suggested that transient and low level of LSD1 could favor singular OR H3K9 demethylation, whereas prolonged and high levels of LSD1 would favor multiple OR H3K9/H3K4 demethylation events (Tian et al., 2016). Given the abundance of LSD1 protein in OSN precursors where OR choice occurs, it is possible that multiple ORs are demethylated at H3K9, and subsequently acquire H3K4 methylation marks. In this scenario, it becomes necessary to remove additional H3K4 methylation marks from these competing ORs to retain singularity. My
results support such a model: depletion of LSD1 reduces the cell’s ability to remove these additional H3K4 methylation marks, causing a shift in the equilibrium state from one to a few active ORs per cell. This conclusion is further supported by the observation that H3K4 methylation appears to be systematically increased at OR loci in LSD1-depleted cells.

We suggest a model for initial OR gene selection that depends on a delicate balance in the management of both the repressive H3K9 and activating H3K4 marks at OR loci. A prediction of this hypothesis is that perturbation of other proteins involved in maintaining this balanced state would produce similar phenotypes. For example, the loss of G9a, also thought to contribute to the silencing of competing ORs, might phenocopy the LSD1 depletion experiment, to lesser or greater extent. I designed a G9a-depletion experiment to test this prediction in the next set of experiments described in Chapter 3. Also, this hypothesis predicts that over-expression of LSD1 might rescue a multi-OR phenotype in a G9a-depleted background, or vice versa, over-expression of G9a might rescue a multi-OR phenotype in an LSD1-depleted background – each of these experiments would test the hypothesis that key chromatin regulatory proteins (like G9a and LSD1) collaborate to ensure an equilibrium state that is strongly biased towards OR silencing, but that disruption of one or more of these regulators might shift the equilibrium towards higher probability of multiple activations. I discuss these and other future experiments concerning this “equilibrium” model for OR regulation more extensively in Chapter 4 ("Future
Directions”), because, to our surprise, the G9a-depletion experiment described in the next Chapter produced a much more severe and interesting phenotype in which the differentiated state of OP6 cells was destabilized.

References


facilitates the diversification of olfactory neurons. Cell Rep, 9(3), 884-892. doi:10.1016/j.celrep.2014.10.001


optimization design. *Proc Natl Acad Sci U S A.*
doi:10.1073/pnas.1601722113


Chapter 3: G9a perturbation transforms an olfactory neuronal cell line to a pluripotent state capable of self-organization

The specific aim of this chapter of my thesis was to investigate the role of G9a in OSN development. As discussed in Chapter 1, the heterochromatin-based silencing of OR genes provide a uniform platform on which a stochastic and a singular OR gene choice can be made (Magklara et al., 2011; Lyons et al., 2013; Lyons et al., 2014). OR heterochromatinization prevents premature and spurious OR gene activations in the OSN lineage (Lyons et al., 2014), thereby enabling functional identity and proper axonal targeting of OSNs (Reviewed in Monahan and Lomvardas, 2016).

G9a is a histone methyl transferase that adds H3K9me2/1 repressive histone marks at variety of genes in various developmental contexts, including at OR loci (Tachibana et al., 2002; 2005; Chaturvedi et al., 2012; Sankaran et al., 2013; Ingawa et al., 2013; Lyons et al., 2014). It has been proposed that G9a/GLP-mediated addition of repressive marks ensures that all OR clusters acquire a similar silenced ground state and recruitment to a heterochromatic compartment, after which singular OR gene de-repression occurs to generate an OR-specific OSN (Magklara et al., 2011). Depletion of G9a/GLP in vivo reduces OR expression levels, as well as OR representation, and as anticipated, some OSNs become multigenic with respect to OR expression per cell (Lyons et al., 2014). However, an important conclusion from these experiments is that not all ORs
become “liberated” for multigenic expression in this genetic background, as
might be predicted by the hypothesis that G9a-mediated repression is a
necessary pre-requisite for keeping competing ORs silenced through the
selection process. In addition, one study suggested that H3K9me3
heterochromatinization is reduced in the G9a depleted OSNs, however, ORs were
still able to aggregate in this study, suggesting G9a-mediated H3K9me3 is not
determining the formation of OR heterochromatic foci (Lyons et al., 2014).
Moreover, there is evidence for additional, superimposed mechanisms beyond
just the G9a/GLP-mediated silencing that enforce OR singularity. For example,
H3K4 demethylation via LSD1 (Vyas et al., 2017), multiple enhancer interactions
(Markenscoff-Papadimitriou et al., 2014), and post selection refinement (Abdus-
Saboor et al., 2016) all point to redundant and additional, more complex layers
of OR co-regulation.

After demonstrating that LSD1 depletion in OP6 cells resulted in a
phenotype in which single cells expressed multiple OR genes/alleles (Chapter 2),
we considered an iterative equilibrium model in which multiple proteins
collaborate to ensure that one and only one OR allele acquires a permissive
chromatin state. For example, opposing enzymatic activities of LSD1 and G9a
might function to add or remove H3K9 methylation, respectively. LSD1, and
perhaps other demethylases might further collaborate to ensure that active
marks, such as H3K4 methylation, are kept in check. Specifically, we propose
that G9a and LSD1 might function collaboratively to retain an epigenetic
landscape that is significantly biased towards ubiquitous H3K9 methylation and rare H3K4 methylation at OR loci, respectively. A prediction of this hypothesis is that perturbation of G9a might produce a multi-genic OR phenotype similar to that observed in LSD1 depleted cells, that the double mutant might have a more extreme multi-genic phenotype, and that over-expression of one or both enzymes might partially or fully rescue these phenotypes. With these questions in mind, I designed a G9a depletion and an LSD1 over-expression experiment in OP6 cells. As discussed below, the former produced a surprising phenotype in which OP6 cells appeared to transform their identities into cell types that resembled stem cells, both in terms of behavior and gene expression. The latter experiment (LSD1 over-expression) was not fully investigated due to lack of available time, and is the subject of discussion in my “Future Directions” chapter of this thesis (Chapter 4).

Portions of the following are derived from a draft of a manuscript co-authored by myself and my thesis advisor, Dr. Lane, written during the summer of 2018. Ultimately, we did not submit this manuscript, as we believed the results were too preliminary and required additional experimental evidence for an initial publication on this work (manuscript in preparation).
3.1 Introduction

The process of cell differentiation enables multicellular organisms to develop complex tissues consisting of multiple cell types, as well as to maintain homeostatic balance of these tissues during growth or following injury (Grafi, 2004; Lee and Seo, 2018; Takahashi & Yamanaka, 2006; Jopling, Boue, and Izpisua-Belmonte, 2011; Jopling et al., 2010; Kopp, Grompe, and Sander, 2016; Kragl et al., 2009; Kumar et al., 2011; Tata et al., 2013). The transition from the earliest stem cells of the embryo to a fully mature and differentiated cell type is a multistep process involving the repression of pluripotent factors, such as *nanog*, *Oct4*, and *Sox2* (Hawkins et al., 2010; J. Y. Li, Patterson, Mikkola, Lowry, & Kurdistani, 2012; Loebel, Watson, De Young, & Tam, 2003; Swierczek, Ciemerych, & Archacka, 2015; Takahashi & Yamanaka, 2006; Zhu et al., 2013), and the activation of factors that progressively restrict the potential along a well-defined cell lineage (Reviewed in Jaenisch & Young, 2008; M. Li & Belmonte, 2017). Accordingly, cell differentiation is accompanied by the establishment of a post-mitotic state and enhanced senescence (van Deursen, 2014), chromatin-mediated silencing of pluripotent genes (Hochedlinger & Jaenisch, 2015; Jaenisch and Young, 2008), and activation of genetic networks that establish appropriate cell morphology, structure, and behavior (Jopling, Boue, & Izpisua Belmonte, 2011; Jopling et al., 2010; Kopp, Grompe, & Sander, 2016; Kragl et al., 2009; Kumar et al., 2011; Tata et al., 2013).
A perspective of the differentiated state is offered by Waddington’s classic epigenetic landscape model (Takahashi, 2012), in which cells are trapped by an epigenetic barrier, requiring an active process to reverse back to a more immature state (Waddington, 1957). Hence, a dogma in development has been that lineage progression is unidirectional. However, several recent examples of de-differentiation have been documented in a number of model systems, evident by the acquisition of stem-cell-like qualities, such as re-entry into the cell cycle and re-acquisition of pluripotency (Brawley & Matunis, 2004; Brockes & Kumar, 2002; Hsu, Pasolli, & Fuchs, 2011; Jopling et al., 2010; Rinkevich, Lindau, Ueno, Longaker, & Weissman, 2011; van Es et al., 2012). The capability of cell populations to de-differentiate is important in the context of shifting cellular resources to meet specific demands in the context of injury (Lin et al., 2017). The process of de-differentiation as documented in several contexts depends on the reprogramming of the epigenetic landscape in order to reactivate key pluripotent factors that suppress lineage-specific genetic networks (Eguizabal et al., 2016; Hochedlinger & Jaenisch, 2015; K. Lee & Seo, 2018; M. Li & Belmonte, 2017; K. H. Ma, Hung, & Svaren, 2016; Pociask et al., 2017). A well-documented example of de-differentiation is in plant cells. Plant cell de-differentiation progresses through many distinct stages of chromatin de-condensation in order to acquire a cell fate switch (Grafi, 2004). This de-differentiation is accompanied by redistribution of HP1, inactivation of pRB/E2F (master regulators of the cell cycle, differentiation, and apoptosis), transcriptional activation of pRB/E2F-
regulated genes, and disruption of the nucleolar domain to metabolically activate 18S ribosomal genes (Endo and Nadal-Ginard, 1998). In differentiated plant cells, pRB/E2F target genes (e.g., RNR2 and PCNA) are condensed and silenced, however during de-differentiation these genes become decondensed and transcriptionally active (Sage et al., 2003; Ma et al., 2000). The HP1-mediated process of chromatin re-organization resets the genetic expression program, in particular, to activate pluripotent gene networks (Williams et al., 2003).

The development and maintenance of the mammalian olfactory epithelium is a particularly interesting context for investigating developmental plasticity for three reasons. First, the mammalian olfactory lineage generates >1,000 different types of specialized olfactory sensory neurons (OSNs) defined by the specific odorant receptor (OR) expressed on their ciliated surfaces (Reviewed in Buck, 2004; Monahan & Lomvardas, 2015; I. Rodriguez, 2013), and this differentiation process involves complicated interplay between extrinsic and intrinsic cues that generate significant epigenetic differences within OSN subpopulations (Lyons & Lomvardas, 2014). Second, these OSN populations must be actively maintained through the life of the organism due to injury and depletion, requiring a persistent population of stem cells that support ongoing neurogenesis in the adult in such a way to repopulate a broad range of sensory neuronal and non-neuronal cell types (Schwob et al., 2017). And third, with certain environmental triggers (e.g., acute injury), it has been shown that the earliest cells of the OSN lineage de-differentiate to pluripotent stem cells capable
of regenerating a host of neuronal and non-neuronal cell types of the olfactory epithelium (Lin et al., 2017; discussed in more detail below).

The olfactory epithelium (OE) contains two distinct stem cell populations located in the basal layer: globose basal cells (GBCs) and horizontal basal cells (HBCs) (Graziadei & Graziadei, 1979; Holbrook, Szumowski, & Schwob, 1995; Huard & Schwob, 1995). GBCs are mitotically active and heterogeneous with respect to marker gene expression, capable of giving rise to both neuronal (OSNs) and non-neuronal (sustantacular cells, microvilli cells, gland and duct cells) cell types (Asan & Drenckhahn, 2005; M. Chen et al., 2014; Hegg, Jia, Chick, Restrepo, & Hansen, 2010; Huard, Youngentob, Goldstein, Luskin, & Schwob, 1998; Moran, Rowley, & Jafek, 1982; Morrison & Costanzo, 1992). GBCs, as expected, express proteins shown to be important for pluripotency (e.g., Sox2), as well as factors associated with the olfactory lineage (e.g., Pax6) (Gordon, Mumm, Davis, Holcomb, & Calof, 1995; Guillemot et al., 1993; Guo et al., 2010; Manglapus, Youngentob, & Schwob, 2004). The first stage of GBC differentiation is defined by the activation of a lineage marker gene; for example, GBC differentiation along the duct/gland lineage is associated with activation of the Ascl3 and Sox9 markers (Holbrook, Wu, Curry, Lin, & Schwob, 2011; Weng, Vinjamuri, & Ovitt, 2016), GBC differentiation along the sustantacular lineage is associated with activation of the Hes1 marker (Herrick, Guo, Jang, Schnitttke, & Schwob, 2018; S. Rodriguez et al., 2008), and GBC differentiation along the OSN lineage is associated with the activation of the Ascl1 marker (Krolewski, Packard,
The latter cells are referred to as transit amplifying progenitors, and can further differentiate into immature neurons that activate neural markers, such as *Neurog1* and *NeuroD1* (Krolewski, Packard, & Schwob, 2013; Packard, Giel-Moloney, Leiter, & Schwob, 2011). These immature neurons further progress to mature bipolar OSNs, which express terminal markers such as *GAP43* and *OMP* (Iwema & Schwob, 2003; A. C. Lee, He, & Ma, 2011; Nickell, Breheny, Stromberg, & McClintock, 2012). The second stem cell type, the HBCs, are normally mitotically inactive and seem to function as a “reserve” stem cell population whose proliferation is triggered by severe injury or replacement demands ( Reviewed in Duggan & Ngai, 2007; Herrick, Lin, Peterson, Schnittke, & Schwob, 2017; Huard & Schwob, 1995; Packard, Schnittke, Romano, Sinha, & Schwob, 2011 ), associated with the down-regulation of the *Trp63* gene (not expressed in GBCs) that might be a key factor in establishing HBC quiescence (Schnittke et al., 2015).

Interestingly, an elegant lineage tracing strategy demonstrated that OSN-committed, cycling progenitors (e.g., *Asc1*, *Neurog1*) dedifferentiated into GBC-like cells expressing the *Sox2* pluripotency marker in response to acute injury (Lin et al., 2017). This dedifferentiation response was enhanced by inhibition of a chromatin modifying gene, *EZH2*, which is a component of the *polycomb* repressive complex responsible for H3K27 methylation associated with developmental gene silencing (Bernstein et al., 2006b; Bracken, Dietrich,
Pasini, Hansen, & Helin, 2006; O’Carroll et al., 2001). These results implicate global epigenetic reprogramming as conducive to dedifferentiation in the OSN lineage. This perspective is consistent with studies characterizing chromatin properties during differentiation, where progressively more compacted chromatin was observed with OE neurogenesis (Le Gros et al., 2016b), whereas stem cell populations have been characterized as having more open chromatin states (Gaspar-Maia, Alajem, Meshorer, & Ramalho-Santos, 2011; Ugarte et al., 2015). The heterochromatin reader protein HP1b regulates chromatin compaction and distribution of the euchromatin and heterochromatin in mature OSNs (Le Gros et al., 2016). In the absence of HP1b, heterochromatin de-compacts and re-distributes towards the nuclear periphery mimicking the nuclear organization states of the olfactory stem cell population (Le Gros et al., 2016). Taken together, it is possible that merely the de-repression of key pluripotency genes, such as Sox2, might be sufficient in certain developmental contexts to drive gene networks that maintain stem-like qualities while suppressing lineage progression (Reviewed in M. Li & Belmonte, 2018; P. Liu, Chen, Liu, Qi, & Ding, 2018; Masui et al., 2007; Ng & Surani, 2011).

One key observation in the above study by Lin et al is that the Ascl1+/Neurog1+ cells capable of dedifferentiation in this paradigm are very early OSN progenitors that are mitotically active. It is not clear whether cell cycle factors are critical collaborators in this process; e.g., that might otherwise preclude a similar dedifferentiation process in post-mitotic cells further along in
For several years, we have utilized an immortalized cell line transformed by a temperature-sensitive large-T antigen, derived from the E10.5 mouse olfactory placode (Illing et al., 2002), in order to study OSN development and differentiation. The OP6 founder cell was staged as a post-mitotic intermediate-late immature neuron along the OSN lineage (Kilinc, Savarino, Coleman, Schwob, & Lane, 2016; Pathak, Johnson, Getman, & Lane, 2009b; Vyas, Meredith, & Lane, 2017), expressing GAP43 and BF1 consistent with a more advanced immature neuron (Duggan, DeMaria, Baudhuin, Stafford, & Ngai, 2008; Weiler & Benali, 2005), but not the early transit-amplifier marker, Ascl1 (Murray et al., 2003) nor the mature OSN marker, OMP (A. C. Lee et al., 2011). When deactivating the large-T antigen and inducing further differentiation with retinoic acid, OP6 cells mature into bipolar neurons expressing mature markers such as G-Olf, ACIII, OCNC1, and OMP (Illing et al., 2002; Nickell et al., 2012; Pathak et al., 2009b).

In this study, with the initial goal of studying the impact of histone methylation in the context of OR gene regulation in OP6 cells, we perturbed the chromatin modifying enzyme, G9a, a histone methyltransferase gene whose protein product catalyzes H3K9 methylation thought to contribute to the silencing of all but one OR gene in each cell (Lyons et al., 2013; Lyons et al., 2014). Surprisingly, we observed apparent dedifferentiation of OP6 cells into what might be a pluripotent cell capable of robust stem cell-like colony formation and eventual re-differentiation along various cell lineages. Here, we
describe the up-regulation of stem-like markers in transformed colonies, the reorganization of heterochromatin with transformed nuclei, and the apparent trans-differentiation into elaborate, non-neuronal tissue.

3.2 Materials and Methods

3.2.A OP6 and 293T Cell Cultures. The OP6 cell line was cultured at 33°C in Dulbecco's Modified Eagle's Medium (DMEM, Life Technologies) supplemented with 10% fetal bovine serum (FBS; Gibco), as described previously (Illing et al., 2002). The 293T cell line was cultured at 37°C in DMEM supplemented with L-glutamine (Gibco), 10% FBS, and 1% penicillin/streptomycin (Gibco).

3.2.B CRISPR-Cas9 Gene Targeting. We adopted the CRISPR-Cas9 protocol according to manufacturer protocols (Addgene). Briefly, Cas9 target sites were selected using the sgRNA design tool available at the Broad Institute (http://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design). We designed sgRNAs targeting Exon 8 of the G9a gene, as well as a scrambled sequence negative control that should not target endogenous mouse loci (sequences available upon request). Complementary oligonucleotides 20 bp in length were synthesized, annealed, and cloned into the lentiviral CRISPR-Cas9 delivery vector containing spCas9-eGFP under control of the elongation factor-1α (EF-1α) promoter and the sgRNA under control of the U6 promoter (Addgene). Following transformation into competent Stbl3 bacteria (Invitrogen), positive colonies were verified by sequencing and plasmids prepared using the
Qiagen miniprep kit (Qiagen). Resulting CRISPR/Cas9 lentiviral constructs were transiently transfected along with packaging plasmids pCMV-dR8.2dvpr and pCMV-VSV-G (Addgene) into HEK-293T cells (ATCC) using standard calcium phosphate protocols (Invitrogen). Virus-containing supernatants were collected 36- and 60-hours post-transfection and concentrated by ultracentrifugation. Approximately $10^5$ host OP6 cells were subjected to spin infection at RT for 90 minutes in the presence of 8 μg/ml polybrene as described elsewhere (Stewart et al., 2003). Lentivirus-infected OP6 cells/colonies were selected using GFP for subsequent secondary colony analyses.

3.2.C cDNA Analyses. To enrich for the cells at the core of transformed GFP+ colonies, we subjected dissected colonies to a series of mild trypsinization-centrifugation-wash cycles to remove more loosely attached peripheral cells. Total RNA was extracted from the remaining portions of these colonies using Trizol (Thermo Fisher), and first-strand cDNA produced using superscript reverse transcriptase III per manufacturer protocols (Invitrogen). All cDNA preparations included parallel negative control samples in which the reverse transcriptase was omitted (“no-RT controls”). Quantitative PCR (qPCR) was conducted on an ABI 7300 real-time PCR machine (Applied Biosystems) for several housekeeping and developmental genes (primer sequences and PCR conditions available upon request).

3.2.D Immunocytochemistry. Cells and colonies were fixed for in 3% paraformaldehyde (for 10 minutes) and 8% paraformaldehyde (for 30 minutes),
respectively, in PBS buffer containing 0.5% Triton-X (Sigma). Fixed cells were blocked in PBS containing 1% BSA and 0.1% Tween-20 for 20 minutes at 37°C. Primary and secondary antibody incubations were performed at 37°C for 45 minutes in a humidified chamber (except for primary incubations on colonies, which were performed overnight at 4°C). Non-specific antibody interactions were washed in PBS containing 0.1% Tween-20 at 5 minutes intervals and with gentle agitation. The following primary antibodies and dilutions were used in this study: rabbit anti-G9a/EHMT2 (C6H3) (Cell Signaling, 3306T, 1:200), anti-histone H3K9me3 (Abcam, ab8898, 1:200), and rat anti-HP1 beta/CBX1 (Abcam, ab10811, 1:200). The following secondary antibodies and dilutions were used: donkey anti-rabbit Alexa 488 (Jackson Immunoresearch, 711-545-152, 1:100), goat anti-rabbit-Cy3 (Millipore, AP132C, 1:800), and donkey anti-rat-Cy3 (Jackson Immunoresearch, 712-165-153). Cell nuclei were stained with DAPI (1 mg/ml) supplied within the Vectashield mounting solution (Vector Labs).

Scramble OP6 population was used as negative control for all IF analysis. An alkaline phosphatase substrate, BCIP/NBT (Sigma, B3804), was used to stain stem-like colonies. Internal negative control staining for alkaline phosphatase (AP) suggested that majority of the undifferentiated OP6 cells were negative for AP staining. Cells within the central portion of the colonies containing rounded morphology are stained intensely positive for AP. The surrounded non-GFP epithelial-like cells do not stain positive for AP, suggesting only the CRISPR infected OP6 cells possesses the ES-like molecular properties. This internal
negative control analysis confirms the absence of high false-positives. However, there may be instances where fibroblast like cells lowly stain positive for AP, this rare population of cells can be accounted for false positive, alternatively represent late re-differentiation staged cells. It may be possible that AP chemical reaction generated blue coloration that may have spread in cells residing in the nearby vicinity to the colony through diffusion.

3.2.E Image Analysis. Immunofluorescence images were acquired using a Deltavision RT imaging system (Applied Precision) adapted to an Olympus (IX71) microscope equipped with XYZ motorized stage. Images were processed using Softworx (Applied Precision). Individual nuclei were optically sectioned at 0.5 μm intervals to resolve spatial signal patterns, and in some cases, these confocal images were projected onto a single plane for visualization. To score for G9a+/G9a- cells within the colonies, we counted cells in each Z-sections of colonies ranging from small to medium sizes and categorized cells according to G9a immunostaining intensities. Emergence of developmental biological structures were rare (~7 structures derived from ~500 colonies) and distributed across different plates. At least one structure per plate was evident after ~3 weeks of passaging secondary colonies in various plates. Usually, bigger colonies (comprising >500 cells) were more likely to further develop into a structure. Due to lack of live cell imaging facilities, we took snapshots of emerging structures in pseudo-time (Fig. 5).
3.3 Results

The *G9a* methyltransferase is a member of the *Suv39h* group of SET domain-containing proteins that catalyze addition of methyl groups on the lysine-9 residue of histone-3 (*H3K9*) (Tachibana, Sugimoto, Fukushima, & Shinkai, 2001). *H3K9* tri-methylation (*H3K9me3*) is required for maintenance of pericentric heterochromatin and transcriptional silencing, including the silencing of developmentally regulated genes (Becker, Nicetto, & Zaret, 2016; Grewal & Klar, 1996; E. Li, 2002).

In this study, we sought to investigate a role for *G9a* in the competitive silencing of odorant receptor (OR) genes during olfactory sensory neuronal (OSN) development, motivated by the observation that the *H3K9me3* histone mark, as well as pericentric (chromocenter) co-localization in the nucleus, is associated with the silencing of all but one OR gene in OSNs (Clowney et al., 2012; Kilinc, Meredith, & Lane, 2014; Lyons et al., 2014; Magklara et al., 2011; Vyas et al., 2017). We generated a CRISPR-Cas9 cassette with a synthetic guide RNA directed against mouse *G9a* in order to perturb heterochromatin regulation within a cell line (OP6) derived from immature cells along this lineage (Heck et al., 2014; Sanjana, Shalem, & Zhang, 2014).

The OP6 cell line is clonally derived from a founder cell that had been previously staged as a post-mitotic immature sensory neuron from E10.5 mouse olfactory placode (Illing et al., 2002). OP6 cell populations appear morphologically homogeneous and express several pre-OSN marker genes,
including ~one OR gene per cell (Illing et al., 2002; Pathak, Johnson, Getman, & Lane, 2009a). On the other hand, OP6 cells are not fully differentiated neurons of the lineage given their lack of bipolar morphology, as well as the lack of expression of key developmental genes associated with mature OSNs, e.g., OMP, (Illing et al., 2002; Kilinc et al., 2016; Vyas et al., 2017). Therefore, the OP6 cell line is likely a committed progenitor along the OSN lineage; indeed, culturing OP6 populations in retinoic acid results in further OSN maturation, including extension of bipolar axons and expression of more mature neuronal marker genes (Illing et al., 2002; Kilinc et al., 2014). Over decades of culturing, to our knowledge there is no report of OP6 cell de-differentiation or alternative developmental potential. As detailed below, we report here that the perturbation of G9a expression in OP6 cells resulted in a rapid morphological change, colony formation, and apparent reversion to an earlier stem-like/pluripotent state, with resulting transformed colonies apparently capable of differentiating and organizing cells into elaborate structures not associated with the OSN lineage.


We infected OP6 populations with a pL-CRISPR.EFS.GFP lentivirus all-in-one CRISPR cassette (Heck et al., 2014), containing synthetic guide RNAs (sgRNAs) targeting exon 8 of the mouse G9a gene locus (see Methods, section 3.2B). Post-infection, we observed rare (~5%) OP6 cells expressing GFP, as reported by
stable expression of the GFP cassette (Fig. 1A, left panels). We note that within 48-72 hrs post infection, some GFP-expressing cells exhibit altered morphology, as compared to surrounding GFP-negative cells in the culture that serve as an internal control (Fig. 1A, right panels). The positive cell shown in the right panels of Figure 1A does not appear to have yet divided – there are no surrounding GFP-positive/rounded cells in the vicinity – although it is possible that it has already progressed through the cell cycle prior to detaching and reattaching in the location shown in the image. We monitored the growth of GFP-positive transformed cells in these primary post-infection cultures for ~one week, observing rapid growth and the formation of 3D colonies (Fig. 1B). We estimate that ~20% of the original subpopulation of GFP-positive cells formed robust colonies (Fig. 1B, left panel). After 7 days in culture, we observed few, if any, isolated GFP-positive cells. We do not know if GFP-positive cells that fail to form colonies are selectively lost from the population, were consumed within other expanding colonies, or if they persisted yet with extinguished GFP expression.

After 7 days post-infection, we enriched for transformed cells by scraping off a primary colony, and dispersed its cells into a fresh culture environment. We found that it was difficult to completely disrupt the cell mass of this colony, and even after rigorous agitation, we still observed small clumps of cells, suggesting that these transformed colonies are tightly compacted and stabilized by cell-cell contacts, an attribute of stem cell colonies (Jaenisch and Young, 2008). We
passed the disrupted cells from one primary colony into multiple culturing wells. Within 48 hours, we observed new (secondary) colonies attached and growing in each of these wells (typically, 3-4 colonies per well). Examples of colony formation at various stages are shown in the panels of Figure 1B.

A noteworthy feature of these transformed colonies is that they resemble something like “embryoid bodies” (Koike, Sakaki, Amano, & Kurosawa, 2007), which are defined as three-dimensional aggregates of pluripotent stem cells capable of executing embryonic development (e.g., blastocyst stages of mouse embryos). Transformed colonies are border-defined, exhibit brownish pigmentation (Fig. 1C, right panel), and as noted previously, exhibit a surrounding halo of less compacted, “differentiated” cells, all characteristic features of embryoid bodies (Alison, Wobus, & Boheler, 2006; Sakai, Yoshiura, & Nakazawa, 2011). As a preliminary assay, we stained these transformed colonies for alkaline phosphatase (AP) expression, a convenient marker commonly used to distinguish embryonic stem cells (Andrews, Meyer, Bednarz, & Harris, 1984; Onder et al., 2012; Thomson et al., 1998), and find that these colonies are indeed, AP positive (Fig. 1C, left panel). The non-transformed cells that are not part of colonies do not express AP. We will return to a more thorough investigation of stem-like properties in subsequent section of this report (Section 3.3.D).
Figure 1. G9a-targeted CRISPR-Cas9 Treatment Transforms OP6 Cell Morphology and Growth.

A. Transformed OP6 cell morphology. Rare infected cells express GFP (left two panels), with initial morphology that resembles wild-type OP6 cells. Within first 48-72 hrs, some GFP-positive cells exhibit altered morphology (compacted, rounded) as compared to surrounding GFP-negative cells (right two panels). Scale bars = 10μm. B. Colony formation in primary cultures. A colony of GFP-positive cells is initially sparsely organized with only rare morphologically altered (rounded) cells (arrows, left panel). Scale bar = 20μm. A lower magnification image (second panel, scale bar = 50μm) of a region of the primary culture containing several colonies at various stages of growth, including sparsely organized colonies with rare rounded cells (arrow 1), larger and more compacted colonies with increased rounded cells within the central core regions (arrow 2), and highly compacted, well delineated colonies of various sizes (arrow 3). The third panel (scale bar = 50μm) shows further maturation of colonies, with unpigmented smaller colonies (arrow 1) and larger colonies exhibiting pigmentation (arrow 2). Colony formation is also evident in OP6 cultures treated with a G9a inhibitor drug (right panel, arrow). C. Transformed colonies resemble embryoid bodies. Following passaging of a primary transformed colony, multiple colonies develop in secondary cultures that are alkaline-phosphatase positive (a representative AP+ colony, left panel). These secondary colonies grow with a dense, GFP-positive 3D cell mass in the interior with a monolayer of flattened/elongated cell types forming a peripheral halo (middle panels). Secondary colonies are also pigmented (right panel).
In order to further support our conclusions that G9a perturbation resulted in OP6 cell transformation, we conducted an independent perturbation experiment using a small molecule G9a inhibitor/drug, BIX01294 (Chang et al., 2009; Kubicek et al., 2007) on previously untreated OP6 cells. We explored a range of drug dosages in the media between 1-15 μM, and found that concentrations >5 μM were toxic to the cells, whereas a 1 μM dosage did not produce any observable effect. We grew 3 culture wells each at a 2-μM and 5-μM dosage of BIX01294 and observed colony formation in four of these six culture dishes. These colonies resembled growth and morphology characteristics observed previously for colonies generated following the G9a-CRISPR treatment (Fig. 1B, right panel). A total of six primary colonies were observed in these cultures, all evident between 24-48 hours after introduction of the drug. Additional work with this drug by another graduate student in the lab validated these initial observations (Joyce Noble, personal communication). These results strongly implicate G9a perturbation (via genetic or drug treatment) as inducing OP6 transformation and colony growth.

3.3.B  CRISPR-Cas9 Appears to Perturb G9a Expression without Gene Deletion.

The CRISPR-Cas9 procedure utilizes a guide RNA to target the Cas9 endonuclease to a specific genomic locus where it would typically make a small deletion designed to interrupt the function of the targeted gene (Jiang & Doudna,
We therefore expected that the transformed GFP-positive OP6 cells had incurred a targeted deletion at the G9a gene locus. We isolated gDNA from transformed cells at the heart of several secondary colonies, and conducted PCR using primers designed to span the targeted region in the G9a gene for subsequent sequencing. We note that there was a very high DNA sequence failure rate (we obtained only 6 high-quality sequence reads from 48 gDNA templates), and we believe this high failure rate was due to the GC-rich regions flanking the target site. Nevertheless, none of the 6 sequences exhibited mutations/indels, suggesting that these transformed colonies were not the result of a clonal expansion from a homozygous-deleted founder. In retrospect, perhaps this is not surprising given the polyploidy in the cell line (ploidy=4; (Kilinc et al., 2014), where deletion of all G9a alleles would presumably be rare/improbable. Additional sequencing is required in order to more confidently assess genotypes within transformed OP6 colonies.

To gain further insight, we conducted immunofluorescence (IF) experiments on transformed colonies using an antibody against the G9a protein. We analyzed the expression of G9a in transformed colonies of various sizes, carefully examining each cell nucleus in each ~0.5 micron-thick confocal plane (the images shown in Fig. 2a-c are projections of all planes). We observed that G9a expression level is significantly reduced, but not abolished, in these colonies (Fig. 2a-c), as well as in dispersed cells from colonies (Fig. 2d), as compared to control populations where G9a expression is robust (Fig. 2e). This result
confirmed that colonies contain cells capable of expressing G9a, and therefore, cannot be clonal expansions of a homozygous G9a-deleted founder.

**Figure 2. G9a Expression is Heterogeneous within Transformed Colonies.**

G9a immunofluorescence is evident within approximately half (~46%) of the cells within transformed colonies of various sizes. (a) A typical small colony (33 cells containing ~67% G9a-positive cells; (b) a typical medium-sized colony (53 cells containing ~66% G9a-positive cells; (c) a portion of a large colony (190 cells containing ~24% G9a-positive cells. Colonies were disrupted and reseeded as a monolayer (panel d) in order to better compare G9a expression to a monolayer population of non-targeted control cells (panel e). Scale bars = 25μm. The histogram (panel f) summarizes the average percentages of positive cells in 9 sampled colonies of various sizes (S=small, M=medium, L=large), normalized to a non-targeted control in which ~100% of cells are G9a-positive. Error bars represent one standard deviation for multiple analyzed colonies (sample size, n, shown above bars).
Overall, ~half of the cells in the nine analyzed colonies are (faintly) positive for the G9a antigen, although there is high variability from colony to colony, especially among the smallest colonies (<50 cells, Fig. 2f). We note that there were substantially fewer (<25%) G9a-positive cells in the largest colony examined (~200 cells) as compared to the four medium-sized colonies examined (~50 cells), which consistently contained >60% positive cells. It is possible that G9a-negative cells are selectively advantaged (e.g., due to more rapid growth rates) such that they are more likely to dominate as colony size increases.

The heterogeneity in G9a expression within transformed colonies could arise by at least four kinds of mechanisms. First, colonies might have formed clonally from a heterozygous founder, and the observed protein expression heterogeneity might be due to significant reduction in expression (dosage) approaching sensitivity threshold levels for the antibody. Second, colonies might have begun with a heterozygous founder, and subsequent gene conversion/recombination events might have produced various genotypes within the population (e.g., a mixture of deleted and wild-type/heterozygous cells). Third, colonies might not be clonal, and surrounding wild-type cells might have been recruited into the colony. A problem with all three of these hypotheses is that they would predict the presence of mutant/deleted G9a template in our sequencing sample to account for transformation and loss of G9a expression in a large fraction of cells in the colony. Moreover, as noted previously, polyploidy in the OP6 cell line further complicates any genetic
explanation to account for significant loss of G9a expression. That is, it is difficult to imagine such a significant dosage effect in a heterozygous lineage, if only one (of say, four) G9a gene loci have been deleted, and it is difficult to imagine the generation of homozygous subpopulations (even via gene conversion/recombination) whereby each copy of the gene is deleted in G9a-positive cells. A caveat in this argument is the possibility that a dominant-negative allele could be disruptive, yet variable in penetrance, arising from merely one mutant copy per cell that might not be evident in the small sample of sequences so far obtained.

We speculate about a fourth plausible mechanism that could generate low and heterogeneous levels of G9a expression with high penetrance irrespective of gene copy number: a knockdown effect. CRISPR-mediated transcriptional repression has been documented in other systems – e.g., via steric hindrance to transcription elongation (Qi et al., 2013) (Zebec, Zink, Kerou, & Schleper, 2016) or via recruitment of heterochromatic factors (Anton & Bultmann, 2017; Lo & Qi, 2017). This epigenetic hypothesis, unlike the previous genetic deletion explanations, provides a parsimonious explanation for wild-type G9a genotypes in transformed cells (sequencing data), and predicts consistently reduced G9a expression levels as compared to untreated samples (see Fig. 2d and 2e; also see Fig. 4). Moreover, this epigenetic hypothesis provides a plausible explanation to account for potentially reducing expression from all/most alleles of a polyploidy cell, as well as a plausible explanation for inconsistent heritability during colony
expansion (i.e., to account for a mixture of G9a-positive and -negative cells),
since maintained silencing would presumably depend on the sustained robust
expression of the G9a cassette and sustained interference at all G9a loci, as well
as other intrinsic factors (e.g., stage of the cell cycle, etc.) (Hale et al., 2009;
Yuchen Liu et al., 2016; Price, Sampson, Ratner, Grakoui, & Weiss, 2015).
However, at this time, we must remain cautious about our conclusions regarding
G9a perturbation in this experiment – further sampling of gene sequences,
transcript levels, and protein expression in more enriched cell subpopulations
are required in order to confidently conclude the exact manner of G9a
perturbation we attained through our CRISPR-Cas9 guide RNA. This work is
ongoing in the lab and discussed further in Chapter 4.

3.3.C G9a Perturbation is Correlated with Reorganization of
Heterochromatin.

Careful examination of the G9a IF in Figure 2 revealed unusual DAPI (DNA)
staining within nuclei that lack G9a protein expression. Conspicuously, these
cells seem to lack chromocenters, which are DNA-dense pericentric
heterochromatin compartments distributed throughout the interior of the
nucleus (Probst & Almouzni, 2008). This observation is particularly noteworthy
since H3K9me3, the histone mark regulated by G9a, is crucial in the formation of
pericentric heterochromatin (Becker et al., 2016; Grewal & Klar, 1996). The left
panel in Figure 3A shows a pair of neighboring cells in a transformed population,
one of which expresses G9a protein and contains normal chromocenter nuclear organization, and the other lacking G9a expression and chromocenters (the DAPI stain is homogeneous in this latter cell, without the typical DAPI-dense compartments evident in the middle panel of Fig. 3A). We analyzed 9 colonies of various sizes, scoring both G9a-positive and G9a-negative cells within these colonies for the presence/absence of nuclear chromocenters. Overall across these nine colonies, approximately 60% of all (n=318) G9a-negative cells analyzed lack chromocenters; exceptions seemed to be associated with very large (>100 cell-) or very small (<20 cell-) colonies, noting that all 6 colonies between 20-60 cells showed a near-perfect correlation (>99% of G9a-negative were also chromocenter-negative in these medium-sized colonies). In contrast, <3% of all (n=281) G9a-positive cells analyzed exhibited abnormal heterochromatin organization. All 7 of these exceptional cells (G9a-positive, yet chromocenter-negative) were observed within one small colony, where G9a expression was evident but exhibited very weak signal that could represent false-positive staining. Across the other 8 colonies, each of the (n=274) G9a-positive cells had normal chromocenters. We also did not observe any abnormal chromocenter phenotypes in the scrambled, non-targeted control populations, nor in untreated OP6 cell populations. From these data, we speculate that G9a perturbation is contributing to the disruption of normal pericentric heterochromatin.
Chromocenters are typically enriched for H3K9me3-marked chromatin (associated with constitutively silenced DNA; Fig. 3B, right panel) and the HP1 protein (Fig. 3C, middle and right panels). HP1 contains a chromodomain that interacts with H3K9me3 to initiate DNA compaction and silencing (Verschure et al., 2005; Williams et al., 2003). We used antibodies against both H3K9me3 and HP1 antigen to investigate the impact of G9a perturbation on these heterochromatin components. We found that H3K9me3-marked heterochromatin had relocated to the nuclear periphery (Fig. 3B, middle panel), while the HP1 protein was absent from nuclei (Fig. 3C, left panel) within transformed cells lacking chromocenter structures. Together, our data suggests that G9a perturbation in OP6 cells is associated with heterochromatin reorganization, including loss of chromocenters, re-localization of H3K9me3 chromatin, and down-regulation of the HP1 heterochromatin regulator.
Figure 3. Correlation between G9a Depletion and Heterochromatin Reorganization.

A. Isolated cells in culture (left panel), with one nucleus exhibiting G9a staining (green) and normal chromocenter organization (DAPI, blue) and the other nucleus with no G9a staining and lacking chromocenters (DAPI, blue). Untreated OP6 cell nuclei always exhibit normal chromocenter patterning (DAPI, middle panel). Scale bar = 5μm. The histogram (right panel) shows a correlation between loss of G9a expression (white bars) and the percentage of cell nuclei lacking chromocenters, as compared to G9-positive cells (black bars), for colonies of various sizes (S=small, M=medium, L=large). B. The absence of chromocenter organization is correlated with the movement of H3K9me3 staining to the nuclear periphery (left, middle panels), as compared to a typical OP6 cell where H3K9me3 staining is associated most strongly with central chromocenters (right panel). Scale bar = 1μm. C. HP1 protein is typically enriched within chromocenters (middle, right panels), however, the HP1 protein is not detected within nuclei of transformed OP6 cells lacking chromocenters (left panel). Scale bar = 5μm.
3.3.D Gene Expression Analysis in Transformed OP6 Cell Populations

We isolated RNA and produced cDNA from cells extracted from central portions of several secondary colonies (i.e., to enrich for the rounded, rapidly-dividing subpopulation of transformed cells). Initial qPCR experiments were conducted to survey a range of pluripotent, developmental, and cell-cycle factors (Fig. 4). As anticipated, we observed a significant knockdown of G9a expression (~45-fold reduction) in these populations as compared to scrambled non-targeted control populations earlier. Eight of the 16 other genes tested showed only slightly increased/decreased expression levels (<two-fold). These genes include *Nestin* (a marker for cancer stem cells or neuronal stem cells), *BMP4* (bone differentiation factor), *LSD1* (a histone demethylase), *GFAP* (a mitotic factor commonly found in astrocytes progenitors), *MAP5* (early neuronal marker), *SMARCA4* (chromatin remodeler), *Ascl1* (neurogenesis factor), and *CCNE2* (cyclin).

The remaining eight genes exhibited >two-fold difference in the treated cell population; all but one of these exhibited significant up-regulation. The most significant difference was for the *FGF2* gene (~30-35 fold increase), a growth factor associated with stem cells that is important for maintaining undifferentiated states (De Los Angeles, Loh, Tesar, & Daley, 2012; Greber et al., 2010). We also note significant increased expression of *KLF4* (~5-fold increase), *LGR5* (~5-fold increase), and *Nanog* (~3-fold increase), all of which are stem markers that likewise function to maintain undifferentiated states (Hochedlinger
Thus, all four of the putative stem cell markers tested in this panel of 17 genes showed a significant increase in expression relative to control populations. Additionally, the *TERT* (telomerase) gene, also associated with stem cell maintenance (Marion et al., 2009), as well as oncogenic transformation (reviewed in Hanahan & Weinberg, 2011; Stewart et al., 2002), was expressed at higher levels in treated populations (~5-fold increase). The other three genes exhibiting a significant change in expression in the treated populations are associated with cell cycle control: the tumor suppressor *Ink4a/ARF* (~25-fold increase), the growth-promoting transcription factor *Stat3* (~2.5-fold increase), and the *CyclinD1* protein (~3-fold decrease) (H. Li et al., 2009; Utikal et al., 2009).

To summarize these qPCR results: (i) we observed only modest or insignificant changes in the expression of various developmental regulatory genes tested (*BMP4, LSD1, MAP5, SMARCA4, Ascl1*); (ii) we observed ambiguous changes in the expression of various cell cycle- and oncogenic-associated genes tested (*Ink4A/ARF* and *TERT* up-regulated; *CyclinD1* down-regulated; *GFAP*, *CCNE2*, and *Nestin* not significantly different); and (iii) we observed significant up-regulation of all stem cell-associated genes tested (*FGF2, KLF4, LGR5, Nanog*, and *TERT*).
Figure 4. Gene Expression Analysis in Treated and Untreated OP6 Populations for a Panel of Pluripotent and Cell Progression Gene Markers. Histogram showing the average whole population expression levels in two biological replica experiments relative to the non-targeted control for 17 genes of interest. Each replica is normalized to an internal control gene (actin), and plotted on a log-2 scale to report average fold increase (and decrease) in expression in the two treated samples versus a non-targeted control sample. The dotted rectangle represents a two-fold cutoff threshold within which increases/decreases are likely not significant. As anticipated, G9a expression is significantly reduced in treated populations as compared to control populations (~45-fold).

Together, our qPCR data support the hypothesis that G9a perturbation might have induced de-differentiation of OP6 cells into a stem-like/pluripotent state that develops colonies resembling embryoid bodies. However, we must again remain cautious given the small number of experiments conducted and the range of marker genes explored. Experiments are ongoing in the lab to enrich various cell populations at different stages of transformation and for colony size, as well as expand the range of marker genes, in order to build upon these promising preliminary data.
3.3.E Immunofluorescence for Stem Marker Proteins.

We further investigated markers commonly utilized in the characterization of stem cell populations using immunohistochemistry. We conducted immunofluorescence (IF) experiments on transformed secondary OP6 colonies using available antibodies against FGF5, Sox2, Tra1-81, and Ssea1. The FGF5 is a secretory protein and it is expressed during late stem cell stages associated with early differentiation into primitive ectoderm-like cell types (Trott & Martinez Arias, 2013). Sox2 (transcription factor), Tra1-81 (cell surface glycoprotein), and Ssea1 (fucosyltransferase protein) are commonly used to distinguish pluripotent stem cells from differentiated lineages (Hochedlinger & Jaenisch, 2015; Jaenisch & Young, 2008). Transformed colonies expressed FGF5 predominantly in the outer cells of colonies (Fig. 5a-b), and in particular, within the monolayer of cells with flattened morphology containing elongated nuclei at the peripheral regions of colonies (as described in Fig. 1C). In contrast, Sox2 was expressed more within the interior cells of colonies (Fig. 5c), and expression was especially robust in a small handful (2-3 cells per colony) of the most centrally located cells of the colony. Two-color IF shows mostly non-overlapping expression of these two markers, including ~two bright green Sox2-positive cells in the interior (Fig. 5d). These results suggest that the internal portions of transformed colonies contained more early stem-like cells (with high Sox2/low FGF5 expression), whereas the more peripheral portions of these colonies contained cells at later
stem-like stages en route to differentiating (with high FGF5/low Sox2 expression).

**Figure 5. Transformed Colonies Express Pluripotent Marker Genes, as well as Marker Genes from the Three Canonical Germ Layers.**

Representative immunofluorescence images are shown for various pluripotent marker proteins up-regulated within transformed colonies: FGF5 (panel a; interior of colony, panel b; periphery of colony; scale bars = 25mm), Sox2 (panel c; scale bar = 25 mm), FGF5+Sox2 (panel d; scale bar = 15μm), Tra1-81 (panel e; scale bar = 15μm), and Ssea1 (panel f; scale bar = 15μm).

In addition, transformed colonies, but not untreated or scrambled non-target controls were positive for the other two stem-markers tested (Tra1-81,
Ssea1; Fig. 5e-f), although unlike FGF5/Sox2, there was no apparent patterning in the distribution of these two proteins within colonies. These two markers were even positive in the peripheral flattened cells that tended to be Sox2 negative/FGF5 positive. We note that isolated cells in the culture (i.e., that were not associated with a colony) were negative for these markers, suggesting that expression of these proteins might be dependent on signals present in the immediate vicinity of the colony or that expression persists only in cells that had recently arisen from the colony lineage.

Together, our IF data support the hypothesis that G9a perturbation might have induced characteristic Sox2 and FGF5 expression into de-differentiated of OP6 cells that form colonies that resemble ES cells and behave similar to embryoid bodies. We again note the preliminary nature of these experiments, and work is ongoing in the lab to more extensively compare stem marker gene expression in transformed OP6 colonies at various stages of development/growth versus untreated controls, as well as positive control ES cell populations (Joyce Noble, Ghazia Abbas, personal communication; also see Chapter 4 that discusses these future directions more fully).

### 3.3.3 Initial Observations on the Differentiation Potential of Transformed OP6 Cells.

While we have not yet studied the developmental potential of the transformed OP6 colonies (some preliminary experiments and future directions are
described in Chapter 4), we did make one striking observation that suggests that these transformed colonies are capable of self-organization, just like embryoid bodies in vivo. When we allowed the colonies to grow in undefined media for ~3 weeks, we observed that several colonies began forming differentiated structures (Figure 6). These structures contained defined boundaries and apparent invaginations between groups of cells (Figure 6a), outward extending micro-protrusions ("leaf-like" structures, Figure 6b), and oblong-shaped colonies that appeared to be developing complex layering (Figure 6c). In rare cases, colonies grew long protrusions and apparently segmented extensions (Figure 6d). In the particular image shown (Figure 6d), the extending "braid" is approximately 5 mm in length, extending from a large ~500-cell colony with a diameter of 5-8 mm; the "braid" itself seems to be comprised of a dense mass of cells without any apparent internal structure. More remarkably, occasionally we observed what appeared to be macroscopic 5 mm, intricately developed "worm-like" structures floating in various culture dishes (Figure 6e). From a total of ~50 T100 culture flasks, each containing ~40-50 tertiary colonies, we found an object resembling the one depicted in Figure 5e in seven separate dishes. These objects were never observed in parallel untreated or scrambled non-targeted control cultures. These objects exhibited exquisite complexity, including reiterated striations/compartments, complex layering, and multiple cell types, including what appear to be ciliated cells on one surface (Figure 6f).
Figure 6. Transformed Colonies Differentiate and Self-organize into Putative Developmental Structures.

(a) Apparent invagination and delineating boundary structure between two masses of cells. (b) Micro-protrusions of various sizes at the edges of some colonies are evident after several days in culture. (c) Oblong-shaped colony with a pair of protrusions at one end. (d) A much longer extension emerging from a larger colony. (e, f) Macroscopic structures form in some cultures after 3 weeks that exhibit complex cellular organization. Scale bars = 50μm.
We note that the majority of the tertiary colonies within the ~50 culture flasks did not “develop” such structures – they simply grew in size, requiring eventual passaging in order to maintain. The fact that such a small percentage of colonies exhibited this organizing potential could be explained by stochastic events during transformation – e.g., the exact dosage of G9a expression/degree of perturbation (and resulting differences in gene expression), the generation of a suitable ratio and locations of heterogeneous cell types, and/or the local environmental concentrations of cytokines (including those in the growth media), etc. It is also possible that these G9a-transformed colonies, if they are truly embryoid-like, are potentially capable of organizing themselves along multiple kinds of pathways; e.g., in various growth media conditions. We also note that all of the secondary and tertiary colonies reported in this study, ultimately derived from only a single primary colony from the initial infection, so greater sampling of primary colonies might reveal alternative, much broader developmental potentials. Much more in-depth characterization is needed in order to understand the progression apparent in these provocative cases, including de-differentiation/transformation mechanisms, endogenous activation of pluripotency networks, upregulation of various germ layer markers including spatial and temporal patterning, and more complete characterization of apparent developmental structures/organoids and differentiated cell types that form; each of these future experimental goals is discussed in Chapter 4.
3.4 Discussion

Olfactory neurogenesis is thought to be a continuous and unidirectional developmental process, supported from a population of basal stem cells (GBCs, HBCs) that differentiate to olfactory sensory neurons, as well as other supporting cell types in olfactory tissue (Krolewski et al., 2013; Packard, Giel-Moloney, et al., 2011; Schwob et al., 2017). This study describes a set of unexpected outcomes with what began as a routine experiment to perturb a chromatin regulator (G9a) known to be involved with the proper expression of odorant receptor (OR) genes in a cell line derived from an immature neuron of the developing olfactory lineage (Illing et al., 2002; Lyons & Lomvardas, 2014; Magklara et al., 2011).

We made several surprising observations arising from G9a perturbation: transformation of cell morphology and growth characteristics, colony formation, heterochromatin reorganization, colony differentiation, and apparent self-organized “organoid” development. A parsimonious explanation to account for these observations is that G9a perturbation might have reduced epigenetic barriers that permitted de-differentiation; e.g., via robust activation of the early transcription factors, Oct4-Sox2-nanog, has been shown previously to induce pluripotent stem cells from differentiated cell types (Jaenisch & Young, 2008; Takahashi & Yamanaka, 2006; Yu et al., 2007). The Oct4-Sox2-nanog transcriptional network is autoregulatory and its activation is sufficient to suppress lineage differentiation and activate proliferation genes (Assou et al.,...
2009; Boyer et al., 2005; Reviewed in Dejosez & Zwaka, 2012; reviewed in Kashyap et al., 2009; Sun et al., 2009). We speculate that the activation of these key transcriptional regulators might be a sufficient explanation to account for various stem-like properties, including alteration to a more rounded cell morphology (Fig. 1), a shortened cell cycle and embryoid-like colony formation (Fig. 1), pigmentation and alkaline phosphatase staining (Fig. 1), and expression of various pluripotent markers (Fig. 4). However, the levels of stem-marker activations in transformed OP6 cells do not seem to be as robust as would be expected for an embryonic stem cell (reviewed in Hochedlinger and Jaenisch, 2015), therefore, it is unclear at this time whether the observed changes in cell behavior and gene expression towards a more stem-like quality represent a true de-differentiation or rather towards a more bivalent or uncommitted state. Additional characterization of the individual cells and colonies of the transformed population is necessary to clarify the extent to which G9a-perturbed cells have de-differentiated. For example, we will next want to compare/correlate the loss of neural marker genes with the gain of stem marker genes at the resolution of single cells (see Chapter 4).

How G9a perturbation might result in de-repression of key stem regulators is also not clear. G9a catalyzes the acquisition of repressive H3K9 mono- (H3K9me1) and di- (H3K9me2) methylation states (Tachibana et al., 2001), whereas other enzymes, such as SETDB1 and SUV39H1/2, catalyze the acquisition of H3K9me3 from mono- and di-methylated substrates (Black, Van
Generally, G9a and other transcriptional regulators exit the nucleus during M phase of the cell cycle, and consequently, repressive marks such as H3K9 methylation are lost (Ganier et al., 2011; Halley-Stott, Jullien, Pasque, & Gurdon, 2014; Y. Liu et al., 2017). Thus, there might be a requirement for G9a to reestablish these repressive marks after cell division, whereas the loss of G9a may result in a failure to maintain this epigenetic memory.

For example, during stem cell differentiation, the promoter of the stem-inducing regulator OCT4 recruits a G9a complex containing histone deacetylase to remove H3K9 acetylation marks thereby enabling G9a to methylate these residues en route to silencing the locus (Feldman et al., 2006). Moreover, cell reprogramming is more efficient with the inhibition of H3K9 methyltransferases (e.g., G9a, as well as SETDB1 and SUV39H1/2) and less efficient with the inhibition of H3K9 demethylases (e.g., Kdm3a, 3b, 4a) that regulate H3K9me3 levels at downstream target genes of stem regulators, such as OCT4, Sox2, KLF4, and C-Myc stem regulators (J. Chen et al., 2013; Das et al., 2014; Gladych, Andrzejewska, Oleksiewicz, & Estecio, 2015; Onder et al., 2012; Soufi, Donahue, & Zaret, 2012; Sridharan et al., 2013). Therefore, perturbation of G9a may directly interfere with the ongoing silencing of the stem regulatory network.

On the other hand, perturbation of G9a might indirectly induce activation of stem factors. Key stem regulators, such as Oct4, Sox2, and nanog, are typically described as possessing a “poised” epigenetic state, with both repressive
H3K27me3 and activating H3K4me3 marks present on their cognate promoters (Bernstein et al., 2006a; Harikumar & Meshorer, 2015; Shema et al., 2016). Differentiation along a lineage is in part enabled by the silencing of these stem factors, which is associated with an equilibrium shift towards increased H3K27me3 and decreased H3K4me3 on their promoters (Bernstein et al., 2006a, Rotem et al., 2015; Shema et al., 2016). G9a is not directly involved with either the H3K27 or H3K4 methylation pathways, and therefore, stem marker activation in this context might involve an indirect G9a-mediated de-repression of other enzymes that regulate these marks.

Along these lines, G9a functions within multi-protein complexes that serve to target its repressive activity to various gene promoters, as well as to deliver downstream factors, such as HP1 and de novo DNA methyl transferases (DNMT3a/b), to further stabilize gene silencing (Becker et al., 2016; Verschure et al., 2005). We speculate that the stability of these repressor complexes and their components might have broader epigenetic consequences beyond the H3K9 methylation landscape. The fact that HP1 protein is absent from G9a-depleted nuclei (Fig. 2) exemplifies a potential indirect effect. It is possible that the loss of G9a, a binding partner of HP1 (Sampath et al., 2007), results in the orphaned HP1 becoming destabilized, perhaps as a mechanism to ensure proper balance of the two factors within the nucleus. Regardless of the mechanism of HP1 depletion upon G9a perturbation, the down-regulation of the former protein is likely to have broad impacts on the epigenetic landscape beyond just H3K9 methylation.
targets. For example, the loss of HP1 has been shown previously to have a very significant impact on the overall nuclear organization within the olfactory cell lineage in particular, including the loss of chromocenter structures and the movement of constitutive heterochromatin to the nuclear periphery (Le Gros et al., 2016a), just as we have observed here (Fig. 2). It is possible that this global heterochromatin reorganization might indirectly lead to repositioning of key stem regulators, which would normally be silenced at the nuclear periphery (Jost et al., 2011), or more globally interfere with topological associated domains (TADs) (de Wit et al., 2013; Gorkin, Leung, & Ren, 2014) otherwise contributing to the silencing of numerous genes.

The G9a gene has previously been conditionally deleted within the mouse olfactory lineage (Lyons et al., 2014). While the authors of this study did not investigate de-differentiation effects in these animals, there is no report or obvious evidence to indicate aggressive colony formation in these tissues, as might be predicted if the mutant cells had reverted to embryonic stem-like qualities similar to those described here. We speculate that the expression of the large T-antigen, a viral oncogene that interferes with the retinoblastoma (Rb) cell cycle entry checkpoint protein (Weinberg, 1995), might synergize with G9a perturbation in the OP6 cell context so that both epigenetic and cell cycling barriers are sufficiently reduced to increase the probability of transformation. For example, one study demonstrated that deactivation of the Rb protein is an essential (but not sufficient) step in inducing differentiated myotubes to
efficiently de-differentiate (Schneider, Gu, Zhu, Mahdavi, & Nadal-Ginard, 1994); additionally, acquisition of pluripotency in this case depends on viral-induced chromatin reorganization (Eberharter & Becker, 2002; Endo & Nadal-Ginard, 1998; Monier, Armas, Etteldorf, Ghazal, & Sullivan, 2000). A general theme in various studies is that chromatin reorganization and cell cycle reentry are coupled events during the de-differentiation process (also see (Grafi, 2004; Jopling et al., 2011)). Moreover, several studies illustrate the general requirement for a shortened G1 phase of the cell cycle to maintain pluripotency, and a mitotic advantage (S-phase > G1/G0-phase) in facilitating reprogramming efficiency (Halley-Stott et al., 2014; Y. Ma, Kanakousaki, & Buttitta, 2015). The fact that we observed decreased G1 cyclin gene (cyclinD1) expression (Fig. 4) is consistent with a shortened G1 phase of the cell cycle in transformed OP6 cells (Wianny et al., 1998). Therefore, the presence of the large T-antigen in G9a-pertubed OP6 cells is likely contributing to the high efficiency of transformation by reducing barriers for reentry into S-phase. In Chapter 4, I discuss some future experiments that investigate the dependency on the large T-antigen for establishment and maintenance of stem-like qualities in transformed OP6 cell populations.

We also consider the unique attributes of the olfactory system that might render cells of this lineage particularly well situated for de-, re-, or trans-differentiation in order to respond appropriately to tissue depletion or injury. Cellular plasticity depends on reduced barriers for re-entry into the cell cycle
and re-acquisition of pluripotent factors (Grafi, 2004; Jopling et al., 2011). A recent study demonstrated trans-differentiation in the olfactory neuronal lineage explicitly within the mitotically active transit amplifying population (Lin et al., 2017). These events are triggered by injury or cell loss, however, can be induced with perturbation of a different heterochromatin factor, EZH2, that regulates H3K27 methylation states (Cao & Zhang, 2004). De-differentiation is accompanied by increased Sox2 expression, and re-differentiation occurs to produce non-neuronal olfactory cell types (Lin et al., 2017). Presumably then, this process is regulated within the olfactory niche for efficient trans-differentiation to alternative olfactory cell lineages (e.g., via reversion through the GBC olfactory stem cell), as opposed to acquisition of an even more naïve stem cell pluripotency (e.g., reversion back to an embryonic stem cell). Nevertheless, reduction of epigenetic barriers in an OP6 cell culture context absent of environmental cues, as well as with enhancement of cell-cycle stimulating factors as noted previously, could result in more complete reversion to a more naïve stem-like state. A major aim for future studies is to more fully investigate the pluripotency of transformed OP6 cells, including whether the apparent stem-like reversion is incomplete with lingering memory or bias for the OSN or neuronal lineages more generally (see Chapter 4).

Finally, we observed that transformed OP6 colonies seemed to re-differentiate over time. The first evidence of this is the heterogeneity of the colonies themselves. Despite the fact that colonies are presumably clonal, or at
least comprised of cells at its core that exhibit ~uniform GFP expression (Fig. 1), we nevertheless observed subpopulations that differed in cell morphology (Fig. 1) and gene expression (Figs. 2 and 4). As noted, CRISPR-Cas treatment apparently may not have resulted in an edited G9a locus and fully penetrant knockout, rather, might have produced a knock-down effect, as has been described previously with this method (Zebec et al., 2016), leading to heterogeneous G9a expression (Fig. 2). This might have been quite fortuitous, since subsequent re-differentiation within subpopulations of the colony might indeed depend on the reestablishment of the very G9a-dependent epigenetic barriers that had been perturbed to enable de-differentiation in the first place.

At a gross level, we observed a pattern within colonies that suggests a central stem-like core and a more differentiated outer periphery (Figure 1). The cells at the core of transformed colonies exhibited robust GFP expression from the CRISPR-Cas cassette (Figure 1), minimum G9a expression and loss of chromocenters (Figure 2 and Figure 3), compacted cell morphologies and aggressive 3D growth characteristics (Figure 1), as well as maximum stem-marker features (pigmentation, alkaline phosphatase, Sox 2, etc.) (Figures 1 and 4). In contrast, larger colonies contained a peripheral halo of flattened, attached cells with more differentiated morphologies; e.g., more reminiscent of immature neurons, epithelial cells, or fibroblast cells (Figure 1). The outer portions of colonies also exhibited more cells with normal chromocenter organization (Figure 2) as might be predicted for a more advanced, pre-differentiation stage.
(Trott & Martinez Arias, 2013). Further study of various differentiation markers along a time-course during colony development is warranted in order to better characterize any stereotypic profiles, such as the expression of canonical germ-layer marker genes (see Chapter 4).

Despite the apparent cellular heterogeneity within the growing colonies, perhaps the truest test of organized patterning in a differentiating body of cells is the capability to achieve non-random structure. The development of colonies into structures shown in Figure 5 was reproducible (observed in several independent dishes) and apparently progressive in nature (with increased complexity over time). While much work needs to be done to characterize the developmental potential of G9a-perturbed colonies – i.e., the exact cell types and apparent tissues/structures present, as well as other alternative developmental outcomes that might arise with greater sampling or when colonies are grown in different media environments – these preliminary observations further support the hypothesis that G9a perturbation has induced pluripotency from otherwise committed cells along the olfactory neuronal lineage.

In conclusion, we consider this to be a preliminary report of a potentially novel system for investigating induced pluripotency. Several questions remain for more thorough characterization of this system. Does transformation depend on the large-T antigen? To what extent are lineage or stem marker proteins expressed in a mutually exclusive way? Can we mimic these transformations using a G9a inhibitor drug, or knockout paradigm, where perturbation would be
more uniform and penetrant within the entire cell population? To what extent is the development of complex structures a stereotyped progression that can be characterized with specific stages and gene expression profiles? Can we induce transformed colonies to differentiate along various other lineages, potentially producing different structures, using more defined growth media? These and other future directions aim to establish a cell culture paradigm for the study and application of trans-differentiation protocols, as well as to better understand the epigenetic events that enable cell reprogramming in this system are discussed as future directions in Chapter 4 of this thesis.

References


260


complexes and stem cell microRNAs. *Stem cells and development, 18*(7), 1093-1108.


Gradual Chromatin Compaction and Reorganization during Neurogenesis 
In vivo. Cell Reports, 17(8), 2125-2136. doi:10.1016/j.celrep.2016.10.060

Functional Maturation of Olfactory Sensory Neurons and Development of 
Mother Preference. Journal of Neuroscience, 31(8), 2974-2982. 
doi:10.1523/jneurosci.5067-10.2011


Lee, K., & Seo, P. J. (2018). Dynamic Epigenetic Changes during Plant 
doi:10.1016/j.tplants.2017.11.009

nearial stem cells to tissue maintenance and regeneration. Nature Neuroscience, 10(6), 720-726. doi:10.1038/nn1882

Li, E. (2002). Chromatin modification and epigenetic reprogramming in 
mammalian development. Nature Reviews Genetics, 3(9), 662-673. 
doi:10.1038/nrg887

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Canamero, M.,...Serrano, M. 
(2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. 
Nature, 460(7259), 1136-1139.

Dynamic distribution of linker histone H1.5 in cellular differentiation. 
PLoS Genet, 8(8), e1002879. doi:10.1371/journal.pgen.1002879


doi:10.1038/s41556-018-0067-6

Lin, B., Coleman, J. H., Peterson, J. N., Zunitch, M. J., Jang, W., Herrick, D. B., & 
Schwob, J. E. (2017). Injury Induces Endogenous Reprogramming and 
Dedifferentiation of Neuronal Progenitors to Multipotency. Cell Stem Cell, 
21(6), 761-774 e765. doi:10.1016/j.stem.2017.09.008

Remodeling of the Endogenous Oct4 or Sox2 Locus Enables 
Reprogramming to Pluripotency. Cell Stem Cell, 22(2), 252-+. 
doi:10.1016/j.stem.2017.12.001


Liu, Y., Pelham-Webb, B., Di Giammartino, D. C., Li, J., Kim, D., Kita, K.,... 
Apostolou, E. (2017). Widespread Mitotic Bookmarking by Histone Marks 
and Transcription Factors in Pluripotent Stem Cells. Cell Reports, 19(7), 
1283-1293. doi:10.1016/j.celrep.2017.04.067

265


Pathak, N., Johnson, P., Getman, M., & Lane, R. P. (2009a). Odorant receptor (OR) gene choice is biased and non-clonal in two olfactory placode cell lines, and OR RNA is nuclear prior to differentiation of these lines. *Journal of Neurochemistry*, 108(2), 486-497. doi:10.1111/j.1471-4159.2008.05780.x

Pathak, N., Johnson, P., Getman, M., & Lane, R. P. (2009b). Odorant receptor (OR) gene choice is biased and non-clonal in two olfactory placode cell lines, and OR RNA is nuclear prior to differentiation of these lines. *Journal of Neurochemistry*, 108(2), 486-497. doi:10.1111/j.1471-4159.2008.05780.x


Chapter 4: Future Investigations of OR regulation and OSN de-differentiation

4.1 Introduction

Each olfactory sensory neuron (OSN) normally expresses only one allele (monoallelic) of one OR gene (monogenic) in order to give each OSN a specific odorant-binding capability (Buck and Axel, 1991). In Chapter 2 of my thesis work, I showed that OR regulation is partially dependent on the LSD1 protein (Vyas et al., 2017). We knocked down LSD1 expression in the OP6 cell line and showed that OR re-selections still occurred during culturing, indicating that LSD1 is not required for de novo OR activations in this system. We also showed that OR representation in OP6 cultures was not apparently impacted by LSD1 depletion, suggesting that LSD1 is probably not functioning to influence OR eligibility for selection (e.g., by regulating accessibility at certain OR clusters). Unexpectedly, given the favored model in the field in which LSD1 had been regarded as a putative OR “activator”, we observed that LSD1 depletion caused excess OR activations per cell, suggesting that LSD1 functions as an OR “repressor” that helps ensure only one OR gene is expressed per cell. As might be predicted by this hypothesis, we also showed that prolonged culturing of LSD1-

---

1 Portions of the text in this Chapter are co-authored with R. P. Lane as part of an upcoming grant submission
depleted OP6 cells exhibited gradual and systematic increase of the activating H3K4 methyl mark at OR loci, consistent with a role for LSD1 in maintaining low H3K4 methylation levels, presumably as a mechanism to ensure inactivity of competing ORs. Based on these observations, we have proposed an iterative model in which dynamic addition of H3K9 methylation (by enzymes that add these marks at ORs; e.g., G9a and GLP) and removal of H3K4 methylation (by enzymes that remove these marks at ORs; e.g., LSD1) preserve OR monogenicity by suppressing competing OR genes. According to our model, the disruption of either enzyme (G9a that adds H3K9 or LSD1 that removes H3K4) might shift the equilibrium such that H3K9 methylation becomes less universal and H3K4 methylation becomes less rare, thereby increasing the probability of multiple OR activation events per cell.

To further test this model, we performed a G9a knockdown in OP6 cells (Chapter 3). We had intended to also investigate a G9a-LSD1 double knockdown (predicting an enhanced phenotype; see below), as well as an LSD1 over-expression experiment in G9a depleted cells (predicting a rescue phenotype; see below). However, the surprising results we obtained in the G9a knockdown experiment precluded these studies, and instead shifted our focus to investigating a larger role for G9a in the context of OSN development. As discussed in Chapter 3, G9a CRISPR treatment resulted in morphological transformation of OP6 cells into stem cell-like colonies, global heterochromatin reorganization, upregulation of stem-cell markers, and rare but reproducible
differentiation of stem-like colonies into what appear to be multicellular, highly organized structures (e.g., organoids).

The purpose of this Chapter is to provide future directions and additional unpublished preliminary data concerning three major lines of inquiry. First, there remain several unanswered questions about OP6 cell transformation, including the mechanism by which the CRISPR-Cas9 cassettes have perturbed G9a expression, whether G9a-pertubed OP6 cells have lost their OSN identity, to what extent G9a perturbation has activated canonical stem cell networks (including endogenous cell cycling networks), how chromatin reorganization in transformed OP6 cells has impacted gene expression more generally, and the extent to which transformed OP6 cells exhibit tumorigenic properties (Section 4.1). Second, there remain numerous unanswered questions about the pluripotency and developmental potential of transformed OP6 cells, including whether OP6 colonies can differentiate along multiple cell lineages (e.g., lineages of the three major germ-layers), the frequency and variability of spontaneous development (e.g., more careful analyses of the observed developmental structures), and the competency of transformed OP6 colonies to be induced along differentiation pathways both in vitro and in vivo (Section 4.2). And third, the momentum gained with identifying a novel OR phenotype following LSD1 perturbation (Chapter 2) inspires a number of follow-up experiments that might clarify a role for chromatin “writers” in the establishment of monogenic OR expression, and more specifically, whether G9a and LSD1 collaborate in the
silencing of competing OR genes (Section 4.3). Finally, I end with a brief discussion about the connection of what at first glance appears to be two very different experimental results described in this thesis: an OR regulation story (Chapter 2) and an OSN development story (Chapter 3), with further exploration of the idea that the establishment of mutually exclusive OR expression and the differentiation of mature OSNs are possibly co-dependent events during development, governed by a common set of epigenetic factors.

4.1.A What is the Mechanism of CRISPR-Cas9 Mediated G9a Perturbation?

As discussed in Chapter 3, we targeted exon 8 of the G9a locus using a CRISPR-Cas9-based genome editing approach with the ultimate goal of depleting the G9a protein in OP6 cells by introducing a frame-shift mutation. Although our sequencing sample is not yet deep enough to confidently dismiss the hypothesis, we find no evidence of a genetic deletion/mutation in sequenced DNA from transformed OP6 colonies. As discussed previously, our data seems most consistent with an epigenetic effect; e.g., the gRNA might cause reduced transcriptional output, a hypothesis that consistent with both the observed heterogeneity within colonies and the reduced (but not abolished) level of G9a transcript (qPCR) and protein (immunofluorescence) in transformed colonies.

To more confidently dismiss the genetic deletion/mutation hypothesis, we would need to sequence individual GFP-sorted cells at various time points; e.g., 24 hours, 48 hours, and 72 hours post-CRISPR treatment. In addition, GFP-
sorted OP6 cells at these early time points can be used to perform G9a CRISPR target site-specific PCR followed by restriction digestion using the surveyor / T7 endonucleases to assess genetic heterogeneity introduced by the CRISPR treatment. The surveyor enzyme cuts any unmatched DNA (i.e., mis-matches generated due to indel mutations) within the PCR amplicons, thus would reveal whether new alleles have been introduced by the CRISPR treatment. Both strategies (sequencing and surveyor) at these early time points (e.g. prior to colony growth and differentiation) would preclude possible recombination events that might otherwise correct a mutated G9a allele during prolong culturing. Importantly, however, we want to apply these methods on sorted GFP-positive cells in order to more directly correlate genotyping with expression of the CRISPR cassette, as opposed to working with unsorted populations that might contain numerous wild-type (non-perturbed) cells and alleles. For example, it is possible that primary colony formation may require an assembly of both genetically perturbed and wild-type cells in order to establish necessary cell-cell communications that stabilize the transformed state, a hypothesis that might be supported by the observed heterogeneity in G9a expression, GFP expression, marker gene expression and cell morphology in colonies. It is also possible that a disruptive indel in Exon 8 might produce a dominant allele (truncated protein with a dominant-negative effect) that would require a deep sequencing sample to detect, given both the apparent heterogeneity in the cell population and the polyploidy of the cell line.
In order to verify that the gRNA is, in fact, targeting the G9a locus, we would conduct a two-color immuno-DNA FISH experiment to show co-localization of a G9a DNA FISH probe with a dCas9 antibody at the G9a locus in GFP+ colonies. This experimental result would also help dismiss off-target effects as the cause of transformation. In addition, we have a second gRNA in house against Exon 2 that would be informative on these questions. If this other gRNA produces the same phenotype, it would again argue against off-target effects; it would also be interesting if this second gRNA very clearly causes a genetic change at the G9a locus, and whether the difference between a genetic (Exon 2) versus epigenetic (Exon 8) perturbation results in any phenotypic differences. Along these lines, we are also interested in whether post-transcriptional perturbation of G9a function through drug inhibition mimics the original Exon-8 perturbation. Each of these experiments aim to validate that transformation of OP6 cells is, in fact, caused by G9a perturbation.

It is possible that the gRNA is causing a post-transcriptional effect via its complementarity to G9a mRNA. For example, it is possible that the G9a gRNA might engage with G9a mRNAs to alter its stability or translational potential. Distinguishing a reduction of mRNA production from an increase in mRNA instability is not trivial, without making direct measurements of the rates of mRNA transcription. However, if the G9a gRNA is having an impact on translational efficiency, one prediction would be a disconnect between mRNA and protein production – e.g., we might observe far less G9a protein (i.e., via
Western blot quantitation) per mRNA (i.e., via RT-qPCR quantitation) in GFP-positive versus GFP-negative cells.

The hypothesis we currently favor is a CRISPR interference (CRISPRi) mechanism at the G9a locus, whereby the gRNA is causing changes in G9a chromatin states, thus, efficiency of transcriptional expression (Larson et al., 2013; Qi et al., 2013; Marraffini et al., 2011; Zebec et al., 2016). Predictions of this hypothesis include the reduction (but not loss) of G9a transcripts that are otherwise intact (i.e., without mutation). We have conducted preliminary qPCR to demonstrate reduction of G9a cDNA in transformed OP6 populations, however, we would like to do additional RNA analyses on cell populations to correlate robust G9a repression with robust GFP expression and/or robust gain in stem marker expression (e.g., Sox2). Another prediction of this hypothesis is a change in chromatin states around the G9a locus, especially in cell subpopulations that exhibit high GFP expression. To test this prediction, we would conduct ChIP experiments to investigate a correlation between gRNA expression and acquisition of repressive chromatin marks. Together, these experiments would clarify an important underlying assumption about the root cause of the observed phenotypes: Does the G9a CRISPR cassette target G9a DNA (mutation), RNA (production or stability), or protein (production or stability) or is the observed phenotype due to an off-target effect?
4.1.B Has G9a Perturbation caused Chromatin Reorganization?

Chromatin reorganization during reprogramming serves two major purposes: 1) to establish repressive epigenetic marks on tissue-specific genes to silence differentiation-specific programs; 2) to erase repressive epigenetic marks on pluripotency genes in order to activate stem cell networks (Takahashi and Yamanaka, 2006; Werning et al., 2007; Okita et al., 2007; Mikkelsen et al., 2008).

Along with these local and gene-specific chromatin changes, the global constitutive heterochromatin reorganizes during the reprogramming process (Fussner et al., 2011). These global chromatin changes are measured by the density of heterochromatin domains, number of chromocenter foci, redistribution of heterochromatin (measured by H3K9me3 and DNA methylation immunofluorescence), loss of HP1, and the formation of 10nm chromatin fibers (Peters et al., 2001; Guenatri et al., 2004; Tessadori et al., 2007; Rego et al., 2008; Ahmed et al., 2009). Moreover, in differentiated and partially reprogramed cells, the constitutive heterochromatin is marked by H3K9me3 and H4K20me3 and is highly compartmentalized into densely packed chromocenters, whereas in ES cells heterochromatin is more broadly dispersed within nuclei (Efroni et al., 2008; Hiratani et al., 2009; Ahmed et al., 2010). As discussed in Chapter 3, control OP6 cells contain densely packed heterochromatin organized into DAPI-rich chromocenters that are decorated with H3K9me3 and HP1. G9a-depleted OP6 cells appear to have lost their
chromocenters, and heterochromatin is instead redistributed (e.g., to nuclear periphery).

Chromocenters contain repeat elements, such as major satellites, LINEs, and SINEs (Peters et al., 2001; Meshorer et al., 2006; Efroni et al., 2008). Chromatin decondensation during reprogramming liberates these repeat elements from repressive chromocenters and are transcribed in the ES and iPSCs (Fussner et al., 2011; Novo et al., 2017). We performed RNA FISH using probes that hybridize with satellite repeats on chromosome 2, 9 and 12 in order to measure transcriptional activity of these repeats. We predicted that these repeat elements might be up-regulated due to heterochromatin decompaction (loss of chromocenters) in the G9a-depleted background. A very preliminary result suggested that the satellite repeats are expressed in G9a CRISPR treated cells, in contrast to control cells, suggesting that G9a perturbation has generally reduced the repressive environment (Figure 1). This experiment requires validation by qPCR on cDNA samples, since it is possible that the increased RNA FISH signal in G9a-perturbed cells might in part or largely explained by increased accessibility and denaturation of satellite DNA in the absence of chromocenter structure. Also, these experiments were not controlled using sense RNA FISH probes in order to substantiate that these signals are due to RNA, not DNA, hybridization. I would also perform ChIP-qPCR (e.g., using H3K9me3, H3K27me3, H3K4me3, and H3K27Ac antibodies) to investigate chromatin states at repeat elements and whether they mirror the open states observed generally in stem cells (Jaenisch...
and Young, 2008; Novo et al., 2017). The results of these experiments might add another line of evidence that transformed OP6 cells are behaving in a stem-like manner.

At this point, we do not know the cause of the chromocenter reorganization or its significance. It is possible that merely a significant reduction of H3K9 methylation marks as a direct consequence of G9a perturbation is sufficient to destabilize chromocenters, a hypothesis that might predict a similar outcome if G9a is perturbed in any cell line. On the other hand, disruption of downstream regulatory networks in G9a-perturbed OP6 cells might be responsible for chromocenter destabilization. For example, the down-regulation of HP1 in these cells, which functions downstream of H3K9 methylation in a chromatin compaction pathway (Shridharan et al., 2009), might be the essential factor in this process. We would be interested in deleting HP1 directly, which we might predict phenocopies the chromatin reorganization phenotype of G9a depletion, as well as adding back HP1 (or other downstream chromatin factors identified by RNA-Seq, see 4.1.C below) to G9a-depleted cells to test whether the chromocenter phenotype (or any other phenotype associated with OP6 transformation) can be rescued.
Figure 1: 3D Chromocenter Phenotype and Transcriptional Activity of Satellite Repeats
(A) Characteristic G9a KD nucleus showing chromocenter phenotypes. All Z-stacks projected image in left panel showing radially located diffused chromocenters that looks peripherally organized towards nuclear periphery in individual Z-stacks.
(B) RNA-FISH showing major satellite repeat expression in G9a KD (left panel) and scramble control OP6 cells (right panel). Transcriptional activity of satellite repeat is seen by green intensities organized around diffused chromocenters in G9a KD cell.

4.1.C Has G9a Perturbation caused OP6 Cell De-differentiation?
One of the important observations we made was that not all GFP+ OP6 cells transform into stem-like cells (Chapter 3, Figure 1), suggesting the presence of
reprogramming intermediates. For example, GFP+ intermediates might fail to down-regulate OP6 markers and/or fail to activate pluripotency markers, a phenomenon commonly seen in somatic cell reprogramming (Werning et al., 2008; Stadtfeld et al., 2010b). G9a induces high levels of H3K9 methylation at pluripotency genes (e.g. OCT3/4) in order to silence them in somatic cells (Chen et al., 2013); therefore, G9a acts as a key barrier against somatic cell reprogramming (Matoba et al., 2014; Sridharan et al., 2009). Inhibition of H3K9 methyltransferases (e.g. G9a, Setdb1, Suv39h1) results in decondensation of chromatin at pluripotency enhancers and reduces the levels of H3K9 methylation at pluripotency gene promoters e.g. Oct3/4, Sox2, Klf4, C-myc (OSKM) to permit expression of these genes (Soufi et al., 2012; Shridharan et al., 2013; reviewed in Hochedlinger and Jaenisch, 2015; Becker and Zaret, 2016).

We have perturbed G9a in OP6 cells causing a stem-like transformation in terms of cell behavior, morphology, and gene expression. We have proposed that G9a perturbation has caused de-repression at key pluripotent gene loci, and the reactivation of one or more of these genes (e.g., Sox2) might be sufficient in the OP6 context to activate stem cell regulatory networks that stabilize the de-differentiated state. However, this hypothesis is at this point only casually supported. A major next focus in this project is to more rigorously investigate the downstream effects of the G9a perturbation at a more global level.

The first and most fundamental line of experiments involves more carefully monitoring and recording the cell transformation process. As described
in Chapter 3, we observed that GFP+ cells post-transfection tends to change morphology from a flattened (OP6-like) to rounded (ESC-like) cell that rapidly divides to form small flat colonies that later in their development grow into larger, more differentiated 3D colonies. We would like to monitor this process in real time using live cell imaging. We would like to measure the fraction of GFP-positive cells that forms colonies, and whether these colonies are clonal (i.e., as opposed to apparent recruitment or intermingling of peripheral cells during colony growth). We would like to understand the process of colony differentiation – are there specific events, such as contact with surrounding cells or reaching a critical colony size – that causes colonies to behave differently or execute changes in gene expression patterns (e.g., the observed differentiation within colonies in which GFP+/Sox2_{high} cells occupy the central core and GFP-/Sox2_{low} cells reside in the colony periphery). More full characterization of the transformation process would include measuring changes in cell division rates during colony maturation, and whether all colonies progress in a stereotyped manner. While these would be qualitative and descriptive data, observations would be complemented by detailed analyses of gene expression, as described below.

While we have thus far taken a candidate gene approach (e.g., investigating specific stem cell markers by immunofluorescence or qPCR), the best approach for interrogating the epigenetic and transcriptional landscape within the cells of transformed colonies would be to take advantage of deep
sequencing technology. Specifically, more global ChIP-Seq experiments will permit characterization of the epigenetic landscape in control populations (e.g., regular OP6 cells, ES cells) versus transformed cell populations at various stages of progression. RNA-Seq can be conducted at the resolution of single cells in transformed OP6 populations, which will permit very precise correlations of GFP expression, G9a expression, stem marker up-regulation, and neural marker down-regulation, as well as expression of non-stem/non-neural regulatory networks associated with other lineages. For example, we might observe the gradual extinction of the OP6 genetic program (e.g. down-regulation of GAP43, etc.) and subsequent activation of endogenous pluripotency genetic networks (e.g. activation of OSKM Yamanaka factors). Most importantly however, we might be able to cluster gene expression profiles into discrete categories that appear to define: (a) the wild-type OP6 state, (b) a de-differentiated state that still appears to be largely along the OSN/neural lineage, (c) a stem-like state that has little/no evidence of lineage marker expression and maximum activation of pluripotency genes (also see 4.1.E below), and (d) various “re-differentiation” or “trans-differentiation” states that exhibit reduced/little stem marker expression and neural/OSN marker expression, yet exhibit apparent activation of regulatory networks characteristic of other cell lineages (also see 4.2 below). Through such an analysis of gene expression profiles, we aim to identify key genes (“progression markers”) whose expression status defines each of the discrete clusters, so that we could then use immunofluorescence (RNA in situ or protein
staining) in order to map these defined stages back to the cultures at various stages of colony development. Together, these experiments represent a longer-term and more complete characterization of the changing epigenetic and gene expression landscape associated with OP6 cell transformation and colony formation.

4.1.D Has G9a Perturbation Activated Endogenous Cell Cycling Networks?

The OP6 cell line is immortalized using SV40 large T antigen (TAg) that inactivates retinoblastoma (pRb) and p53 tumor suppressor genes (DeCaprio et al., 1988; Ahuja et al., 2005), forcing the post-mitotic OP6 cells to re-enter the cell cycle (Illing et al., 2002). The expression of SV40 large T antigen inactivates retinoblastoma by phosphorylation that reconstitutes the cyclin D/CDK complex, priming the cell for proliferation via inducing re-entry into the cell cycle (Tanaka et al., 1997; Latella et al., 2001; Sage et al., 2003; Tiainen et al., 1996; reviewed in Endo and Nadal-Ginard, 1998; Echeverri and Tanaka, 2002).

As noted previously, Neurog1+ progenitors in the OSN lineage that were induced by chromatin perturbation to undergo dedifferentiation in vivo were cycling cells (Lin et al., 2017), and generally, de-differentiation is enhanced in mitotically active cells (Schneider et al., 1994; Mal et al., 2000; Egli et al., 2008; Campbell et al., 1996a, b; Cibelli et al., 1998; reviewed in Odelberg et al., 2002). Reprogramming efficiency is greater in actively dividing cells because of reduced repressive H3K9 and H4K20 methylation levels that tend to prevent
destabilization of cell states (Ganier et al., 2011). For example, fibroblast nuclei incubated with *xenopus* egg extract isolated during M phase reprogrammed more efficiently than if isolated from G1 phase (Ganier et al., 2011). In a similar study, permeabilized somatic cell nuclei isolated from M phase (but not other cell cycle phases) reprogrammed efficiently upon incubation with enucleated egg cells (Halley-Stott et al., 2014). This mitotic advantage was dependent on ubiquitination of H2A and H2B histones in order to facilitate chromatin remodeling during M phase (Joo et al., 2007; Halley-Stott et al., 2014; Hannna et al., 2009).

These observations raise an interesting question with regard to OP6 transformation: to what extent is this process enabled or facilitated by the presence of the large T antigen in OP6 cells, which might very well reduce cell-cycling barriers that must be overcome in order to de-differentiate and/or form 3D (ES-like) colonies? A prediction of this hypothesis is that transformation will not occur when the temperature-sensitive large T antigen is deactivated, which we can test by doing a primary transfection at the non-permissive temperature. A related question is whether sustained growth of transformed colonies requires ongoing large T antigen expression, or whether the process of initial transformation has re-activated stem-like cell cycling networks that permit large T antigen-independence thereafter.

As discussed in Chapter 3, initial gene expression analysis suggests that cyclin D1 (the G1-phase specific cyclin) is down-regulated in transformed OP6
cells (Fig. 4). The down-regulation of Cyclin D1 and D2 is also commonly seen in pluripotent ES cells (Stead et al., 2002; Faast et al., 2004; Wianny et al., 1998), suggesting transformed OP6 cells might have acquired ES-like cell cycle properties. To further support this hypothesis, a more extensive gene expression and time course study for the various cyclins and cyclin-dependent kinase (CDKs) is required. For example, Cyclin A/E and CDK2 complex is constitutively expressed throughout all cell cycle stages in ES cells (Wianny et al., 1998; Stead et al., 2002; Wu et al., 2015; Chen et al., 2014), whereas this complex is inhibited by CDK inhibitors (CDKI) such as p21, P27 in differentiated cells (Besson et al., 2008; Faast et al., 2004; Dolezalova et al., 2012). The Rb protein is phosphorylated (thus inactivated) at all stages of the cell cycle in ES cells, whereas Rb is unphosphorylated at G1-S phases in terminally differentiated cells (Sage et al., 2000; White et al., 2005). Due to constitutive expression of CyclinA/E and the CDK2 complex, as well as the lack of CDKI and R-point, the doubling time of ES cells is faster as compared to differentiated cells (reviewed in Boward, Wu, and Dalton, 2016). Therefore, we anticipate cell doubling-time for transformed cells would be faster and similar to ES cells (positive control) as compared to control OP6 cells (Pardee et al., 1974).

Preliminarily, I have observed the loss of large T antigen expression in many of the cells that form colonies (Figure 2), and our lab has observed aggressive colony expansion when grown at the non-permissive temperature (G. Abbas, personal communication), suggesting that transformed OP6 cells might
indeed have achieved independence through reactivation of an endogenous mitotic state. The extensive gene profiling described in 4.1.C above will provide additional clarity on these questions, not only in terms of more fully characterizing changes in cell cycle regulatory networks during the transformation process, but also monitoring these changes as they occur in primary, secondary, and tertiary colonies, as well as during colony development and differentiation.

**Figure 2: RNA-FISH Analysis using Antisense Large T antigen Probe** suggests uniform expression of Large T antigen (coding sequence probe shown in green in cytoplasm) in immortalized OP6 scramble control cells (left panel), whereas, a single G9aKD cell (lacking the chromocenters, red arrow in right panel) within a small colony do not express Large T antigen (lack of green signal in cytoplasm). Therefore, that G9a KD cell might be independent of expression requirement for Large T antigen and may be dependent on pluripotency genes for proliferation. DAPI counter stain in blue. RNA-FISH signal in green.
4.1.E Has G9a Perturbation Activated Stem Cell Networks?

The pluripotency transcription factors (e.g. Oct4, Sox2, Klf4, and c-Myc (OSKM)) play a pivotal role in reprogramming (Takahashi et al., 2006; Yamanaka et al., 2009). These factors serve two major functions: 1) repression of somatic genes (Stadtfeld et al., 2008b); 2) activation of endogenous pluripotency and cell cycle genes (Brambrink et al., 2008; Stadtfeld et al., 2008b; reviewed in Jaenisch and Young, 2008). In differentiated cells, somatic genes and tumor suppressor genes are decorated with histone acetylation marks (and are transcribed at higher levels), whereas pluripotency genes are decorated with repressive histone and DNA methylation marks and are transcriptionally silenced (Kaji et al., 2006; Boyer et al., 2006; Lee et al., 2006; Hanna et al., 2008b; Mikkelsen et al., 2008; Mansour et al., 2012; Shridharan et al., 2009). During reprogramming, the pluripotency factors function as repressors at somatic gene promoters where they recruit histone deacetylase (HDAC1, HDAC2) and polycomb repressor complex 2 (PRC2) to reduce histone acetylation and methylation, respectively (Maherali et al., 2007; reviewed in Jaenisch and Young, 2008; Stadtfeld and Hochedlinger, 2010). Simultaneously, reprogramming factors occupy the promoters of pluripotency genes where they function as activators by recruiting histone acetyltransferase (HAT) and polymerase II (Pol II) in order to stabilize the pluripotent state (reviewed in Jaenisch and Hochedlinger, 2015). These pluripotency factors also bind to their own promoters in a positive autoregulatory transcriptional loop to further stabilize the pluripotent state.
(Kuroda et al., 2005; Odom et al., 2006; Alon et al., 2007; Maherali et al., 2007; Okita et al., 2007; Werning et al., 2007). These autoregulatory positive-feedback loops maintain endogenous pluripotency and self-renewal networks even in the absence of ongoing exogenous transgene expression or other inducing factors (Masui et al., 2007; Werning et al., 2007; Okumura-Nakanishi et al., 2005).

Several molecular and cellular events are commonly observed during reprogramming, such as the down-regulation of key somatic marker genes (e.g., Thy1, Snail, Neurog1, Pax6, Lhx, Dbx), and activation of early pluripotency marker genes (Oct4, Nanog, Sox2, AP, SSEA1, Fbxo15). Other features include the silencing of ectopic retrovirus elements (via HDACs and DNMT3a/b), telomere extension (via TERT activity), and X-chromosome reactivation (Brambrink et al., 2008; Stadtfeld et al., 2008b, 2010b; Werning et al., 2008; Hanna et al., 2008b; reviewed in Jaenisch and Hochedlinger, 2015). In Chapter 3, I reported some preliminary evidence for some of these molecular events by qPCR (Figure 4) and IF (Figure 5). As described in 4.1.C above, a more extensive gene expression profiling at the resolution of single cells and across various stages of colony development will be extremely informative in the complete characterization of these transformation events. The activation of a stem cell regulatory network makes additional predictions of both the chromatin states of downstream loci (e.g., lineage-specifying genes, etc.) and transcription-factor occupancy at key target genes, both of which can be tested experimentally using ChIP.
However, our preliminary data so far does not seem to be consistent with a complete regression to an ES-like state. There are key stem marker genes typically associated with embryonic stem cells, such as Oct4 and Klf4, that do not appear to be up-regulated, and even those up-regulated stem markers do not appear to express at the very high levels observed in ES cells (G. Abbas, personal communication). Moreover, some of the neural marker genes, such as MAP5, Nestin, GFAP, Asc1, and GAP43 do not appear to be down-regulated. It is currently not clear whether these observations are a consequence of significant heterogeneity in the cell population (requiring more rigorous single-cell analyses, as described in 4.1.C above), or whether these observation indicate a “partial” regression in which transformed OP6 cells have acquired an intermediate pluripotency signature while still retaining their OSN ancestry. It should be noted that these transformations were conducted in undefined media containing numerous growth factors, an environment that might be sub-optimal for de-differentiation. For example, the fetal bovine serum present in the OP6 growth media contains bone morphogen proteins (BMPs) that inhibit pluripotency gene networks and promote somatic gene networks (Chen et al., 2013; Herrera et al., 2009; David et al., 2008). We have yet to explore different growth environments, for example, expanding colonies using low serum and/or BMP inhibitors that might advance the transformation process more fully.

As noted, we observe heterogeneity in cell populations even within the compacted, ES-like colonies. For example, our IF data shows up-regulation of
Sox2 in the central portions of colonies (that tend to be GFP positive), and FGF5 in the peripheral regions of colonies (that tend to be GFP negative, as well as exhibiting a more flattened cell morphology), raising the possibility that colonies contain mutually exclusive cellular compartments. We speculate that such colony organization might be achieved via cell-cell communications, as such an organization might be favorable in a stem cell niche in order to maintain a pluripotent pool while “spinning off” differentiated cells through asymmetric cell divisions that can migrate and further progress within a particular lineage (Guo et al., 2017; Kim et al., 2014; Park et al., 2017; reviewed in Apostolou and Stadtfeld, 2018). Moreover, these asymmetrical cell divisions (and differentiation in general) might require G9a activations in order to re-silence the stem network, and thus the suppression of the gRNA cassette. Of course, there should be no biological predisposition for a stem cell to communicate to a progeny cell specified to differentiate that it should deactivate a G9a gRNA (i.e., such that GFP becomes deactivated in a daughter cell destined to differentiate). So, an important outstanding question is, how do colonies arise with apparently more differentiated cells expressing lower levels of GFP (G9a gRNA) on the periphery? One plausible answer is selection: a colony simply cannot grow and develop differentiated cells unless the G9a gRNA is deactivated at high enough frequency for this to be a common enough occurrence, and that once a GFP-negative cell is generated (now primed to differentiate via the reactivation of G9a), it migrates to the periphery according to cell-cell communication signals.
within the colony. Another plausible answer is recruitment: the stem cell core might not asymmetrically divide (rather, just grows larger in size), however, the colony recruits GFP-negative cells into the peripheral regions of the colony where they receive further differentiation instructions. To distinguish between these two scenarios, we would like to conduct time-lapsed live cell imaging where the former hypothesis would predict that one of two daughter cells will on occasion extinguish GFP expression prior to migrating to the periphery and differentiating, and in contrast, if both daughter cells continue expressing GFP they might remain at the colony core. Time-lapse imaging should also dismiss/support the latter hypothesis, since we would be able to observe recruitment/migration, or absorption of peripheral (GFP-negative) cells at the periphery of the colony as it grows, while observing no apparent GFP patterning as cells divide in the core. It would also be interesting to know whether expression of pluripotency markers is dependent on deactivation of G9a protein; similarly, we might predict cells that continue to express G9a (or reactivate G9a) would be more likely to acquire a more differentiated phenotype. These questions could be addressed by two-color IF – if G9a deactivation is required for up-regulation of pluripotency markers and G9a activation is required for more differentiated states, then we would predict that G9a and Sox2 would exhibit mutually exclusive expression within colonies. A similar two-color strategy can be utilized to test for mutually exclusive expression of early- versus late-stem markers, or stem- versus neural- (or other differentiation) markers, as
discussed in more detail in section 4.1.C above (see RNA-Seq and expression profile clustering analyses).

In summary, the main objectives of these particular future directions would be to study the de-differentiation process at higher resolution – for example, gene expression profiles at the single cell level, organization and asymmetrical cell division properties of putative stem cell colonies, and the relationship between G9a expression/non-expression in the de-differentiated/differentiated cells within colonies.

4.1.F Do Transformed OP6 Colonies Exhibit Tumorigenic Properties?

There is a striking similarity between cell dedifferentiation/reprogramming and tumorigenesis (reviewed in Friedmann-Morvinski and Verma, 2014; Zhang et al., 2013; Iglesias et al., 2016; Malta et al., 2018). Both processes involve up-regulation of the Yamanaka factors (Oct4, Klf4, Sox2, cMYC), presumably to facilitate transformation and stabilization of an aggressively mitotic state (reviewed in Iglesias et al., 2016; Malta et al., 2018). For example, Sox2 expression is overexpressed in lung and esophagus cancers, and is found up-regulated in glioblastomas, breast cancers, and Ewing sarcomas (Riggi et al., 2010; Bass et al., 2009; Leis et al., 2012; Iglesias et al., 2014; reviewed in Sarkar and Hochedlinger, 2013). Oct4 is found in germ cell tumors, breast, thyroid and prostate cancer cells (Gidekel et al., 2003; Madjd et al., 2009; Zhou et al., 2011; De Resende et al., 2013; Hochedlinger et al., 2005). Similarly, stem cells, like
tumor cells, exhibit down-regulation of tumor suppressor genes, such as p53 and pRb that enable cell cycle progression and minimize apoptosis and senescence barriers (Kawamura et al., 2009; Tapia et al., 2010; Riggs et al., 2013). Both ES/iPSC cells and cancer cells switch to glycolytic metabolism (Warburg effect) from oxidative phosphorylation normally utilized in somatic cells (Riggs et al., 2013; Vazquez-Martin et al., 2012; Menendez et al., 2014). There is a significant overlap between transcriptome profiles between iPSCs and tumor cells (Riggs et al., 2013; Malta et al., 2018). Both ES/iPSC cells and tumor cells have open chromatin that leads to chromosomal instabilities and broader differentiation potential (Orkin and Hochedlinger, 2011).

How can we distinguish whether G9a-mediated OP6 cell transformation is a de-differentiation or oncogenic process? Despite these striking similarities between pluripotent and tumor cells, there are molecular markers and cellular processes that differ between stem cells and cancer cells. For example, tumors are initiated by epithelial-mesenchymal transitions (EMT) whereas pluripotency is initiated by mesenchymal-epithelial transition (MET) (Mani et al., 2008; Thiery et al., 2009; Yang et al., 2008; Li et al., 2010). Since pluripotency factors inhibit expression of epithelial markers such as Snail1, Zeb1, Twist1, E-Cadherin, I predict that these markers will be down-regulated in OP6 colonies if they are stem-like, but up-regulated if they are tumor-like. Stem cells exhibit elevated lipids with long acyl chains as compared to cancer cells (Guler et al., 2018), and therefore, lipid analyses using biochemical assay kits available from Abcam or by
using Fourier-transform infrared spectroscopy (FTIR) might be informative. The glycogen pool is higher in tumor cells as compared to pluripotent stem cells (Guler et al., 2018; Riggs et al., 2013), and more generally, RNA-Seq data (4.1.C) will be extremely useful for comparing/contrasting transformed OP6 expression profiles to those of stem cells and tumor cells (e.g., published transcriptome profiles for various tumors; Malta et al., 2018). Finally, an important difference between stem cells and tumor cells is the ability of the former to exhibit exquisite responsiveness to environmental signals and to develop highly ordered structures within the embryo. In the following section, I turn to discussing the future directions that investigate the developmental potential of transformed OP6 cells.

4.2 Further Characterization of Transformed Colony Potential

In this section of my thesis, I will discuss assays that can be used to characterize the differentiation potential of stem-like colonies generated from OP6 cells. There are five standard functional assays to assess the developmental potential of stem cells: 1) in vitro differentiation, 2) teratoma formation, 3) chimera formation, 4) germline contribution, and 5) tetraploid complementation (Takahashi and Yamanaka, 2006; reviewed in Jaenisch and Young, 2008). Each of these assays is briefly described below.

Mouse stem cell colonies growing in 3D are differentiated spontaneously into embryoid bodies due to cell-cell interactions within the colonies (Polo et al.,
As a result, some cells within embryoid colonies down-regulate pluripotency factors and acquire properties of differentiated cells that express markers for all three germ-layers (Aasen et al., 2008). For example, β-III tubulin (TUJ1) is a marker gene for the ectoderm germ layer, smooth muscle actin (SMA) is a marker gene for the mesoderm germ layer, and α-fetoprotein (AFP) is a marker gene for the endoderm germ layer (Takahashi and Yamanaka et al., 2006). We wish to test for germ layer differentiation using immunohistochemistry (IHC), where we would expect to detect a mutually exclusive pattern of germ layer marker expression within colonies. The appearance of germ layer markers, especially markers not expected within the OSN or other neuronal or ectodermal lineage, would suggest re- or trans-differentiation potential of transformed OP6 cells.

It should be noted however, that while marker gene expression is informative, it does not address the functionality of the pluripotent stem cells (Sheridan et al., 2012; reviewed in Hochedlinger and Jaenisch). One general test for stem cell pluripotency function is an in-vivo teratoma formation assay (Zhang et al., 2012; Raab et al., 2017; reviewed in Hochedlinger and Jaenisch). A teratoma is a germline tumor with differentiation potential of stem cells because these tumors will generate the various lineages of the 3 germ layers: for example, neurons (ectoderm), cardiomyocytes (mesoderm) and hepatocytes (endoderm) (Takahashi and Yamanaka, 2006; reviewed in Young and Jaenisch, 2008). If transformed OP6 cells are not stem-like cells then they will not
generate an organized teratoma containing all 3 germ layers (Hochedlinger and Jainisch, 2015).

A limitation of the teratoma assay is that it does not test the ability of dedifferentiated / induced stem cells to promote normal development since it instead models neoplastic transformation (Daley et al., 2009; Buta et al., 2013; Nelakanti et al., 2016). Therefore, a chimera formation assay is preferred in which reporter (e.g., GFP) labeled stem cells are injected into a host blastocyst, and resulting embryonic tissue is monitored for having arisen from these labeled cells (Boroviak et al., 2014; Huang et al., 2012; reviewed in Wu and Belmonte, 2015). The appearance of labeled cells in multiple tissue types would suggest that transformed OP6 cells are pluripotent. Another stringent approach to test for pluripotency is a tetraploid complementation assay in which donor cells are injected into a 4n (tetraploid) host blastocyst. Since 4n host cells cannot contribute to the somatic lineages (Eggan et al., 2001), the resulting embryo would be exclusively composed of donor cells, except for the trophectoderm (placental) lineage (Nagy et al., 1990). A limitation of these latter two in vivo experimental approaches is that host cells in the blastocyst may complement donor cells so that they appear to be fully competent and autonomous when they might not be; for example, host blastocyst cells secrete factors that may partially reprogram donor stem cells to acquire full pluripotency (Xiang et al., 2018; Cohen et al., 2018). Therefore, it is common to use a combination of these techniques (e.g., in vitro and in vivo) to establish pluripotency.
4.2.A Are Transformed OP6 Cells Pluripotent?

We were curious as to whether the embryoid-like colonies generated from OP6 cells are capable of differentiating into various cell types. The first question we asked was whether there was any evidence of spontaneous differentiation within transformed OP6 colonies (i.e., in the absence of any applied differentiation protocol). We used antibodies against marker genes commonly used to identify cell types from the three canonical germ layers: Tuj1 (beta-III tubulin as an ectodermal marker), Smα (smooth muscle actin as a mesodermal marker), and AFP (alph-fetoprotein as an endodermal maker). Our preliminary result suggests that stem-like colonies contained cells that expressed each of these marker proteins (Figure 3), suggesting that differentiation along multiple lineages might have occurred as colonies mature. These very initial experiments were conducted just prior to leaving the lab and were added to a larger IF experiment. Due to limiting time and materials, the germ-marker IF experiment shown in Figure 3 was not conducted on untreated OP6 cells (negative control populations) or on ES cells (positive control populations) in parallel. A more rigorously controlled experiment is warranted in order to validate these results. We note however that the IF staining for each of these markers is not ubiquitous within colonies, and the lack of staining in many of the cells suggests that this signal is not due to non-specific cross-reactivity, assuming that a non-specific antigen is likely to be present in every cell. We also note that subcellular staining of these proteins agreed with previous publications; e.g., Tuj1 staining resembles
elongated processes that represent immature neuronal stages, SMA stains actin skeletal in the cytosolic compartment, and AFP accumulates in the cytoplasm as a growth associated protein. However, since germ marker differentiation is spontaneous and random, a co-labeling of all three-germ marker is required to investigate whether these markers are expressed mutually exclusively within these colonies.

Figure 3: Transformed Colonies Express Markers from Three Germ Layers
A representative immunofluorescence images for marker genes from the three canonical germ layers (scale bars = 25µm): (a) Tuj1 (β-III tubulin ectodermal marker), (b) Sma (smooth muscle actin mesodermal marker), and (c) AFP (α-fetoprotein endodermal maker); insets show staining for these three markers in isolated cells within each culture (scale bars = 15µm). Nuclei are stained with DAPI (blue) in all images.

Assuming for the moment that the IF staining for these 3 markers is not arising from cross-reactivity or false-positives, we note that there is no apparent
patterning in terms of the locations of positive cells within colonies. In fact, we were surprised to observe robust expression of germ layer marker proteins in the interior portions of the colony, since we had hypothesized based on cell morphology, growth rates, and Sox2 expression that these central regions would be largely comprised of de-differentiated/undifferentiated “stem-like” cells. At this time, we cannot dismiss the possibility that these interior cells are behaving chaotically (e.g., perhaps cells had transformed into a germ line-like tumor; Kleinsmith and Pierce, 1964; Martin et al., 1981), perhaps co-expressing pluripotent marker genes, as well as marker genes for one or more of the canonical developmental lineages. In order to investigate whether there exists an organized (non-tumor-like) pattern of germ layer expression within transformed colonies, I would perform three-color IF using all three germ marker antibodies in a single IF experiment, where I would predict non-overlapping expression if, in fact, individual positive cells are each differentiating along only one particular lineage.

We speculate that differentiation, if it is occurring, may very well require the presence of G9a to re-silence stem regulatory networks. As reported, G9a expression is heterogeneous within transformed colonies (see Figure 2, Chapter 3). If so, we would predict that the patterns of G9a expression within transformed colonies might complement differentiation markers. To investigate this, I would perform two-color IF using G9a and each of the germ markers.
The microenvironment of the stem cell niche (adult or embryonic stem cells) and tissue origin of these stem cells play an important role in tissue fate determination (reviewed in McCauley and Wells, 2017). The most compelling test for pluripotency of transformed OP6 cells would be the ability to differentiate along defined developmental lineages given the right environments. We propose both in vitro and in vivo experiments to address this question. An in vitro trans-well differentiation assay (Murrell et al., 2005), in which transformed OP6 colonies are seeded and incubated with medium conditioned with soluble factors derived from living neonate tissues (e.g. liver, heart, skeletal muscles, kidney) or commercially available recombinant supplements, is commonly utilized to investigate pluripotency by testing the inductive capability of the seeded cells. The cells from the colonies sense the signals from the tissue extracts (or supplements) and differentiate into cells with morphology and molecular properties appropriate for the signal environment. For example, to differentiate stem cells into hepatocytes, the media is supplemented with activin A, BMP4, bFGF, HGF, and oncostatin M (Si-Tayeb et al., 2010). To differentiate stem cells into cardiomyocytes, the media is supplemented with B27, H-albumin, L-ascorbic acid, lactate, and Wnt-C59 (Burridge et al., 2014). The differentiated hepatocytes can be verified by FOXA2, SOX17, GATA4, HNF4-α, AFP, ferritin and albumin antibody staining (Si-Tayeb et al., 2010), and the differentiated cardiomyocytes can be verified by troponin (TNNT2), α-SMA, α-actin, MLC2A/2C antibody staining (Burridge et al., 2014).
As an initial attempt to address questions of this nature, we grew colonies in the semi-solid cellulose M3434 complete media used previously to specifically induce blood-cell lineage development from pluripotent precursors (Kerenyi et al., 2014; Griseri et al., 2012; Kats et al., 2014). Preliminarily, upon visual examination, we did not observe any evidence of blood cell differentiation when applied to transformed OP6 colonies. Instead, we observed the development of what appeared to be differentiated neurons, some with elongated, elaborate processes and gland-like structures with central lumen (Figure 4). Importantly, these neurons did not resemble olfactory neurons, which are distinctly bipolar in character (Buck and Axel, 1991; Illing et al., 2002; Pathak et al., 2009), and therefore, appear to be a novel differentiation outcome as opposed to the reestablishment of the original OP6 lineage. The fact that the colonies did not apparently develop cells of the blood lineage as we had anticipated might indicate an incomplete regression of OP6 cells upon G9a perturbation (i.e., not all the way back to an ES stage), rather, these cells might retain lingering epigenetic memory of its neuronal heritage that biases differentiation potential upon induction.
We collected secondary stem-like colonies and washed with PBS, made homogenous cell suspension and mixed with semi-solid M3434 (stem cell technology) media supplied with cytokines and interleukins (IL3, IL2, stem cell factors, Erythropoietin). Cell suspension was dispensed into 5mm² patri-dishes and differentiated for two weeks. (A) A representative image showing differentiation of transformed OP6 cells into gland like structures (blue arrows, left panel) with centrally located lumens. (B) a representative image showing differentiation of colonies into neurons (blue arrows, right panel).

Alternatively, the existing in vitro media environment might constrain either the de-differentiation or re-differentiation potential of transformed OP6 cells. While experiments could be conducted to alter the growth media in an attempt to encourage more robust ES-like transformation (e.g., REFs), a definitive test of pluripotency would be one or more of the in vivo assays described previously. As discussed, transformed OP6 colonies could be dissociated into a single cell suspension, labeled with a cell tracking dye / fluorescence protein (e.g. GFP/ LacZ), and injected into the embryo or in various
developmental niches in the adult mouse. The donor cell contribution to
developmental lineage and/or tissues, such as cardiac atrium/ventricles, liver,
blood vessels, brain, spinal cord, skeletal muscles and gut, can be assessed by co-
labeling of cell tracker dye / fluorescence marker with tissue specific markers.
Alternatively and in parallel, I would conduct a teratoma assay in which a
homogenous suspension of transformed OP6 cells from the colony are injected
into immune-compromised mice (SCID mice) within various developmental
niches (e.g., intramuscular, subcutaneous, intra-testicular, and under the kidney
capsule (Hentze et al., 2009; Peterson et al., 2011; Zhu et al., 2014). If these
transplanted cells are pluripotent, they should develop teratomas within 6-12
weeks, induced by local cues (Cooke et al., 2006; Cao et al., 2007; Nelakanti et al.,
2015). When tumors reach a sufficient size, they can be removed and subjected
to histopathological analysis by eosin staining, immunohistochemical analysis by
germ-marker staining, and gene expression analysis by RT-PCR for Tuj1, SMA,
and AFP assays in order to establish their pluripotency.

4.2.B Are Transformed OP6 Colonies Capable of Self-organized Development?

The reproducible differentiation of pluripotent stem cells into self-organized,
multicellular 3D assembly of cells that mimic organogenesis in the canonical
developmental niche has been reported for several adult and embryonic stem
cell-derived organoids (reviewed in McCauley and Wells, 2017; Sato and Clevers,
For example, organoids of cerebellum (brain) (Lancaster et al., 2013), optic cup (Nakano et al., 2012), kidney (Takasato et al., 2014), lung (Dye et al., 2015), stomach (McCraken et al., 2014), and intestinal tissues (Spence et al., 2011) can be derived from embryonic stem cells. Mini-organoids of hindgut (Sato et al., 2009), stomach (Barker et al., 2010), pancreas (Boj et al., 2015), liver (Huch et al., 2013) can be derived from adult stem cells. These in vitro derived organoids contain complex cell types, anterior-posterior polarity, architecture, and functional similarities with in vivo counterparts. Various factors secreted from the developmental niche (such as growth factors and cytokines), along with various developmental signaling pathways, govern specification, patterning of the three germ layers, and tissue morphogenesis (reviewed in Zorn and Wells, 2009). The onset and duration of signaling pathways, the dose and combination of growth factors, and the interplay between multiple signaling pathways are involved in the early and late patterning of these tissues (McCracken et al., 2017; Dye et al., 2015; Takasato et al., 2015). For example, a high level of Nodal factors (e.g. activin A) induces endoderm specification, whereas low levels of these Nodal factors promote mesoderm specification, and the (default) ectoderm state forms in the absence of Nodal factors (Gadue et al., 2006; D’Amour et al., 2005). Early patterning of mesoderm and endoderm differentiation requires growth factor gradients of Wnt, FGF, RA and BMP signaling pathways, whereas neuroepithelium (ectoderm) differentiate in the absence of these growth factors (Nakano et al., 2012; Takasato et al., 2015; McCracken et al., 2017). Late
patterning of mesoderm, endoderm and ectoderm additionally requires tissue specific growth factors (Dye et al., 2015; Spence et al., 2011). For example, retinoic acid signaling determines cerebral fate, whereas sonic hedgehog (Shh) signaling determines optic fate from a common neuroepithelium precursor (Lancaster et al., 2013).

Interestingly, we observed that with prolonged culturing (~3 weeks) of hundreds of stem-like colonies grown in regular OP6 media (undefined with fetal serum) on rare occasion spontaneously produced multicellular and organized developmental structures (Figure 6, Chapter 3). As discussed in Chapter 3 (Section 3.3.F), these structures contained sharp borders, anterior-posterior polarity, repeated segmentation, and differentiated microvilli / ciliated cells. We have two pressing questions about these structures: how do they form, and are they organoids that model mouse development in vivo? With regard to the first question, we need to more carefully monitor colony growth and these spontaneous developmental events using time-lapse photography. We would like to determine factors within transformed colonies that make these events more likely. For example, we don’t know whether smaller or bigger colonies are more likely to develop structures, or whether certain patterns of marker genes within these colonies (e.g., G9a, Sox2, or other marker gene patterns) are stereotypical prior to these developmental events. One way to investigate this question is to capture colonies just as initial stages of development are evident (e.g., evidence of protrusions, invaginations, or other shape changes) and
monitor gene expression patterning using various two-color experiments (e.g., G9a, Sox2 and embryoid body markers, such as TUJ1, SMA, AFP). At this point, the goal would be to simply identify telling features that seem to predict the progression from a colony just expanding in size versus a colony that begins to differentiate in apparently organized ways.

The second question concerns characterization of the developmental objects that have arisen in the current experimental context. The proposed experimental approach would be to stain whole mounts or to perform cryosectioning of a developmental structure into 50-70-micron tissue slices in order to conduct in situ hybridization and immunohistochemistry. In parallel, we propose to conduct single cell RNA-seq to genetically characterize various cell types present in these structures in a more exhaustive way. Can we identify specific cell types with high confidence? Is the organization of specific cell types consistent with a particular tissue (e.g., organoid)?

Finally, we observed very similar characteristics among 7 different structures that emerged under these conditions. At this time, we do not know if this approximate reproducibility is due to genetic/epigenetic predisposition (e.g., all transformed colonies analyzed in this preliminary work were derived from a single primary OP6 colony) or due to a common environmental state (i.e., all transformed colonies were grown in the standard OP6 growth media). To investigate further, we would like to grow additional secondary/tertiary OP6 colonies in the same media from various primary colonies (i.e., not from just the
one founder), and/or, grow existing secondary/tertiary colonies in different media containing growth factors that would normally encourage specific lineage progressions (reviewed in McCauley and Wells, 2017; Sato and Clevers, 2015; also see Section 4.2.C below).

Finally, we speculate that if transformed OP6 colonies are embryoid bodies expressing different germ layers markers, these various cell types might similarly interact with each other to differentiate into 3D floating organoid-like structure observed in this culture. We would like to explore co-culturing techniques used in other contexts for inducing organoid development along various lineages. For example, co-culturing of embryonic stem cells derived hepatic endoderm with mesenchymal stem cells (MSCs) and human umbilical vein endothelial (HUVECs) induces 3D morphogenesis of hepatic endoderm to form a vascularized liver bud (Takebe et al., 2013). Similarly, combining neural crest cells with mid/hindgut spheroids results in 3D intestinal organoids with enteric neurons and glia that responsible for peristaltic gut movements (Workman et al., 2016).

In this section, we discussed the strategies to investigate reproducibility and developmental potential of the transformed OP6 colonies. Specifically, we are interested to reproduce the preliminary data presented here by growing additional primary colonies. We will analyze their developmental potential in a similar or different conditioned media, e.g. co-culturing with various tissue extract and be able to differentiate OP6 stem-like colonies into desired cell types.
Our primary goal would be to fully characterize observed structures to investigate cell composition, nature of tissues and whether it is organoid.

### 4.3 Future Studies on Epigenetic Regulation of OR Genes

LSD1 was proposed to de-repress single OR alleles via H3K9 demethylation *in vivo* (Lyons et al., 2013). Therefore, ORs should not be activated when LSD1 is depleted. However, low levels of ORs were still expressed in the absence of LSD1, suggesting LSD1 mediated OR de-repression may function to increase OR expression levels per se, which requires H3K9 demethylation at a chosen OR allele (Lyons et al., 2013; Vyas et al., 2017). How might LSD1 be involved in OR gene regulation prior to amplification of OR expression level is not known. Furthermore, it is possible that LSD1 may demethylate at more than one OR loci. In this scenario, the error-rate of multiple OR de-repression events (via H3K9 demethylation) should be minimized prior to the UPR feedback mediated downregulation of LSD1 to stabilize initial OR choice (Tan et al., 2013; Nagai et al., 2016; Dalton et al., 2014). In this notion, a transient and compartmentalized LSD1 phenotype was observed that is thought to limit LSD1 activity to a single or a small subset of ORs (Lyons et al., 2013; Kilinc et al., 2016). Another striking result was that the expression levels of majority of ORs are decreased in the absence of G9a/GLP (enzymes that add methyl groups on H3K9) as compare to control mice, and fewer than 10 ORs were upregulated than normal levels. This upregulated ORs were not dependent on LSD1 mediated H3K9 demethylation,
suggesting LSD1 may be partially dispensable and H3K9 removal may not be a rate limiting step for OR activation (Lyons et al., 2014; Tian et al., 2016). Moreover, single cell RNA-seq suggests that multiple ORs are expressed in immature OSNs, suggesting the initial OR choice is not singular (Hanchate et al., 2015; Tan et al., 2015). The UPR feedback mechanism ensures the stability of the singular OR expression by downregulating LSD1, suggesting that OR commitment is downstream of initial OR specification (Dalton et al., 2013). Thus, OR specification and commitment are two separate processes that may be difficult to distinguish in vivo by an existing model that is static, insufficient and focuses only on monoallelic activation (Tan et al., 2013; Tian et al., 2016; Vyas et al., 2017; reviewed in Lomvardas and Maniatis, 2016). Therefore, a dynamic and a comprehensive model for OR singularity independent of feedback inhibition is warranted. The OP6 cell line represents an intermediate stage along the OR activation pathway, where OR selection has occurred but commitment to the chosen OR is not realized, likely due to nonexecution of the feedback inhibition process, resulting in persistent LSD1 expression and frequent OR switching (Illing et al., 2002; Pathak et al., 2009; Noble et al., 2018).

We have shown that LSD1 is not required for de novo OR activation during switching and that LSD1 depletion leads to gradual and systematic accumulation of H3K4 methylation at OR loci. This is consistent with a role of LSD1 in maintaining low H3K4 methylation levels at OR loci, presumably as a mechanism to ensure inactivity of competing ORs in the context of OP6 cells.
(Vyas et al., 2017). We propose that initial OR selection is an iterative and dynamic process (as opposed to the static process proposed by the Lomvardas group), orchestrated by equilibrium between histone methyltransferase (e.g. G9a, GLP) and histone demethylase (e.g. LSD1) activities that create an epigenetic landscape favorable to the global repression and singular OR activation. Our model predicts that management of both H3K9 methylation (by enzymes that add these marks at ORs e.g. G9a and GLP) and H3K4 methylation (by an enzyme that removes these marks at ORs e.g. LSD1) are required for a winnowing down process (i.e. reducing the number of competing OR candidate genes for activation). The model also predicts that LSD1-G9a double knockout will produce an even more enhanced phenotype (e.g., an even greater number of ORs expressed per cell). Here, I would proposed experiments that further investigate the epigenetic regulation of ORs including the role of G9a prior to the de-differentiation phenotype (Section 4.3.A), and the impact of LSD1 and G9a double knock-down and LSD1 overexpression on OR gene regulation (Section 4.3.B), I would also investigate additional provocative mechanisms of the iterative epigenetic model of OR gene selection in LSD1 KD background such as the stability of DNMT1, DNA hypomethylation, the increase in global histone acetylation, the significant down-regulation of the LSD1-corepressor protein, CoREST, and the significant up-regulation of the matrix-associated lamin B receptor. Together, all of the data gathered on these topics so far suggest that LSD1 may play a diverse role in modulating DNA methylation and nuclear
architecture, all of which have been shown to be important in the context of OR gene regulation. (Section 4.3.C).

4.3.A Does G9a Perturbation Impact OR Regulation in OP6 Cells?

G9a is a histone methyl transferase that adds an H3K9me2/1 repressive histone mark at a variety of genes in various developmental contexts (Tachibana et al., 2002, 2005; Chaturvedi et al., 2012; reviewed in Sankaran et al., 2013), including at OR loci (Lyons et al., 2014). One study suggests that the G9a/GLP mediated addition of repressive marks ensures that all OR clusters acquire a similar silenced ground state (Magklara et al., 2011) after which singular OR gene derepression occurs to activate a particular OR in each developing neuron (Lyons et al., 2013; Lomvardas and Maniatis, 2016). We propose that G9a and LSD1 function collaboratively to retain an epigenetic landscape that is significantly biased towards ubiquitous H3K9 and rare H3K4 incidences at OR loci (Vyas et al., 2017). A testable prediction of this model is that perturbation of either enzyme will produce a multi-genic phenotype, while perturbation of both LSD1 and G9a (double knockdown) may exaggerate a multigenic OR phenotype. For example, perturbation of LSD1/G9a may cause a shift in the equilibrium of epigenetic states across OR loci such that H3K9 methylation becomes less dominant and H3K4 methylation becomes less rare, thereby increasing the probability of multiple OR activation events per cell. To further test this model, I performed a G9a knock-down (KD) using the CRISPR-Cas9 approach and
observed transformation of the OP6 cells at 72 hours post-treatment as discussed in Chapter 3, suggesting G9a controls other developmental genes within the OP6 cellular context as well. We began this project with the goal of investigating the impact of G9a depletion on OR gene regulation, however, we ended up with the interesting transformation phenotype described in Chapter 3. The OP6 transformation in G9a depleted cells occurs 72 hours post-CRISPR treatment, therefore, we have a limited window of opportunity to study OR gene regulation before the OP6 cells begin to change their fate. We suggest that the GFP+ cells be sorted within 24-48 hours post-CRISPR treatment and RNA FISH should be performed using pooled-OR RNA FISH probes to compare the OR expression phenotype in empty vector control versus G9a KD GFP+ cells. We predict that depletion of G9a in OP6 cells may lead to a multigenic-multiallelic phenotype due to the lack of heterochromatic histone marks (e.g. H3K9me3) and that many GFP+ cells would exhibit multi-spot RNA FISH signals.

Previous work suggests that depletion of G9a reduces OR expression levels and H3K9 methylation levels at OR loci, in addition to violating the “one receptor per neuron” rule in vivo (Lyons et al., 2014). To investigate the impact of G9a KD on expression levels and H3K9 methylation status at OR loci, we suggest to perform RT-qPCR and ChIP to correlate OR expression and epigenetic status prior to cellular transformation. We predict three outcomes: 1) In the absence of G9a, ORs become devoid of repressive epigenetic marks (e.g. H3K9methylation) that can be measured by H3K9me ChIP. 2) ORs are
upregulated in G9a KD versus control OP6 cells that can be measured by RT-qPCR assays. 3) We also predict that due to limiting factors (e.g. lack of enhancer clustering, transcription factor abundance) OR expression levels may not increase rather multiple ORs may express at low levels in each cell. This can be detected by pooled OR RNA FISH assay as discussed above.

Currently, we do not know whether a cell cycle progression is required for OP6 cell transformation. It may be possible that OP6 cell transformation is a combinatorial effect of the cell cycle and G9a KD. Therefore, one strategy to study the impact of G9a KD on OR gene regulation would be to perform G9a KD on non-cycling, differentiated OP6 cells. We predict that due to the lack of cell cycle, differentiated OP6 cells would not undergo transformation upon depletion of G9a and may exhibit a multigenic OR phenotype. Still, we do not know whether the developmental genes or the olfactory genes are the primary targets for G9a in cycling OP6 cells, meaning we don’t know which genes are the first to be affected by the G9a depletion in undifferentiated OP6 cells? However, a lag in the OP6 cell transformation phenotype that was evident around 72 hours post-CRISPR treatment, meaning a constitutive G9a sgRNA expression was required for the transformation phenotype. Therefore, a TetO-inducible G9a knock-down strategy would minimize the prolonged impact/severity on OP6 cells. The TetO-inducible system would pulse a G9a shRNA dose only in the presence of doxycycline (tetracycline analogue) that will allow only transient downregulation of G9a (e.g. <48 hours). This transient G9a depletion may be
sufficient to impact OR genes but not developmental genes. Thus, the inducible G9a KD strategy has the potential to prevent cellular transformation but still may be able to exhibit a multigenic-OR phenotype that can be investigated by pooled OR RNA FISH assay.

4.3.B Does LSD1 Overexpression Impact OR Regulation in OP6 Cells?

The levels of LSD1 affects the kinetic of OR choice and OSN maturation (Lyons et al., 2013; Tan et al., 2013). For example, a transient and low levels of LSD1 expression is suitable for an activation event at a single OR locus per cell (via H3K9 demethylation) (Lyons et al., 2013; Tian et al., 2016). Furthermore, Adcy3 stabilizes OR choice via downregulating LSD1 (via feedback inhibition) (Lyons et al., 2013; Dalton et al., 2013). The depletion of Adcy3 causes sustained LSD1 expression longer than WT OSNs (due to the absence of feedback inhibition). This presence of LSD1 after the initial de-repression event deactivates the previously chosen OR (via H3K4 demethylation) and de-silences the newly chosen OR allele (via H3K9 demethylation), therefore increasing the OR gene switching frequency (Lyons et al., 2013). Moreover, depletion of Adcy3 is correlated with the perturbation of glomeruli targeting (Chester et al., 2007; Col et al., 2007; Zou et al., 2007), probably due to the multi-OR phenotype.

If LSD1 functions in the H3K9 demethylation pathway of ORs activation, then the overexpression of LSD1 in the OP6 cell line should lead to the following two predictions: 1) an increase in OR switching frequency, which can be
detected by clonal expansion from single LSD1 overexpressing founder OP6 cells, followed by degenerate PCR and sequencing, and 2) multiple de-repression events leading to a multigenic OR phenotype, which can be detected by pooled OR RNA FISH. If LSD1 functions in the H3K4 demethylation pathway as shown in Chapter 2, then overexpression of LSD1 leads to the silencing of competing ORs and behave just like WT OP6 cells (i.e. monogenic), which can be detected by RNA FISH and ChIP assays. If LSD1 functions in both H3K9 and H3K4 demethylation pathways, then we predict increased OR gene switching but not multi-OR phenotypes in an OP6 cellular context.

We overexpressed LSD1 in control OP6 cells using lentivirus approach described in chapter 2 (section 2.2.B). LSD1 over expression vector (pHAGE-CMV-LSD1-Flag-HA) was a gift from Mosammaparast laboratory at Washington University, St. Louis. We overexpressed LSD1 in OP6 cells to study the impact of levels of LSD1 on OR gene regulation. We also overexpressed LSD1 in LSD1 KD OP6 cells, to rescue the multi-OR phenotype observed in the LSD1 KD background. To enrich for LSD1 overexpression population, we propose to expand cultures using a blasticidin marker for positive selection. We performed an immunofluorescence assay using anti-HA antibody to confirm exogenous LSD1 expression (Figure 5). We will use this LSD1 overexpression cell line for further OR gene expression analysis.
**Figure 5: Overexpression of LSD1 in OP6 and LSD1 KD Cells**
A representative image of anti-HA IF (red) showing the efficiency of exogenous LSD1 expression in OP6 (A), the intensity of overexpression in OP6 cells (B), and that in LSD1 KD cells (C). Normal expression of LSD1 in untreated control OP6 cell (D). Compartmentalized LSD1 phenotypes (white arrows) poly-punctate (E) and mono-punctate (F), recapitulated by LSD1 overexpression. M phase nucleus showing exclusion of LSD1 from the condensed chromatin (G). A representative stem-like colony found in LSD1 overexpression culture (H). DAPI counterstain is shown in blue. (Scale bar = 5µm, A-G; 10µm, H)

We also recapitulated consolidated LSD1 compartments (mono- and poly- punctate LSD1 phenotypes) that were shown to co-localize with one or a small number of ORs to restrict the LSD1 demethylase activity or to sequester OR regulators in this compartment (Kilinc et al., 2016). Additionally, we observed that LSD1 is excluded from the condensed chromatin during the M phase of the cell cycle. The M phase specific gene expression changes were observed during the cell cycle progression in embryonic cell lines (Nair et al., 2012; Blobel et al., 2009). Surprisingly, we found that LSD1 overexpression leads to OP6 cells transformation into stem cell-like colonies that phenocopy G9a OP6
cellular transformation, suggesting both LSD1 and G9a may function in the same dedifferentiation pathway. The LSD1 overexpressing OP6 cell line can be used to investigate OR phenotypes as well as to pull down LSD1 interacting partner proteins using an IP- pulldown assay.

4.3.C Additional Predictions of an Iterative Epigenetic Model for OR Selection

Substantial DNA methylation patterning has been observed at OR loci between immediate progenitor cells and mature OSNs (Colquitt et al., 2013). DNMTs are non-histone targets of LSD1 (Wang et al., 2009) and are expressed in OP cells (Colquitt et al., 2013). It is possible that LSD1 plays a role in maintaining DNA methylation via stabilizing DNMT1 during OSN differentiation, as it does in ES cells (Clements et al., 2012; Grayson and Guidotti et al., 2013). DNMTs are also shown to interact with LSD1 at a subset of gene promoters where they function to silence genes independent of their DNA methylation activity (Clements et al., 2012). DNA methylation at an OR enhancer (called the H-region enhancer) is important for maintaining the monoallelic expression nature of the OSNs (Lomvardas et al., 2006). Our hypothesis is that LSD1-dependent DNA methylation may contribute to the silencing of competing ORs; we predict that LSD1-depleted cells will exhibit: 1) down-regulation of DNMTs and 2) global hypo-DNA methylation. Preliminarily, we observed a decrease in DNA
methyltransferase 1 (DNMT1) protein in LSD1 KD cells by western blot (Figure 6). This data is consistent with previous publications (Wang et al., 2008).

Using IF, we tested the global impact of DNA methylation in LSD1-depleted cells, and found a significant reduction (~2.75 fold, consistent with published data) in total intensities measured for both 5mC and 5hmC in LSD1 KD versus control populations (Figure 6). This phenotype was even more severe (~3.75 fold) in the vicinity of nuclear chromocenters, where OR genes and other heterochromatic loci tend to reside. We propose to use bisulfite sequencing methods to specifically measure DNA methylation changes at OR loci. We predict that the OR loci may exhibit reduced DNA methylation, if DNA methylation functions as a barrier to OR activation. Thus, the OP6 cell line may be a good model system to study co-operativity between histone modification and DNA methylation at OR loci.
LSD1 and CoREST proteins are interdependent for their stability in vivo (Shi et al., 2004; Foster et al., 2010). We investigated whether LSD1 interacts with the CoREST complex and whether down-regulation of LSD1 in KD cells results in down-regulation of CoREST. We analyzed LSD1 (n=31) KD and scramble control (n=51) individual cell nuclei for LSD1/CoREST protein expression pattern. We observed a robust reduction in CoREST expression in LSD1 KD cells, confirming that there is a co-regulation of these two protein components (Figure 7A). We will next investigate whether these rare LSD1-CoREST complexes have assembled at OR loci, using DNA FISH and ChIP assays.

Histone deacetylase (HDAC) proteins interact with the LSD1 protein as a component of Co-REST complex in vivo, where they function to deacetylate histone targets (Foster et al., 2010). We tested a possibility that loss of LSD1 may influence the stability of HDACs. Our preliminary western blot data suggests a global increase in histone acetylation presumably via inhibition of HDAC activity.
(Figure 7B), consistent with a loss of CoREST, which contains HDAC components (You et al., 2001). We suggest investigation of how the above provocative observations may impact OR gene regulation.

Lamin B receptors (LBR) are located at the inner nuclear envelope in immature OSNs, where they function as a scaffolding protein to tether a heterochromatic domain called lamin associated domains (LADs) (Steensel and Belmont, 2017). LBR downregulation allows ORs to aggregate in the center of the OSN nucleus forming heterochromatic foci (chromocenters) in the mature OSNs (Clowney et al., 2012). The ectopic expression of LBR disrupts heterochromatic OR foci and causes multigenic OR expression in mature OSNs (Clowney et al., 2012). Since LSD1 KD leads to multigenic OR expression phenotype, we wondered whether LSD1 KD may impact LBR expression in OP6 cells. Surprisingly, we found that LBR protein but not mRNA was up-regulated in LSD1 KD cells as measured by IF and single cell RT-PCR (Figure 8).
Figure 7. Impact of LSD1 Knockdown on Co-REST Protein Levels in OP6 Cells.

(A) Left Panels: LSD1 and CoREST immunofluorescence in LSD1-knockdown (lower two images) and scrambled control cells (upper two images). Right Panels: Average Co-REST staining intensities (top histogram) and average number of distinct Co-REST foci (bottom histogram) in scrambled control (Scr) and LSD1-knockdown (KD) cells. (B) Left Panel: Western blot showing total H4 acetylation levels in two independent LSD1-knockdown populations (KD1, KD2) as compared to the scrambled control (Scr) population; actin blots are shown as a loading control. Right Panel: Quantitated western band intensities for total H4 acetylation in the scrambled control sample (Scr) versus the average of the two LSD1-knockdown (KD) blots (one standard deviation indicated).
Figure 8. Impact of LSD1 Knockdown on Lamin-B Receptor (LBR) Protein Levels in OP6 Cells.

(A) LBR immunofluorescence illustrating a bimodal distribution of LBR expression in LSD1-knockdown cells (KD), with ~24% of cells exhibiting lower levels of LBR staining (<2x DAPI staining level, exemplified by upper rightmost cell in the KD image, comparable to levels observed in scrambled controls, left image) and ~76% of cells exhibiting significantly increased levels of LBR staining (>2x DAPI staining level, exemplified by the other cell in the KD image).

(B) Summary of LBR staining intensities showing bimodal distribution in KD cells (white bars, left histogram) and a large standard deviation in the average intensity (right histogram), and a higher overall level of LBR staining as compared to scrambled controls (black bars).

(C) RT-PCR results suggesting that LBR mRNA is not significantly different between two knockdown populations (KD1, KD2) as compared to two scrambled control populations (Scr1, Scr2). GAPDH RT-PCR on these same populations is shown as a loading control.
4.4 Summary of the Thesis

In my thesis work, I discovered two novel functions of LSD1 in the context of the OP6 cell line: 1) Depletion of LSD1 perturbed monogenic/monoallelic OR expression; 2) Overexpression of LSD1 dedifferentiated OP6 cells into stem-like colonies, which appeared to phenocopy G9a depletion in this cell line. We proposed that LSD1 may not be involved in initial OR gene selection, since clonal expansion of LSD1-depleted OP6 cells did not perturb de-novo OR activation, a result not predicted if LSD1 function is required in the OR activation pathway. Moreover, in vivo studies add further evidence that LSD1 is not likely to be involved with establishing or maintaining singular OR expression since ORs remain active in LSD1 knockout mice (Lyons et al., 2013). Since LSD1 possesses a dual substrate specificity to demethylate both H3K4 (causing de-activation) and H3K9 (causing de-repression) (Shi et al., 2004; Metzger et al., 2005), we speculate that LSD1 might contribute to both suppression and activation of OR genes during OSN differentiation. Of note, depletion of LSD1 in OSN progenitors (GBCs and HBCs) leads to an OSN differentiation defect, suggesting that LSD1 activity is important for both OR regulation and OSN differentiation (Coleman et al., 2018; Nagai et al., 2016).

The G9a protein, like LSD1, may also be tied to both OR regulation and OSN differentiation. In Chapter 3 of my thesis work, we found that depletion of G9a allows neuronally committed OP6 cells to apparently dedifferentiate and
give rise to stem-cell like colonies with potential to differentiate into various cell types. We show that this transformation was correlated with up-regulation of pluripotency transcription factor, Sox2, a result that is reminiscent of dedifferentiation of cycling, immature neuronal progenitors in vivo (Lin et al., 2017). Furthermore, heterochromatic odorant receptor (OR) gene silencing is mediated by histone methyl transferase (G9a) and loss of G9a leads to multigenic OR expression in the OE (Lyons et al., 2014). Therefore, there is evidence for G9a functioning in both silencing OR genes and, potentially, silencing genes important for maintaining stem-ness.

While beyond the scope of this thesis, the roles of LSD1 and G9a in both OR regulation and in establishing or maintaining differentiated states along the OSN lineage raises an interesting question: are the two processes intimately connected? It is interesting to note that most of the published reports in which the onset of OR expression is perturbed also results in diminished or lost OSN neuronal populations in vivo (Iwema and Schwob, 2003; Shykind et al., 2004; Rodriguez-Gil et al., 2015; Lyons et al., 2013; Coleman et al., 2018). A parsimonious interpretation is that chromatin “writers”, like G9a and “erasers” like LSD1, collaborate to ensure that progression from stem-precursors to a post-mitotic, committed immature OSN fate during differentiation is coupled to the specification of OSN identity via mutually exclusive OR expression. For example, G9a/LSD1 complexes activated during OSN differentiation might function to repress stem markers and cell cycling enablers to enable progression
to a post-mitotic immature OSN and simultaneously function to initiate an OR selection mechanism that ensures the new OSN will be functional within the olfactory system. In this way, there might be an intimate connection between what at first glance appears to be two very different experimental results described in my thesis: an OR regulation story (Chapter 2) and an OSN development story (Chapter 3). One of the opportunities with RNA-seq studies proposed in Chapter 4 will be to more deeply investigate the interconnected regulatory networks that become activated/deactivated in the context of LSD1/G9a inhibition or over-expression.

References


Chen, J., Liu, H., Liu, J., Qi, J., Wei, B., Yang, J., . . . Pei, D. (2013). H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat Genet, 45*(1), 34-42. doi:10.1038/ng.2491


Rego, A., Sinclair, P. B., Tao, W., Kireev, I., & Belmont, A. S. (2008). The facultative heterochromatin of the inactive X chromosome has a distinctive


