UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE QUÍMICA

Programa de Pós-Graduação em Ciências Biológicas (Bioquímica)

ARTUR GUAZZELLI LEME SILVA

**A influência de polimorfismos de base única
na metilação de DNA em genes de
receptores olfatórios**

Versão corrigida da Tese defendida

São Paulo

Data do depósito na SPG:

28/02/2018

ARTUR GUAZZELLI LEME SILVA

**A influência de polimorfismos de base única
na metilação de DNA em genes de
receptores olfatórios**

*Tese apresentada ao Instituto de Química da
Universidade de São Paulo para a obtenção do
Título de Doutor em Ciências (Bioquímica).*

Orientadora: Profa. Dra. Bettina Malnic

São Paulo

2018

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha Catalográfica elaborada eletronicamente pelo autor, utilizando o programa desenvolvido pela Seção Técnica de Informática do ICMC/USP e adaptado para a Divisão de Biblioteca e Documentação do Conjunto das Químicas da USP

Bibliotecária responsável pela orientação de catalogação da publicação:
Marlene Aparecida Vieira - CRB - 8/5562

S586i Silva, Artur Guazzelli Leme Silva
A influência de polimorfismos de base única na metilação de DNA em genes de receptores olfatórios / Artur Guazzelli Leme Silva Silva. - São Paulo, 2018. 83 p.

Tese (doutorado) - Instituto de Química da Universidade de São Paulo. Departamento de Bioquímica.
Orientador: Malnic, Bettina Malnic

1. Epitélio olfatório . 2. Genes de receptores olfatórios. 3. Metilação de DNA. 4. Polimorfismo de base única (SNP). 5. Variabilidade genética. I. T. II. Malnic, Bettina Malnic, orientador.

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

"A influência de polimorfismos de base única na metilação de DNA em genes de receptores olfatórios"

ARTUR GUAZZELLI LEME SILVA

Tese de Doutorado submetida ao Instituto de Química da Universidade de São Paulo como parte dos requisitos necessários à obtenção do grau de Doutor em Ciências - Área: Bioquímica.

Aprovado (a) por:

Prof. Dra. Bettina Malnic
(Orientadora e Presidente)

Prof. Dra. Carla Columbano de Oliveira
IQ - USP

Prof. Dr. Alexandre Bruni Cardoso
IQ - USP

Prof. Dra. Lygia da Veiga Pereira Carramaschi
IB - USP

Prof. Dr. Fabio Papes
IB - UNICAMP

SÃO PAULO
24 de abril de 2018

*Dedico este trabalho a todos aqueles que
de alguma forma me incentivaram e me
inspiraram na busca pelo conhecimento.*

AGRADECIMENTOS

Agradeço a toda minha família que de alguma forma contribuiu para minha formação, especialmente à minha mãe pelo carinho, zelo, broncas e pelo ensinamento de valores importantes para vida, e ao meu pai que me ensinou a ter empenho e dedicação em tudo que eu faço. A meu irmão pela companhia, trocas infinitas de experiência e por me inspirar na busca dos meus objetivos. Agradeço a meus avós, em especial a minha avó Yvone e meu avô Jacomo pelo apoio e serenidade infinita que sempre foram exemplos para mim.

À Profa. Dra. Bettina Malnic por confiar a mim a realização desse projeto, pela valorosa orientação ao longo de todo o doutorado, por sua disposição e disponibilidade para discutir os rumos da pesquisa e das recentes descobertas científicas, contribuindo imensamente para minha formação acadêmica.

À Dra. Máira Harume Nagai por sua inestimável paciência, sabedoria e empenho que muito me inspiram e contribuíram para meu desenvolvimento científico. À Dra. Lúcia Maria Armelin-Corrêa pelos ensinamentos e discussões sempre produtivas e estimulantes. Aos demais colegas de laboratório Cassandra S. de Lyra, Cleiton F. Machado, João B. P. do Nascimento, Luiz E. R. Firmino, Tiago J. de Almeida e Victor P. de Sá Xavier pela convivência harmônica e profundamente enriquecedora. À Érica Michelle Rodrigues Bandeira por todo o auxílio técnico e burocrático, e companheirismo do dia-a-dia. Ao Edson Alves Gomes pelo auxílio técnico e sempre instrutiva troca de ideias. À Luci Navarro pela prontidão e eficiência no controle do equipamento de sequenciamento de DNA. Ao Fernando Delgado Pretel (técnico do CEFAP-USP/FLUIR) pelo auxílio na operação do citômetro de fluxo para separação dos neurônios GFP+. A todos os funcionários do biotério pelo auxílio no cuidado dos animais de experimentação.

A todos os meus professores pela inestimável contribuição para o meu desenvolvimento intelectual, inspiração para a vida e incentivo na busca constante pelo conhecimento.

A todos os colegas que tive a oportunidade de conhecer ao longo da minha graduação, em especial a meu amigo e cunhado Arcelino Bezerra da Silva Neto pelo companheirismo, motivação e uma convivência enriquecedora. Ao Charles Roberto de Almeida Silva pela parceria e no compartilhamento de ideias e reflexões aprofundadas sobre o mundo. Ao Danilo Oliveira dos Santos pela imensa paciência e por sua forma de contestar as contradições sempre acompanhada com uma elegante ironia. À Fernanda Conceição Matos pelas lições de responsabilidade e visão humanista do mundo.

A todos que não foram citados nominalmente, mas que de alguma forma contribuíram para que eu pudesse cumprir essa etapa da minha vida.

À Andreza Cristina Souza Silva meus mais profundos agradecimentos, por sua cumplicidade, alegria, carinho peculiar, companheirismo e constante aprendizado na vida que decidimos compartilhar. Pelo seu empenho em sempre estar se aprimorando e coragem nos momentos de mudança, que sempre me inspiram e motivam. Seguramente, sua contribuição foi fundamental para a realização desse trabalho.

Finalmente, agradeço à CAPES e a FAPESP pelo apoio financeiro.

*O mais incompreensível do mundo é
que ele seja compreensível.*

(Albert Einstein)

*Se cheguei até aqui foi porque me
apoiei no ombro dos gigantes.*

(Isaac Newton)

RESUMO

Silva, A.G.L. **A influência de polimorfismos de base única na metilação de DNA em genes de receptores olfatórios.** 2018. 83 páginas. Tese de Doutorado – Programa de Pós-Graduação em Ciências Biológicas (Bioquímica). Instituto de Química, Universidade de São Paulo, São Paulo.

Os genes de receptores olfatórios (OR) pertencem a uma família de proteínas de membrana formada por cerca de 1000 genes no genoma de camundongo. Os genes OR são expressos de forma monogênica e monoalélica nos neurônios olfatórios (OSNs). No entanto, ainda não está claro o mecanismo que permite essa forma de expressão peculiar, sobretudo, qual o papel da metilação de DNA nesse processo. Nosso estudo determinou o padrão de metilação de DNA da região promotora e codificadora do gene *Olfr17*. Em células de epitélio olfatório (MOE) de camundongos adultos, observamos na região codificadora (CDS) do gene uma frequência de metilação em dinucleotídeos CpG $\geq 58\%$, enquanto que na sua região promotora ela foi bem mais baixa. Os níveis de metilação do *Olfr17* em MOE de embrião (E15.5) e fígado foram similares aos observados em MOE de animais adultos. Em seguida, analisamos se a metilação de DNA pode regular a expressão gênica do *Olfr17*. Utilizando animais transgênicos onde os neurônios olfatórios que expressam *Olfr17* também expressam GFP, pudemos selecionar neurônios olfatórios GFP+ e analisar a metilação do gene *Olfr17*, que está ativo nestas células. Verificamos que o padrão geral de metilação do *Olfr17*, tanto na região CDS como na região promotora, não se altera quando este gene está ativo. Este resultado indica que alterações na metilação do gene *Olfr17* não são necessárias para que este receptor seja expresso. Finalmente, verificamos que a região promotora do gene *Olfr17*, de duas linhagens de camundongos diferentes, a C57BL/6 e a 129, possuem dois polimorfismos de base única (SNPs) que alteram o conteúdo CpG. Devido a estes SNPs, a linhagem 129 apresenta dois sítios CpG adicionais, inexistentes na linhagem C57BL/6. Nossas análises mostraram que estes CpGs são frequentemente metilados, o que torna o promotor do *Olfr17* de 129 significativamente mais metilado que o promotor de C57BL/6. Em seguida, nós analisamos o nível de expressão no MOE dos dois alelos de *Olfr17*, o 129 e o C57BL/6, utilizando ensaios de RT-qPCR. Estes experimentos demonstraram que o nível de expressão do alelo 129, que possui 3 CpGs metiladas em seu promotor, é menor que o do alelo C57BL/6, que apresenta apenas uma CpG que é pouco metilada em seu promotor. Nossos resultados sugerem que as alterações na região promotora influenciam a probabilidade com que o gene OR é escolhido para ser expresso no MOE.

Palavra-chave: Metilação de DNA, genes de receptores olfatórios, sequenciamento pelo método de bissulfito, variação genética, expressão gênica, polimorfismo de base única.

ABSTRACT

Silva, A.G.L. **Single nucleotide polymorphisms lead to differential DNA methylation in odorant receptor genes.** 2018. 83 pages. PhD Thesis - Graduate Program in Biochemistry. Institute of Chemistry, University of São Paulo, Sao Paulo.

Olfactory receptor (OR) genes belong to a large family of membrane proteins composed of ≈ 1000 genes in the mouse genome. The OR genes are expressed in the olfactory sensory neurons (OSNs) in a monogenic and monoallelic fashion. However, the mechanisms that govern OR gene expression are unclear. Here we asked whether DNA methylation plays a role in the regulation of OR gene expression. We first determined the DNA methylation pattern in the coding (CDS) and promoter regions of the odorant receptor gene *Olfr17*. In olfactory epithelium (MOE) cells, the CpG methylation level in the CDS is $\geq 58\%$ but is much lower in the promoter region of the gene. In embryonic MOE (E15.5) and liver, the levels of *Olfr17* DNA methylation are similar to the ones shown in adult MOE. We next analyzed whether DNA methylation is involved in *Olfr17* regulation. We isolated GFP+ neurons from transgenic mice that coexpress GFP with *Olfr17*, and analyzed the DNA methylation pattern of the *Olfr17*, which is active in these cells. We found that the general methylation pattern, both, in the coding and promoter regions is not altered in the active gene. These results indicate that changes in DNA methylation are not required for the activation of *Olfr17*. Finally, we found that the *Olfr17* promoter region from two different mouse strains, C57BL/6 and 129, has two single-nucleotide polymorphisms (SNPs) that alter the CpG content. The SNPs lead to the existence of two additional CpGs in the 129 allele, which are absent in the C57BL/6 allele. These CpGs are frequently methylated, making the 129 *Olfr17* promoter significantly more methylated than the *Olfr17* promoter from C57BL/6. We next performed RT-qPCR experiments to analyze the expression levels of the 129 and C57BL/6 *Olfr17* alleles in the MOE. These experiments showed that the expression level of the 129 *Olfr17* allele, which contains three methylated CpGs in its promoter region, is lower than the one from C57BL/6, which contains only one, undermethylated CpG, in its promoter. Our results suggest that these promoter modifications regulate the probability of the OR gene choice.

Keywords: DNA methylation, olfactory receptor gene, bisulfite sequencing genetic variation, gene expression, single nucleotide polymorphism.

LISTA DE ABREVIATURAS E SIGLAS

5hmC – 5-hidroximetilcitosina
5mC – 5-metilcitosina
Adcy3 – adenilil ciclase 3
AOB (*accessory olfactory bulb*) – bulbo olfatório acessório
ATF5 (*activating transcription factor 5*) – fator de ativação da transcrição 5
BER (*base excision repair*) – reparo por excisão de bases
CDS (*coding DNA sequence*) – região codificadora
CGI – ilha CpG
DAPI – cloreto de 4',6-diamino-2-fenilindol
Dnmt – DNA metiltransferase
FACS (*Fluorescence-activated cell sorting*) – separação de células ativada por fluorescência
FPKM (*fragments per kilobase million*) – fragmentos por kilobase por milhão
FSC-A (*forward scatter-area*) – área do sinal gerado pela luz espalhada à frente
GFP (*green fluorescence protein*) – proteína fluorescente verde
GG – gânglio de Grueneberg
GPCR – receptor acoplado a proteína G
HB (*heterochromatic block*) – bloco de heterocromatina
HD – homeodomínio-like
iOSN (*immature olfactory sensory neuron*) – neurônio olfatório imaturo
IRES (*internal ribosome entry site*) – sítio interno de entrada do ribossomo
KO – *knockout*
LCR (*locus control region*) – região de controle do locus.
LSD1 (*Lysine-specific demethylase 1*) – demetilase lisina-específica 1
MBD (*methyl binding domain*) – domínio de ligação a grupo metil
MOB (*main olfactory bulb*) – bulbo olfatório principal
MOE (*main olfactory epithelium*) – epitélio olfatório principal
Olfr17 – receptor olfatório 17
OMP (*olfactory marker protein*) – proteína marcadora de neurônio olfatório maduro
OR (*olfactory receptor*) – receptor olfatório
OSN (*olfactory sensory neuron*) – neurônio olfatório
PCR (*polymerase chain reaction*) – reação em cadeia da polimerase
RT-PCR – transcrição reversa seguido de PCR
SO – órgão do Septo de Masera
TAAR (*Trace amine-associated receptor*) – receptor associado a traços de aminas
TFBS (*transcription factor binding sites*) – sítios de ligação para fatores de transcrição
TSS (*transcription start site*) – sítio de início da transcrição.
UPR (*unfolded protein response*) – resposta a proteínas desenoveladas
VNO (*vomer nasal organ*) – órgão vomeronasal

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	Sistema olfatório.....	15
1.2	Aspectos regulatórios da expressão dos genes de receptores olfatórios	17
1.3	Aspectos gerais sobre metilação de DNA.....	24
1.4	Metilação de DNA no sistema olfatório	28
2	OBJETIVOS.....	31
3	MATERIAL E MÉTODOS.....	33
3.1	Análise do conteúdo de CpG em regiões promotoras e codificadoras de genes OR ..	33
3.2	Determinação do padrão de metilação pelo método de bissulfito (DNAbiss-seq).....	35
3.2.1	Animais de experimentação.....	36
3.2.2	Extração de DNA genômico (gDNA).....	36
3.2.3	Tratamento de DNA com bissulfito de sódio	36
3.2.4	Precipitação do DNA tratado com bissulfito de sódio (DNAbiss).....	38
3.2.5	PCR com DNAbiss	38
3.2.6	Reação de ligação de produto de PCR em pCRII.....	40
3.2.7	Transformação de bactérias competentes	40
3.2.8	Purificação de plasmídeo de minicultura bacteriana	41
3.2.9	Reação de sequenciamento	41
3.2.10	Purificação de DNA de reação de sequenciamento.....	41
3.3	Busca de TFBS na região promotora do gene <i>Olfr17</i>	42
3.4	Separação de células ativadas por fluorescência (FACS).....	42
3.4.1	Linhagem transgênica P2-IRES-tauGFP	42
3.4.2	Genotipagem da linhagem P2-IRES-tauGFP.....	42
3.4.3	Dissociação celular.....	44
3.4.4	Separação de células ativada por fluorescência (FACS)	45
3.5	RT-PCR e PCR em tempo real.....	46
3.5.1	Extração de RNA com reagente TRIzol	46
3.5.2	Síntese de DNA complementar (cDNA)	46
3.5.3	PCR quantitativo (qPCR).....	47

4	Resultados.....	49
4.1	Conteúdo CpG em genes OR	49
4.2	Teste de eficiência do método DNAbiss-seq.....	52
4.3	Metilação de DNA no gene de receptor olfatório <i>Olfr17</i>	54
4.4	Metilação de DNA em outros genes OR	56
4.5	SNPs alteram o conteúdo CpG na região promotora de <i>Olfr17</i>	58
4.6	O padrão de metilação se altera com o conteúdo CpG.....	60
4.7	Expressão do gene <i>Olfr17</i> em diferentes contextos genéticos.....	63
5	Discussão	67
5.1	Aspectos genéticos e expressão dos genes OR	67
5.2	Metilação de DNA e expressão dos genes OR.	69
5.3	Contexto genômico e variabilidade na expressão dos genes OR.....	72
6	Conclusões	75
7	Referências.....	77
8	ANEXOS.....	85
8.1	Anexo A – Súmula curricular.....	85
8.2	Anexo B – Artigo em preparação e submetido a disciplina QBQ5749 – Redação Científica no 1º semestre de 2017.	88
8.3	Anexo C – Capítulo para <i>Methods in Molecular Biology on Olfactory Receptors</i> , submetido para avaliação.	111

ÍNDICE DE FIGURAS

Figura 1.1. Subistemas olfatórios.	15
Figura 1.2. Anatomia do sistema olfatório de roedores.	18
Figura 1.3. Representação esquemática das principais etapas da escolha do gene de OR.	23
Figura 1.4. A metilação de DNA ocorre frequentemente em citosina.	26
Figura 1.5. Expressão de proteínas relacionadas à metilação de DNA ao longo do desenvolvimento do Epitélio Olfatório.	29
Figura 3.1. Fluxograma de trabalho para a análise de metilação do DNA.	35
Figura 3.2. Deaminação seletiva de citosina com bissulfito de sódio.	37
Figura 3.3. Representação esquemática do método de conversão com bissulfito de sódio.	37
Figura 3.4. Representação esquemática do alelo transgênico da linhagem P2-IRES-tauGFP.	43
Figura 3.5. Representação esquemática de um equipamento de FACS	45
Figura 4.1. Distribuição do número de CpGs na região promotora de genes OR.	49
Figura 4.2. Ocorrência de dinucleotídeos CpG em 198 regiões promotoras de genes OR.	50
Figura 4.3. Distribuição do número de CpGs em 864 regiões codificadoras dos genes OR.	50
Figura 4.4. Ocorrência de dinucleotídeos CpG em 864 CDSs de genes OR.	51
Figura 4.5. Distribuição de Obs/Exp CpG em promotores e regiões codificadoras de genes OR	52
Figura 4.6. Determinação da frequência de mC (%mC) num grupo de células.	53
Figura 4.7. Frequência de mCpG ao longo da região promotora do gene GAP43.	53
Figura 4.8. Representação esquemática do gene <i>Olfr17</i>	54
Figura 4.9. Frequência de metilação (%mC) no gene <i>Olfr17</i>	55
Figura 4.10. Metilação em dinucleotídeos CpA do gene <i>Olfr17</i> em três diferentes tecidos.	56
Figura 4.11. Metilação em CpG na região codificadora dos genes <i>Olfr714</i> , <i>Olfr17</i> , <i>Olfr1507</i>	57
Figura 4.12. SNPs no promotor de <i>Olfr17</i> alteram o conteúdo CpG.	58
Figura 4.13. Sítios de ligação O/E-like e homeodomínio-like na região promotora do gene <i>Olfr17</i>	59
Figura 4.14. Representação esquemática da separação das células <i>Olfr17</i> -tauGFP por FACS	60
Figura 4.15. Padrão de metilação do gene <i>Olfr17</i> em neurônios olfatórios GFP+.	61
Figura 4.16. Padrão de metilação do alelo ativo do gene <i>Olfr17</i>	63
Figura 4.17. Comparação da %mC na região promotora do alelo C57BL/6 e 129P2/OlaHsd.	64

Figura 4.18. O alelo <i>Olf17</i> da linhagem B6 é preferencialmente expresso ao alelo transgênico	66
Figura 5.1. Modelo de ativação de gene OR dependente da região promotora.....	72
Figura 5.2. O repertório de genes OR varia entre as linhagens de camundongos.....	73

ÍNDICE DE TABELAS

Tabela 3.1: Códigos OLFR dos 198 genes de receptores olfatórios correspondentes aos promotores analisados. Sequências disponibilizadas em (MICHALOSKI; GALANTE; MALNIC, 2006).	33
Tabela 3.2. Códigos MOR dos 864 genes de receptores olfatórios utilizados para a análise de conteúdo CpG.....	34
Tabela 3.3 Oligonucleotídeos para amplificação de regiões dos genes estudados.	39
Tabela 3.4. Concentração dos componentes na reação de PCR com DNAbiss.....	39
Tabela 3.5. <i>Primers</i> utilizados na genotipagem de camundongos transgênicos P2-IRES-tauGFP.	43
Tabela 3.6. Concentração dos componentes na reação de PCR com gDNA.	44

1 INTRODUÇÃO

1.1 Sistema olfatório

O olfato desempenha um papel fundamental para a sobrevivência de diversos animais. Através do contato com compostos químicos no ambiente, eles podem reconhecer animais da mesma espécie, identificar um predador ou uma presa, buscar alimento e procurar um parceiro para o acasalamento. O reconhecimento, pelo sistema olfatório (ou sistemas análogos), de uma grande diversidade de compostos químicos chamados de odorantes, ocorre em um grande número de espécies, de insetos a mamíferos (ACHE; YOUNG, 2005; FIRESTEIN, 2001; KAMBERE; LANE, 2007; LANCET, 1986). Nos mamíferos, a percepção do olfato é realizada em diversos tecidos encontrados na cavidade nasal. Em camundongos, por exemplo, são encontrados o Sistema Olfatório Principal (MOS), o Sistema Olfatório Acessório (AOS), Órgão do Septo de Masera (SO) e o Gânglio de Grueneberg (GG); Figura 1.1. Esses subsistemas podem ser caracterizados pela localização anatômica, pelos componentes celulares (receptores, proteínas de transdução de sinal, canais iônicos, etc) e pela inervação de regiões específicas nos bulbos olfatórios principal (MOB) e acessório (AOB) (MUNGER; LEINDERS-ZUFALL; ZUFALL, 2009).

Figura 1.1. Subsistemas olfatórios.

(A) Representação esquemática de uma secção sagital do nariz de um camundongo mostrando quatro subsistemas olfatórios e as enervações de cada um deles no bulbo olfatório principal (MOB) e acessório (AOB). (B) Coloração com X-gal de camundongo transgênico Omp-IRES-tau:LacZ apresentado um corte sagital correspondente a representação do item (A). Sistema Olfatório Principal (MOS), Sistema Olfatório Acessório (AOS), Órgão do Septo de Masera (SO) e o Gânglio de Grueneberg (GG). Figura adaptadas de (MUNGER et al., 2009).

O epitélio olfatório principal (MOE) é um neuroepitélio pseudoestratificado composto por quatro tipos celulares principais: células de suporte (sus), células basais (HBC), células globosas (GBC) e neurônios olfatórios (OSN) (MOULTON; BEIDLER, 1967). As células basais são células tronco que dão origem a todos os tipos celulares presentes no MOE (GRAZIADEI; GRAZIADEI, 1979). Já os neurônios olfatórios maduros são as células especializadas, que projetam seus dendritos para a cavidade nasal quando maduros, possibilitando a interação com os odorantes presentes no ambiente.

Até o início da década de 90, não estava claro como o sistema olfatório era capaz de discriminar a grande diversidade de compostos químicos presentes no meio. Nesse período, foi verificado que o produto de PCR com primers degenerados, sintetizado a partir de cDNA de epitélio olfatório principal (MOE), produziu diversos amplicons de tamanhos parecidos, o que indicava a existência de uma família gênica com um padrão de expressão tecido-específico (BUCK, L.; AXEL, 1991). As sequências amplificadas compartilhavam similaridades com uma superfamília dos receptores acopladas a proteína G (GPCRs) e, por isso, foi proposto que estes genes fariam parte de uma família de proteínas de membrana, os receptores olfatórios (ORs), cuja função seria a discriminação dos diferentes odorantes presentes no ambiente (AXEL, 1995; BUCK, L. B., 1996).

Os receptores do tipo GPCR apresentam 7 α -hélices transmembrana (TM), em que TM-I está mais próxima da porção N-terminal (extracelular) e TM-VII está mais próxima da porção C-terminal (intracelular). Além disso, a porção extracelular dos GPCRs apresenta três alças de comprimentos variados conectando as α -hélices (L2, L4 e L6) e a porção intracelular apresenta outras três alças (L1, L3 e L5). Os receptores olfatórios possuem trechos da sequência de aminoácidos conservados entre si (GAILLARD; ROUQUIER; GIORGI, 2004; PROBST et al., 1992). Uma dessas sequências está na porção que compreende o final da α -hélice TM-III e o início da alça intracelular L3, e essa sequência é importante para a ligação de uma das subunidades da proteína G heterotrimérica expressa especificamente no MOE, $G\alpha_{olf}$ (JONES, D. T.; REED, 1989). Em contrapartida, a alça L6 e os domínios TM-IV e TM-V são hipervariáveis, indicando que esses trechos das sequências de aminoácidos dos

receptores podem ser os responsáveis pela interação com grande número de compostos químicos de estruturas variadas, condizente com a diversidade de odorantes presentes no ambiente (BUCK, L.; AXEL, 1991; MOMBAERTS, 1999; PILPEL; LANCET, 1999). Quando um odorante interage com um OR ocorre uma mudança conformacional desse receptor permitindo que ele catalise a troca do GDP, ligado à $G_{\alpha_{olf}}$, por GTP. A proteína conjugada $G_{\alpha_{olf}}$ -GTP se dissocia do OR e das demais subunidades ($G\beta$ e $G\gamma$) para ativar a adenilil ciclase III (ACIII) e assim continuar o processo de transdução de sinal (BELLUSCIO et al., 1998; FIRESTEIN, 2001; WONG et al., 2000).

Partindo de sequências de OR determinadas experimentalmente, foi possível prever a existência de cerca de 1300 genes OR em camundongos. No entanto, aproximadamente 20% deles eram pseudogenes, resultando em um repertório de mais de 1000 receptores diferentes dedicados à percepção do cheiro (YOUNG et al., 2002; ZHANG, X.; FIRESTEIN, 2002). Esses genes estão dispersos por todo genoma de camundongo, com exceção dos cromossomos: 5, 12, 18 e Y. Além disso, os genes OR encontram-se peculiarmente agrupados em 51 *loci*, denominados *clusters*, com números de genes e distâncias intergênicas variáveis entre si (GODFREY; MALNIC; BUCK, 2004; SULLIVAN et al., 1996). Analogamente, foi verificado que os genes OR em humanos são distribuídos em 51 *clusters* ao longo do genoma e nenhum deles está no cromossomo Y (MALNIC; GODFREY; BUCK, 2004). Apesar dessas similaridades, mais de 60% dos genes OR humanos são pseudogenes e o número de ORs intactos é cerca de 350, um terço da quantidade encontrada em camundongos (GAILLARD et al., 2004; MALNIC et al., 2004).

1.2 Aspectos regulatórios da expressão dos genes de receptores olfatórios

Os genes OR são expressos nos neurônios olfatórios de forma monogênica e monoalélica (CHESS et al., 1994; ISHII et al., 2001; MALNIC et al., 1999; SERIZAWA et al., 2000). Em outras palavras, cada célula que se diferencia em um OSN escolhe, estocasticamente, apenas um dos genes OR dentre todos os genes dessa família, para ser expresso naquela célula. Além disso, somente um dos dois alelos (materno ou paterno) é

expresso pelo OSN. Esse peculiar padrão de expressão dá origem a um neuroepitélio com uma grande diversidade fenotípica, gerada, a grosso modo, por apenas um tipo de proteína – o OR – que diferencia um neurônio do outro. Apesar disso, os neurônios que expressam o mesmo tipo de receptor projetam seus axônios para uma mesma região no bulbo olfatório, conhecida como glomérulo (MOMBAERTS et al., 1996), criando no OB um nível de organização de ordem topológica para os diferentes estímulos percebido pelos OSNs (Figura 1.2) (RESSLER; SULLIVAN; BUCK, 1994; VASSAR et al., 1994).

Figura 1.2. Anatomia do sistema olfatório de roedores.

À esquerda é apresentada a vista sagital do nariz de um roedor ressaltando as principais estruturas envolvidas na percepção dos odorantes. À direita da imagem uma representação ampliada do epitélio olfatório (OE), dos neurônios olfatórios (OSN) e do bulbo olfatório principal (MOB) ressaltando que cada OSN expressa apenas um dos genes OR e os axônios de neurônios expressando o mesmo OR são projetados para o mesmo glomérulo no OB. VNO: órgão vomeronasal; VSN: neurônios de receptores vomeronasais; AOB: bulbo olfatório auxiliar. Adaptado de (DEMARIA; NGAI, 2010)

Há ainda outro nível de organização descrito para o MOS, em que os OSNs expressando um mesmo gene OR ocupam uma das quatro zonas do epitélio olfatório, designadas Z1, Z2, Z3 e Z4, sendo que a Z1 é localizada na região mais apical do MOE e Z4 na região mais ventral (RESSLER; SULLIVAN; BUCK, 1993; VASSAR; NGAI; AXEL, 1993). Posteriormente, foi verificado que a distribuição de neurônios que expressam o mesmo OR

poderia ocupar regiões intermediárias entre as zonas, mas se mantendo em uma região contínua do MOE (MIYAMICHI et al., 2005). O padrão de expressão em zonas se reflete na posição dos glomérulos no OB, ou seja, neurônios localizados numa posição mais ventral do MOE projetam seus axônios para uma porção mais ventral do OB e, da mesma forma os OSNs mais dorsais do MOE projetam seus axônios para regiões mais dorsais do OB. Interessantemente, essas características de expressão em zonas e projeção no bulbo olfatório são comuns a indivíduos da mesma linhagem (RESSLER et al., 1994), sugerindo que a organização do sistema olfatório é importante para a interpretação adequada dos estímulos do ambiente. **Tendo em vista o complexo padrão de expressão apresentado pelos genes OR, qual seria o mecanismo molecular que possibilitaria a seleção de apenas um alelo dentre todos os genes OR para serem expressos no neurônio olfatório?**

O padrão de expressão dos genes OR compartilha muitas similaridades com aquele apresentado para os genes de imunoglobulinas, uma vez que cada célula do sistema imune expressa um único tipo de receptor responsável pelo reconhecimento antigênico (JUNG; ALT, 2004), além de ambos atuarem no reconhecimento de uma grande diversidade de compostos originários do ambiente externo. Para expressar um único tipo de receptor, as células do sistema imune editam a sequência do gene de imunoglobulina a nível de DNA (através de recombinação gênica), gerando, dessa forma, um genoma recombinado capaz de expressar apenas uma variante daquele gene. Esse mecanismo de expressão gênica era um potencial candidato para explicar o padrão de expressão dos genes OR. Para testar essa hipótese foram gerados clones de camundongos a partir de material genético extraído de neurônios olfatórios maduros que, portanto, já haviam selecionado um único OR para ser expresso. Assumindo que a recombinação é um evento irreversível, era esperado que os clones gerados a partir do material genético de um OSN, deveriam ser capazes de expressar somente o OR escolhido por esse neurônio. Dois grupos, de forma independente, verificaram que os clones nascidos mantinham a expressão de outros OR e não apenas um, como seria esperado, descartando a hipótese de recombinação (EGGAN et al., 2004; LI, J. et al., 2004).

Outro mecanismo provável para a expressão considera a existência de um elemento em *cis* que seria capaz de ativar a transcrição de apenas um dos membros de uma família gênica, assim como verificado para o *locus* de β -globina (LI, Q.; HARJU; PETERSON, 1999) e para o *locus* dos fotorreceptores vermelho e verde (SMALLWOOD; WANG; NATHANS, 2002). Nesses dois casos, existe uma região regulatória (*locus control region*, LCR) que coordena a expressão dos genes daquele *locus*. O primeiro elemento regulatório em *cis* descoberto, com a função de controlar a expressão dos genes OR, foi a região H (SERIZAWA et al., 2000; SERIZAWA et al., 2003) que foi descrita como uma região de ~2 kb conservada entre camundongos e humanos a ~75kb do gene *MOR28 (Olfr1507)* que, ao ser deletada, suprimiu a expressão de alguns OR localizados próximo a esse elemento. Posteriormente, foi demonstrado que o controle regulatório da região H ocorria, diretamente, apenas em *cis* (FUSS; OMURA; MOMBAERTS, 2007), apesar do posicionamento nuclear desse elemento ser próximo a outros *clusters* de genes OR em núcleos de células do MOE (LOMVARDA et al., 2006). O segundo elemento regulatório descrito foi o elemento P, localizado entre os genes ORs P3 (*Olfr713*) e P4 (*Olfr714*), que compartilha características de um promotor de gene OR (BOZZA et al., 2009). A ausência do elemento P alterou a expressão de 10 dos 24 ORs localizados naquele *cluster*, evidenciando que essa sequência possui uma função regulatória (KHAN; VAES; MOMBAERTS, 2011). Apesar dessas descobertas, apenas uma pequena fração dos genes OR tem sua expressão governada diretamente pelos elementos H e P, o que sugere a existência de outras regiões próximas a *clusters* de OR e que desempenham a mesma função. Por meio de busca de características epigenéticas comumente encontrados em *enhancer* (monometilação da lisina 4 da histona 3, H3K4me1; e acetilação da lisina 27 da histona 3, H3K27ac) (SMITH; SHILATIFARD, 2014), foi possível identificar mais de 30 novos *enhancers* em potencial, com função constatada para 12 deles em um ensaio de expressão em *Danio rerio* (MARKENSCOFF-PAPADIMITRIOU et al., 2014). Nesse mesmo trabalho, demonstrou-se que esses *enhancers* frequentemente se encontravam agrupados no núcleo de OSN, indicando que a expressão de um único gene OR pode ser uma consequência da criação de um *hub* transcricional onde somente a expressão de um gene pode ser mantida.

Mais recentemente, foi descrita a existência de outro *enhancer* (elemento J) que teria a função de regular, exclusivamente, a expressão de ORs classe I, expressos apenas na zona 1 (região mais dorsal do MOE) (IWATA et al., 2017). Todos esses achados não permitem, contudo, uma descrição completa do mecanismo molecular de como um gene OR é selecionado e de como sua expressão é estabilizada pelo OSN.

Paralelamente, outros trabalhos descreveram como algumas modificações epigenéticas, arquitetura nuclear e mecanismos celulares de resposta a estresse estariam envolvidos no padrão de expressão monogênica e monoalélica dos genes OR. Foi verificado, por meio de ensaios de microarranjo de DNA com cromatina imunoprecipitada (ChIP-on-chip) e PCR quantitativo, que a trimetilação na lisina 9 da histona 3 (H3K9me3) e a trimetilação na lisina 20 da histona 4 (H4K20me3), comumente encontradas em regiões heterocromáticas, estavam frequentemente associadas a genes OR em neurônios olfatórios maduros e imaturos, mas não estavam presentes em células horizontais e no fígado (MAGKLARA et al., 2011). O mesmo trabalho verificou a presença de um marcador de eucromatina, trimetilação na lisina 4 da histona 3 (H3K4me3), associado ao gene OR que estava ativo naquele OSN, concomitantemente à redução da marca H3K9me3, sugerindo que essas marcas de histonas são importantes para inibir a expressão dos ORs em neurônios olfatórios, mas não em outras células possivelmente carentes dos fatores necessários para a transcrição dos ORs. Posteriormente, foi descrito que os genes OR encontravam-se agregados ao bloco heterocromático (HB) centralmente localizado em OSN (ARMELIN-CORREA et al., 2014; CLOWNEY et al., 2012) e a ausência de HB levava à perda da singularidade na expressão de um OR, corroborando com a ideia da repressão global dos *loci* de ORs. Outro subcompartimento verificado em núcleos de OSN foi um domínio de heterocromatina facultativa, caracterizado pela trimetilação da lisina 27 da histona 3 (H3K27me3) (ARMELIN-CORREA et al., 2014; TROJER; REINBERG, 2007), justaposto ao domínio HB e que não colocaliza com a eucromatina. Curiosamente, a marca H3K27me3 ocorre na região que flanqueia os *enhancers* de OR (MARKENSCOFF-PAPADIMITRIOU et al., 2014), levantando a hipótese de que essa marca pode auxiliar na formação do *hub* transcrricional formado pelos

enhancers específicos do MOE. Recentemente, foi descrito que genes de uma classe de quimiorreceptores expressos em neurônios olfatórios, os TAARs (*trace amine-associated receptors*), encontram-se associados à periferia nuclear próximos à carioteca (YOON et al., 2015), reforçando a ideia de que a organização nuclear é importante para a expressão singular dos quimiorreceptores no sistema olfatório.

Durante o amadurecimento do OSN, o gene OR ativado não se encontra associado à marca H3K9me3 (MAGKLARA et al., 2011). A demetilase lisina-específica (Lsd1) possui um elevado nível de expressão em células progenitoras e neurônios imaturos, além disso, Lsd1 possui função catalítica sobre H3K9me2, sugerindo que Lsd1 participa dos estágios iniciais da escolha do OR a ser expresso (LYONS et al., 2013). Após a ativação do gene OR, as células impedem que novos OR sejam ativados por meio da inibição de Lsd1. Isso ocorre através da resposta celular a proteínas desenoveladas (*unfolded protein response*, UPR), que leva à tradução da forma nuclear de ATF5 (*activating transcription factor 5*), um fator de transcrição necessário para expressão de adenilil ciclase 3 (Adcy3) (DALTON; LYONS; LOMVARDAS, 2013). Por sua vez, Adcy3 inibe Lsd1 impedindo que um novo gene OR seja ativado (LYONS et al., 2013). A Figura 1.3 apresenta graficamente as etapas de ativação descritas até esse momento.

Figura 1.3. Representação esquemática das principais etapas da escolha do gene de OR.

(A) O neurônio olfatório apresenta em seu núcleo um bloco central de heterocromatina onde estão compartimentalizados os loci de genes OR. (B) A enzima Lsd1 catalisa a remoção da marca repressiva da histona, associada a um gene OR. Isso possibilita a interação do *enhancer* com promotor do gene OR que passa a ser transcrito, enquanto os demais são mantidos. (C) A manutenção de um único OR expresso ocorre por meio da interação entre as diversas regiões regulatórias descritas para os loci de OR. (D) Se um OR não funcional é escolhido, um segundo OR é ativado por Lsd1, e somente após a escolha de um OR funcional é que o mecanismo de UPR no retículo endoplasmático começa a atuar para impedir a ativação de um novo OR por Lsd1. HB: Bloco heterocromático; iOSN: neurônio olfatório imaturo; Lsd1: demetilase lisina-específica. Adaptada de (DEGL'INNOCENTI; D'ERRICO, 2017).

Mais recentemente, por meio da análise de transcriptoma de célula única foi demonstrado que, no estágio inicial de seleção de um gene OR, diversos deles são transcritos numa mesma célula, em níveis centenas de vezes menores daqueles observados em neurônios olfatórios maduros (HANCHATE et al., 2015; TAN; LI; XIE, 2015). Outro trabalho demonstrou que a expressão forçada de um OR, em estágios mais tardios do desenvolvimento do neurônio olfatório, ou seja, após a expressão de uma proteína específica de neurônios olfatórios maduros (OMP) (RODRIGUEZ-GIL et al., 2015), frequentemente, levava à supressão da expressão do OR endógeno, indicando que, apesar da escolha do OR já estar estabilizada, existem mecanismos de controle que continuam atuando no neurônio olfatório a fim de manter a singularidade na expressão de apenas um gene OR (ABDUS-SABOOR et al., 2016). Esse mecanismo foi denominado *post-selection refinement* (PSR).

Apesar dos achados, ainda não é possível descrever em detalhes como somente um OR é expresso por célula, dado o grande número de genes disponíveis para ser expresso além deles estarem presentes por quase todos os cromossomos murinos. Alguns detalhes a nível molecular tais como: quais macromoléculas e como é coordenado o agrupamento dos *enhancers*, quais elementos definem a expressão por zonas no MOE, por que alguns ORs são expressos por um maior número de neurônios que outros, etc., estão entre as questões em aberto que são fundamentais para compreender o ajuste fino que rege a expressão dos genes OR.

1.3 Aspectos gerais sobre metilação de DNA

A metilação de DNA está entre as mais difundidas modificações epigenéticas que alteram quimicamente a estrutura do DNA e está presente em diversas espécies de fungos, plantas e animais (JONES, P. A., 2012; SUZUKI; BIRD, 2008). Essa modificação se caracteriza pela adição de um grupamento metil ($-CH_3$) na posição 5 do anel pirimidínico de citosinas, gerando, dessa forma, a base nitrogenada 5-metilcitosina (5mC), Figura 1.4A. Tanto a citosina como a citosina metilada podem sofrer deaminação espontânea, levando a duas bases nitrogenadas abundantes nas células, uracila e timina, respectivamente. A adição do grupo metil às citosinas é catalisada pelas enzimas DNA metiltransferases (Dnmts) que transferem o grupo metil da molécula S-adenosilmetionina (SAM) para o seu alvo no DNA. Dentre as diversas enzimas com função de metiltransferase, destaca-se Dnmt1, que atua na manutenção do padrão de metilação durante a replicação do DNA, copiando na fita recém-sintetizada o padrão da fita parental, Figura 1.4B. As enzimas Dnmt3a e Dnmt3b, por sua vez, possuem atividade metiltransferase *de novo*, que permite a essas enzimas adicionar grupamentos metil a regiões do DNA previamente não metiladas em nenhuma das fitas, Figura 1.4C (MOORE; LE; FAN, 2013). Camundongos *knockout* (KO) para Dnmt1 e Dnmt3b morrem durante a embriogênese (LI, E.; BESTOR; JAENISCH, 1992; OKANO et al., 1999), enquanto o KO para Dnmt3a não impede o nascimento do camundongo, mas leva à sua morte prematura, decorrente de problemas relacionados à neurogênese pós-natal (WU et al., 2010).

O principal alvo das Dnmts são citosinas que antecedem nucleotídeos guanina, denotado CpG (a letra *p* representa a ligação fosfodiéster entre citosina e guanina). Esses dinucleotídeos são sub-representados no genoma de diversos animais, e de forma mais acentuada em mamíferos. Em camundongos e humanos, por exemplo, cerca de 1% dos dinucleotídeos no genoma são CpG (EHRLICH et al., 1982; JABBARI; BERNARDI, 2004), discrepante dos 5,25% se a ocorrência dos nucleotídeos no genoma fosse ao acaso. Evolutivamente, essa depleção poderia ser explicada à luz das transformações químicas sofridas pela citosina (Figura1.4A). A deaminação de citosina produz uracila, prontamente reparada, pois esta base nitrogenada não é um constituinte natural do DNA, ao passo que deaminação de 5-metilcitosina produz timina, uma das quatro bases presentes no DNA, o que dificulta o reparo correto da base alterada (DUNCAN; MILLER, 1980). Outros dinucleotídeos (CpA, CpT e CpC) ou trincas (CHG, CHH, onde H = A, T ou C) são alvos pouco frequentes de Dnmt3a/3b em células de mamíferos, somando cerca de 25% em células tronco embrionárias (ESC) e em células do giro denteado do hipocampo, mas não alcançam 5% em outros tipos celulares (GUO et al., 2014; LISTER et al., 2009; RAMSAHOYE et al., 2000; ZILLER et al., 2011).

Figura 1.4. A metilação de DNA ocorre frequentemente em citosina.

(A) Representação estrutural das bases nitrogenadas uracila, citosina, 5-metilcitosina (5mC) e timina, indicando algumas transformações responsáveis pela interconversão dessas moléculas. ① Deaminação. (B) A replicação geral produz fitas com citosinas desmetiladas. Por meio da ação da enzima DNA metiltransferase 1 (Dnmt1) o padrão de metilação pré-existente à replicação é repassado para a fita recém-sintetizada. (C) As enzimas Dnmt3a e Dnmt3b adicionam grupo metil preferencialmente a citosinas que precedem guaninas (CpG) em fitas de DNA que não possuem um padrão prévio de metilação, processo denominado metilação *de novo*. Os itens B e C foram adaptados de (MOORE et al., 2013).

Apesar de alguns genomas apresentarem uma frequência de dinucleotídeos CpG menores do que a esperada, existem regiões altamente adensadas em nucleotídeos G e C e, conseqüentemente, em CpG. Inicialmente, regiões ricas em CpG foram descritas como trechos do DNA vulneráveis ao tratamento com endonucleases de restrição sensíveis a DNA metilado (BIRD et al., 1985). Tais regiões (aproximadamente 1 kb) foram denominadas ilhas CpG (CGI) e, frequentemente, estão associadas a regiões a 5' de unidade transcricionais de genes tecido específico e, principalmente, a genes *housekeeping* (GARDINER-GARDEN;

FROMMER, 1987). Posteriormente, foi verificado que cerca de 70% dos genes anotados no genoma humano possuíam CGI na região promotora (SAXONOV; BERG; BRUTLAG, 2006) e, na maior parte desses promotores, os CpGs apresentavam-se desmetilados (abundância de 5mC < 20%) (ECKHARDT et al., 2006). Em contrapartida, CpG fora de ilhas CpG encontram-se majoritariamente metiladas (SUZUKI; BIRD, 2008). Além disso, foi demonstrado que o padrão de metilação para algumas células pode ocorrer de forma tecido específica e dependente do gênero do animal (ECKHARDT et al., 2006; EHRLICH et al., 1982).

Uma das primeiras evidências da função biológica da metilação de DNA foi obtida a partir de trabalhos com o gene de β -globina, no qual foi verificado uma menor ocorrência de DNA metilado em eritrócito quando comparado a um tecido que não expressa o gene da β -globina (RAZIN; RIGGS, 1980). Outro exemplo envolvendo metilação de DNA mostrou na inativação de um dos cromossomos X em fêmeas que as regiões próximas ao sítio de início da transcrição (TSS) apresentavam-se altamente metiladas (MOHANDAS; SPARKES; SHAPIRO, 1981), enquanto que os CpGs localizados no corpo do gene encontravam-se majoritariamente desmetilados (HELLMAN; CHESS, 2007). A metilação de DNA também atua na regulação de genes por meio do *Imprinting* genômico, em que apenas um dos alelos (paterno ou materno) de um gene deve ser expresso. Um caso de *imprinting* pode ser visto nos *loci* dos genes *H19* e *Igf2*, onde existe uma região regulatória altamente metilada apenas no alelo paterno. Como consequência dessa metilação, *Igf2* passa a ser expresso apenas nesse alelo, ao passo que o alelo materno expressa apenas o gene *H19* (BELL; FELSENFELD, 2000; SINGH; LEE; SZABO, 2012).

A metilação de DNA foi associada a uma alteração na afinidade de algumas proteínas ao DNA, por conta do impedimento estérico proporcionado pelo grupo metil. No fim da década de 80 foi descoberta a proteína MeCP2, que possui domínio de ligação a metil CpG (MBD) (MEEHAN et al., 1989). Posteriormente, outras 4 proteínas (MBD1, MBD2, MBD3 e MBD4) foram descritas por possuírem o domínio MBD (HENDRICH; BIRD, 1998; LEWIS et al., 1992). Tais proteínas possuem funções no remodelamento da cromatina, além de mutações em

MeCP2 terem sido associadas a algumas doenças neurocognitivas, como a síndrome de Rett (AMIR et al., 1999; MOORE et al., 2013).

Uma característica particularmente interessante sobre a metilação de DNA é a sua reversibilidade. Dessa forma esse poderia ser um potencial alvo de como ocorre a reprogramação e diferenciação celular ao longo do desenvolvimento. A remoção do grupo metil ocorre em múltiplas etapas catalisadas pelas enzimas da família TET (*Ten Eleven Translocation*), capazes de oxidar gradativamente o grupamento metil incorporado à citosina, até seu intermediário mais oxidado, 5-carboxicitosina (5-caC). A etapa final de demetilação envolve o reparo por excisão de bases (BER) mediada pela *thymine DNA glycosylase* (TDG), regenerando a citosina (KOHLI; ZHANG, 2013).

1.4 Metilação de DNA no sistema olfatório

Uma maior atenção foi lançada sobre metilação de DNA no sistema olfatório após a publicação de um trabalho que caracterizou a expressão das três principais Dnmts nos diferentes tipos celulares do MOE (MACDONALD; GIN; ROSKAMS, 2005). Foi demonstrado que Dnmt1 é expressa por células de todos os estágios de desenvolvimento do neurônio olfatório, ao contrário de Dnmt3b, que tem um padrão de expressão restrito às células globosas (GBC), e de Dnmt3a, cuja expressão coincide com marcadores de neurônio olfatório imaturo (Gap43 e NST). Posteriormente, outro trabalho descreveu que a expressão das proteínas MeCP2 e uma isoforma da proteína MBD2 (MBD2a) são expressas majoritariamente em OSNs e, na ausência dessas proteínas, foi verificado o acúmulo de iOSNs, além da elevação na taxa de regeneração do MOE (MACDONALD et al., 2010), Figura 1.5. Além disso, MeCP2 altera a transcrição de uma proteína de adesão celular no MOE (*Pcdh20*) como resposta à estimulação por odorante (LEE et al., 2014). Estes resultados sugerem que a metilação de DNA atua em estágios específicos ao longo do desenvolvimento dos OSNs que são fundamentais para a formação do MOE.

Figura 1.5. Expressão de proteínas relacionadas à metilação de DNA ao longo do desenvolvimento do Epitélio Olfatório.

À esquerda estão as proteínas que ligam DNA metilado (MeCP2 e MBD2) e as enzimas que adicionam o grupo metil à citosina do DNA. O tamanho do retângulo corresponde à região no MOE onde essas proteínas foram localizadas. Ao centro temos a representação do MOE, com os tipos celulares que o compõem e à direita o nome das células e seus marcadores entre parênteses, colocados na altura relativa em que são encontradas em cada tipo celular nesse neuroepitélio. Adaptado de (MACDONALD et al., 2005; MACDONALD et al., 2010).

Tendo em vista que a enzima Dnmt3a é a principal responsável pela metilação *de novo* em OSNs maduros, seria preciso encontrar quais sequências do DNA são os alvos principais dessa enzima. O segundo alvo mais frequente das Dnmts *de novo* são os dinucleotídeos CpA (ZILLER et al., 2011). Colquitt e colaboradores verificaram que a metilação em CGI e *enhancer* de MOE são dependentes de Dnmt3a, que pode estar relacionada à metilação na região H (LOMVARDAS et al., 2006), em dinucleotídeos CpA. Além disso, o nível de metilação foi menor em animais heterozigotos para Dnmt3a quando comparados a animais com os dois alelos funcionais, indicando que a inserção dessa modificação no DNA depende da abundância de Dnmt3a (COLQUITT et al., 2014). Curiosamente, outro trabalho mostrou que animais KO para Dnmt3a apresentavam acentuada redução no tamanho do bulbo olfatório, indicando uma grave perturbação no desenvolvimento do sistema olfatório (WU et al., 2010).

O aduto de DNA 5-hidroximetilcitosina (5hmC) é formado no primeiro estágio da demetilação de DNA por proteínas do complexo TET (KOHLEI; ZHANG, 2013). No entanto, mais recentemente foi verificado um acúmulo dessas modificações em células tronco

embrionárias, células da medula óssea e do cérebro (RUZOV et al., 2011) indicando que esse intermediário da via de demetilação pode ter alguma função nessas células. Em MOE, a presença de 5hmC no corpo do gene parece facilitar a transcrição do gene associado a essa marca, especialmente na transição de células progenitoras para neurônio olfatório maduro (COLQUITT et al., 2013).

Mais recentemente, foi descrito que a prole (F1 e F2) de camundongos condicionados aversivamente ao ligante do receptor M71 (*Olf151*) apresentou um aumento no número de OSNs que expressavam M71 (DIAS; RESSLER, 2014), mostrando que uma experiência vivida pelo progenitor pode levar a uma mudança anatômica em sua prole. Esse mesmo trabalho sugere que uma redução no nível de metilação em CpG da região codificadora de M71 em DNA de espermatozoide é a forma pela qual essa informação é transmitida através das gerações. Mais trabalhos são necessários para elucidar o papel da metilação de DNA na regulação da expressão gênica e, mais especificamente, de que forma genes OR são afetados por essa marca epigenética.

2 OBJETIVOS

Objetivo geral:

- Compreender o papel da metilação de DNA na regulação da expressão dos genes de receptores olfatórios.

Objetivos específicos:

- Caracterizar o conteúdo CpG na região promotora e codificadora dos genes OR;
- Determinar o padrão de metilação do gene *Olf17* em diferentes tecidos.
- Determinar o padrão de metilação de DNA em outros genes OR e regiões regulatórias.
- Caracterizar as alterações genéticas e epigenéticas no promotor do gene *Olf17*, em diferentes contextos genéticos.
- Verificar a expressão do gene *Olf17* em diferentes contextos genéticos.

3 MATERIAL E MÉTODOS

3.1 Análise do conteúdo de CpG em regiões promotoras e codificadoras de genes OR

Foram analisadas 198 regiões promotoras e 864 regiões codificadoras de genes de receptores olfatórios quanto ao conteúdo e a posição dos dinucleotídeos CpG. O primeiro passo dessa análise consistiu em gerar um comando no *software* Microsoft Excel® que permitisse a separação dos nucleotídeos de uma determinada sequência, individualmente em cada célula. O comando utilizado foi *EXT.TEXTO(\$A1;COL()-COL(\$C1)+1;1)*, onde A1 refere-se a célula na qual está a sequência de nucleotídeos em uma linha; C1 trata-se da célula na qual a fórmula foi inserida e que contém o primeiro nucleotídeo da sequência. A propagação dessa fórmula para as células seguintes, na mesma linha, levou a separação dos nucleotídeos, um a cada célula. A fim de contabilizar os nucleotídeos CpGs de uma sequência, outra aba foi criada e a cada célula de uma mesma linha foi aplicado o comando *SE(E(Seq!C1="c";Seq!D1="g");1;0)*, onde Seq! refere-se a aba em que estão os nucleotídeos separados em células únicas, C1 e D1 são as células da aba Seq! que correspondem a nucleotídeos subsequentes numa mesma sequência. Assim, quando uma citosina estiver antecedendo uma guanina, a célula apresentará o número 1, se não, o número 0 estará na célula. Os promotores de genes OR analisados estão na Tabela 3.1 sob a nomenclatura OLFR e as regiões codificadoras analisadas estão na Tabela 3.2, sob a nomenclatura MOR.

Tabela 3.1: Códigos OLFR dos 198 genes de receptores olfatórios correspondentes aos promotores analisados. Sequências disponibilizadas em (MICHALOSKI; GALANTE; MALNIC, 2006).

OLFR1180	OLFR1101	OLFR1308	OLFR140	OLFR18	OLFR279	OLFR389	OLFR497	OLFR741	OLFR1080	OLFR1077P
OLFR566	OLFR1105	OLFR131	OLFR1414	OLFR180	OLFR286	OLFR39	OLFR506	OLFR75P	OLFR1090	OLFR490
OLFR947P	OLFR1123	OLFR1324	OLFR1415	OLFR181	OLFR287P	OLFR390	OLFR510	OLFR77	OLFR1096P	OLFR1294
OLFR1	OLFR1157	OLFR1330	OLFR1416	OLFR183	OLFR288	OLFR392	OLFR53	OLFR796	OLFR1295	OLFR491
OLFR1000	OLFR1162	OLFR1339	OLFR1434	OLFR186	OLFR291	OLFR396P	OLFR541	OLFR801	OLFR1297	OLFR1366
OLFR1006	OLFR1182	OLFR1340	OLFR1437	OLFR187	OLFR309	OLFR398	OLFR544	OLFR806	OLFR1301	OLFR494
OLFR1012	OLFR1198	OLFR1342	OLFR1440	OLFR19	OLFR314	OLFR403	OLFR545	OLFR845	OLFR1368	OLFR171
OLFR1013	OLFR1209	OLFR1346	OLFR1461	OLFR190	OLFR315	OLFR413	OLFR56	OLFR857	OLFR1377	OLFR720
OLFR1014	OLFR121	OLFR1347	OLFR1463	OLFR191	OLFR32	OLFR414	OLFR59	OLFR859	OLFR139	OLFR232
OLFR1016	OLFR123	OLFR1349	OLFR1484	OLFR192	OLFR325	OLFR429	OLFR60	OLFR860	OLFR173	OLFR725
OLFR1020	OLFR1231	OLFR1351	OLFR1487	OLFR195	OLFR328	OLFR43	OLFR601	OLFR867	OLFR174	OLFR370
OLFR1021P	OLFR124	OLFR1352	OLFR1496	OLFR196	OLFR330	OLFR449	OLFR61	OLFR868	OLFR177	OLFR726
OLFR1025P	OLFR1259	OLFR1353	OLFR1501	OLFR2	OLFR332	OLFR457	OLFR648	OLFR869	OLFR24	OLFR488
OLFR1026	OLFR1260	OLFR1354	OLFR1506	OLFR20	OLFR338	OLFR459	OLFR656	OLFR872	OLFR259	OLFR984
OLFR1034	OLFR1261	OLFR1356	OLFR1509	OLFR204	OLFR339	OLFR473	OLFR672	OLFR873	OLFR27	OLFR711
OLFR1037	OLFR1270	OLFR1357	OLFR1510	OLFR209	OLFR348	OLFR478	OLFR690	OLFR90	OLFR374	OLFR985
OLFR1045	OLFR1273	OLFR1359	OLFR160	OLFR211	OLFR362	OLFR480	OLFR702	OLFR951	OLFR376	OLFR972
OLFR1065	OLFR1286	OLFR1364	OLFR165	OLFR214	OLFR367P	OLFR483	OLFR708	OLFR969	OLFR378	OLFR994

Tabela 3.2. Códigos MOR dos 864 genes de receptores olfatórios utilizados para a análise de conteúdo CpG.

MOR10-1	MOR118-1	MOR138-2	MOR162-6	MOR174-11	MOR186-2	MOR204-8	MOR223-9	MOR233-5	MOR251-2	MOR261-1	MOR274-1
MOR101-1	MOR119-1	MOR138-3	MOR162-7	MOR174-12	MOR187-1	MOR204-9	MOR224-1	MOR233-6	MOR251-3	MOR261-10	MOR274-2
MOR101-2	MOR119-2	MOR139-1	MOR163-1	MOR174-13	MOR187-2	MOR205-1	MOR224-10	MOR233-7	MOR252-1	MOR261-11	MOR275-1
MOR10-2	MOR1-2	MOR139-2	MOR164-1	MOR174-16	MOR187-3	MOR202-8	MOR224-12	MOR233-8	MOR252-2	MOR261-12	MOR275-2
MOR102-1	MOR120-1	MOR139-3	MOR164-2	MOR174-2	MOR188-1	MOR202-9	MOR224-2	MOR233-9	MOR253-1	MOR261-13	MOR276-1
MOR103-1	MOR120-2	MOR139-4	MOR164-3	MOR174-3	MOR188-2	MOR203-1	MOR224-3	MOR234-1	MOR253-2	MOR261-2	MOR276-2
MOR103-10	MOR120-3	MOR140-1	MOR165-1	MOR174-4	MOR188-3	MOR203-2	MOR224-4	MOR238-1	MOR253-3	MOR261-3	MOR277-1
MOR103-11	MOR12-1	MOR14-1	MOR165-4	MOR174-5	MOR188-4	MOR203-3	MOR224-5	MOR238-2	MOR254-2	MOR261-4	MOR278-1
MOR103-15	MOR121-1	MOR14-10	MOR165-5	MOR174-9	MOR188-5	MOR204-4	MOR224-6	MOR239-1	MOR255-1	MOR262-4	MOR279-1
MOR103-16	MOR122-1	MOR14-3	MOR165-6	MOR175-1	MOR190-2	MOR204-5	MOR224-9	MOR239-2	MOR255-2	MOR262-5	MOR283-4
MOR103-2	MOR122-2	MOR143-1	MOR165-7	MOR175-2	MOR19-1	MOR206-1	MOR225-1	MOR239-3	MOR255-3	MOR262-6	MOR283-5
MOR103-3	MOR123-1	MOR14-4	MOR165-8	MOR175-3	MOR191-1	MOR206-2	MOR225-2	MOR239-4	MOR255-4	MOR262-7	MOR283-6
MOR103-4	MOR123-2	MOR144-1	MOR167-1	MOR175-4	MOR19-2	MOR206-3	MOR234-2	MOR239-5	MOR255-5	MOR262-8	MOR283-7
MOR103-5	MOR12-5	MOR14-5	MOR167-2	MOR175-5	MOR192-1	MOR206-4	MOR234-3	MOR239-6	MOR255-6	MOR262-9	MOR283-8
MOR103-6	MOR1-4	MOR145-1	MOR167-3	MOR176-1	MOR194-1	MOR207-1	MOR235-1	MOR240-1	MOR256-1	MOR263-3	MOR283-9
MOR103-7	MOR111-4	MOR145-2	MOR167-4	MOR176-2	MOR195-1	MOR208-1	MOR235-2	MOR240-2	MOR256-10	MOR263-1	MOR285-1
MOR103-8	MOR111-5	MOR145-3	MOR168-1	MOR176-3	MOR196-1	MOR208-2	MOR237-1	MOR240-3	MOR256-11	MOR263-10	MOR285-2
MOR103-9	MOR111-6	MOR145-4	MOR169-1	MOR177-1	MOR196-2	MOR208-3	MOR237-2	MOR241-1	MOR256-12	MOR263-2	MOR29-1
MOR104-1	MOR111-7	MOR145-5	MOR170-1	MOR177-5	MOR196-3	MOR208-4	MOR225-3	MOR241-2	MOR256-13	MOR263-3	MOR30-1
MOR104-2	MOR11-2	MOR145-6	MOR170-11	MOR177-6	MOR196-4	MOR209-1	MOR225-4	MOR241-3	MOR256-14	MOR263-4	MOR30-2
MOR104-3	MOR125-1	MOR14-6	MOR170-2	MOR177-7	MOR199-1	MOR2-1	MOR225-5	MOR24-2	MOR256-15	MOR263-5	MOR30-3
MOR104-4	MOR126-1	MOR146-1	MOR170-3	MOR177-8	MOR199-2	MOR210-1	MOR226-1	MOR242-1	MOR256-16	MOR263-6	MOR3-1
MOR105-1	MOR126-2	MOR146-2	MOR170-4	MOR177-9	MOR20-1	MOR210-2	MOR227-1	MOR24-3	MOR256-17	MOR263-9	MOR31-1
MOR105-10	MOR127-1	MOR146-3	MOR170-5	MOR178-1	MOR201-1	MOR210-3	MOR227-2	MOR244-1	MOR256-18	MOR264-1	MOR31-10
MOR105-2	MOR127-2	MOR147-1	MOR170-6	MOR179-1	MOR201-2	MOR210-4	MOR227-3	MOR244-2	MOR256-19	MOR264-17	MOR31-11
MOR105-3	MOR127-3	MOR147-2	MOR170-7	MOR179-2	MOR202-1	MOR210-5	MOR227-4	MOR244-3	MOR256-2	MOR264-18	MOR31-12
MOR105-4	MOR127-4	MOR147-3	MOR170-8	MOR179-3	MOR202-10	MOR21-1	MOR227-5	MOR253-4	MOR256-20	MOR264-19	MOR31-2
MOR105-5P	MOR128-2	MOR148-1	MOR170-9	MOR179-6	MOR202-11	MOR211-1	MOR228-2	MOR253-5	MOR256-21	MOR264-2	MOR31-3
MOR106-1	MOR129-1	MOR165-2	MOR17-1	MOR179-7	MOR202-12	MOR211-2	MOR228-3	MOR253-6	MOR256-22	MOR264-20	MOR31-4
MOR106-11	MOR129-2	MOR165-3	MOR171-1	MOR180-1	MOR202-13	MOR212-1	MOR228-4	MOR253-7	MOR256-23	MOR264-3	MOR31-5
MOR106-12	MOR129-3	MOR141-1	MOR171-10	MOR18-1	MOR202-14	MOR212-2	MOR229-1	MOR253-8	MOR256-24	MOR264-4	MOR31-6
MOR106-2	MOR1-3	MOR141-2	MOR171-11	MOR181-1	MOR202-15	MOR212-3	MOR230-1	MOR253-9	MOR256-25	MOR264-5	MOR31-7
MOR106-3	MOR130-1	MOR141-3	MOR171-12	MOR181-2	MOR202-16	MOR213-2	MOR230-2	MOR254-1	MOR256-26	MOR264-6	MOR31-8
MOR106-4	MOR130-2	MOR14-2	MOR171-13	MOR18-2	MOR202-17	MOR213-3	MOR230-3	MOR245-1	MOR256-27	MOR265-1	MOR31-9
MOR106-5	MOR13-1	MOR142-1	MOR171-14	MOR182-1	MOR202-18	MOR213-4	MOR230-4	MOR245-10	MOR256-28	MOR266-1	MOR32-1
MOR106-6	MOR131-1	MOR14-9	MOR171-15	MOR182-2	MOR202-19	MOR213-5	MOR230-5	MOR245-11	MOR256-29	MOR266-2	MOR32-11
MOR107-1	MOR13-2	MOR149-1	MOR171-16	MOR182-3	MOR202-2	MOR213-6	MOR230-6	MOR245-12	MOR256-3	MOR266-3	MOR32-2
MOR108-1	MOR132-1	MOR149-2	MOR171-17	MOR182-4	MOR202-20	MOR214-1	MOR230-7	MOR245-13	MOR256-30	MOR266-4	MOR32-3
MOR108-4	MOR13-3	MOR149-3	MOR171-18	MOR182-5	MOR202-21	MOR214-2	MOR230-8	MOR245-2	MOR256-4	MOR266-8	MOR32-4
MOR109-1	MOR13-4	MOR15-1	MOR171-23	MOR182-6	MOR202-33	MOR214-3	MOR23-1	MOR245-20	MOR256-48	MOR266-9	MOR32-5
MOR1-1	MOR134-1	MOR151-1	MOR171-24	MOR18-3	MOR202-34	MOR214-4	MOR231-1	MOR245-21	MOR261-5	MOR267-1	MOR33-1
MOR110-1	MOR135-1	MOR152-1	MOR171-25	MOR183-1	MOR202-35	MOR214-5	MOR231-10	MOR245-3	MOR261-6	MOR267-14	MOR33-2
MOR110-10	MOR135-10	MOR152-2	MOR174-6	MOR183-2	MOR202-36	MOR215-1	MOR231-11	MOR245-4	MOR261-9	MOR267-16	MOR34-1
MOR110-2	MOR135-11	MOR152-3	MOR174-7	MOR183-3	MOR202-37	MOR215-2	MOR231-12	MOR245-5	MOR26-2	MOR267-2	MOR34-2
MOR110-3	MOR135-12	MOR154-1	MOR174-8	MOR183-4	MOR202-4	MOR215-3	MOR231-13	MOR245-6	MOR262-1	MOR267-3	MOR34-3
MOR110-4	MOR135-13	MOR155-1	MOR171-19	MOR183-8	MOR202-5	MOR216-1	MOR231-14	MOR245-7	MOR262-10	MOR267-4	MOR34-4
MOR110-5	MOR135-14	MOR155-2	MOR171-2	MOR183-9	MOR202-6	MOR217-1	MOR231-18	MOR245-8	MOR262-2	MOR267-5	MOR34-5
MOR110-6	MOR135-2	MOR156-1	MOR171-20	MOR184-1	MOR202-7	MOR218-1	MOR231-2	MOR245-9	MOR262-3	MOR267-6	MOR34-6
MOR11-1	MOR135-26	MOR156-2	MOR171-21	MOR184-2	MOR203-4	MOR218-10	MOR231-3	MOR246-2	MOR256-5	MOR267-7	MOR34-7
MOR111-1	MOR135-28	MOR156-3	MOR171-22	MOR184-3	MOR204-1	MOR218-2	MOR231-4	MOR246-3	MOR256-6	MOR267-8	MOR35-1
MOR111-3	MOR135-3	MOR156-4	MOR171-3	MOR184-4	MOR204-10	MOR218-3	MOR231-5	MOR246-4	MOR256-7	MOR268-1	MOR36-1
MOR112-1	MOR135-4	MOR156-5	MOR171-4	MOR184-5	MOR204-11	MOR218-7	MOR231-6	MOR246-5	MOR256-8	MOR268-2	MOR38-1
MOR112-2	MOR135-6	MOR157-1	MOR171-42	MOR184-6	MOR204-12	MOR218-8	MOR231-7	MOR246-6	MOR256-9	MOR268-3	MOR38-2
MOR113-1	MOR135-7	MOR158-1	MOR171-44	MOR184-7	MOR204-13	MOR218-9	MOR231-8	MOR247-1	MOR257-1	MOR268-5	MOR39-1
MOR113-2	MOR135-8	MOR159-1	MOR171-46	MOR184-8	MOR204-14	MOR219-1	MOR231-9	MOR247-2	MOR257-2	MOR279-2	MOR40-1
MOR113-3	MOR135-9	MOR159-3	MOR171-5	MOR185-1	MOR204-15	MOR219-2	MOR232-1	MOR248-1	MOR257-3	MOR280-1	MOR40-13
MOR113-4	MOR13-6	MOR160-1	MOR171-6	MOR189-1	MOR204-16	MOR220-1	MOR232-2	MOR248-10	MOR257-4	MOR28-1	MOR40-2
MOR114-1	MOR136-1	MOR160-2	MOR171-7	MOR189-2	MOR204-17	MOR220-2	MOR232-3	MOR248-11	MOR258-1	MOR281-1	MOR40-3
MOR114-11	MOR136-10	MOR160-4	MOR171-8	MOR189-3	MOR204-18	MOR220-3	MOR232-4	MOR248-2	MOR258-2	MOR282-1	MOR40-4
MOR114-12	MOR136-11	MOR160-5	MOR171-9	MOR190-1	MOR204-19	MOR22-1	MOR232-5	MOR248-3	MOR258-3	MOR283-1	MOR4-1
MOR114-2	MOR136-12	MOR16-1	MOR17-2	MOR177-10	MOR204-2	MOR22-2	MOR232-6	MOR248-4	MOR258-4P	MOR283-11	MOR41-1
MOR114-3	MOR136-13	MOR161-1	MOR172-1	MOR177-14	MOR204-20	MOR222-1	MOR232-7	MOR248-5	MOR258-5	MOR283-2	MOR42-1
MOR114-4	MOR136-14	MOR161-2	MOR172-2	MOR177-2	MOR204-21	MOR222-2	MOR232-9	MOR248-6	MOR258-6	MOR283-3	MOR5-1
MOR114-5	MOR136-2	MOR161-3	MOR172-3	MOR177-3	MOR204-22	MOR222-3	MOR233-1	MOR248-7	MOR259-1	MOR269-1	MOR5-2
MOR114-6	MOR136-3	MOR161-4	MOR172-4	MOR177-4	MOR204-23	MOR22-3	MOR233-10	MOR248-8	MOR259-10	MOR269-2	MOR6-1
MOR114-7	MOR136-4	MOR161-5	MOR172-5	MOR185-2	MOR204-32	MOR223-1	MOR233-11	MOR248-9	MOR259-8	MOR269-3	MOR7-1
MOR114-8	MOR136-5	MOR161-6	MOR172-6	MOR185-3	MOR204-34	MOR223-2	MOR233-12	MOR249-2	MOR260-1	MOR270-1	MOR7-2
MOR114-9	MOR136-6	MOR162-1	MOR173-1	MOR185-4	MOR204-35	MOR223-3	MOR233-13	MOR250-1	MOR260-2	MOR27-1	MOR8-1
MOR115-1	MOR136-7	MOR162-2	MOR173-2	MOR185-5	MOR204-36	MOR223-4	MOR233-18	MOR250-2	MOR260-3	MOR271-1	MOR8-2
MOR115-4	MOR136-8	MOR162-3	MOR173-3	MOR185-6	MOR204-37	MOR223-5	MOR233-2	MOR250-3	MOR260-4	MOR272-1	MOR8-3
MOR116-1	MOR136-9	MOR162-4	MOR174-1	MOR185-7	MOR204-6	MOR223-6	MOR233-3	MOR25-1	MOR260-5	MOR273-1	MOR9-1
MOR117-1	MOR138-1	MOR162-5	MOR174-10	MOR186-1	MOR204-7	MOR223-8	MOR233-4	MOR251-1	MOR26-1	MOR273-2	MOR9-2

Para a determinação da fração CpG observada e esperada (Obs/Exp CpG) foi utilizada a Equação 3.1 (GARDINER-GARDEN; FROMMER, 1987).

$$\text{Obs/Exp CpG} = \frac{\text{número de CpG}}{(\text{número de C}) \times (\text{número de G})} \times N \quad (\text{Eq. 3.1})$$

Onde N é o número total de nucleotídeos na sequência analisada.

3.2 Determinação do padrão de metilação pelo método de bissulfito (DNAbiss-seq).

O método escolhido para o estudo do padrão de metilação de DNA foi sequenciamento pelo método de Sanger de DNA tratado com bissulfito de sódio (DNAbiss-seq). Para isso, adotou-se o procedimento que está representado resumidamente na Figura 3.1.

Figura 3.1. Fluxograma de trabalho para a análise de metilação do DNA.

3.2.1 Animais de experimentação

Os animais de experimentação utilizados foram mantidos no Biotério de Produção e Experimentação da Faculdade de Ciências Farmacêuticas e Instituto de Química USP, com permissão de uso concedida pela Comissão de Ética em Uso de Animais (CEUA), sob protocolos no.19/2013 e no.50/2017. Os procedimentos utilizados no manuseio dos animais estão disponíveis no site do biotério (<http://www.usp.br/bioterio/>) e foram avaliados pela CEUA.

3.2.2 Extração de DNA genômico (gDNA)

O DNA genômico foi extraído de epitélio olfatório (OE) dissecado de camundongos da linhagem C57BL/6 selvagem com 6 a 8 semanas de idade (adulto). O OE foi dissecado em solução de 1xPBS (tampão de fosfato 0,01M, NaCl 0,138 M e KCl 2,7 mM em pH 7,4) e mantido a 4 °C até a extração do DNA. Essa foi realizada por meio do sistema DNeasy Blood & Tissue (Qiagen, no. cat. 69504) no item relativo à extração de DNA de tecido. Resumidamente, a extração consistiu na exposição de até 25 mg de tecido em tampão de lise com proteinase K (600 mAU/mL) até a lise total do tecido (≈ 10 min), 56 °C sob agitação periódica. Adicionou-se 200 μ L de etanol absoluto e transferiu-se a mistura para uma mini coluna de separação fornecida com o kit. Centrifuga-se a mini coluna de separação (7000 x g) e descarta-se a solução. Duas lavagens com solução contendo etanol são realizadas seguidas de centrifugação. O DNA foi eluído da coluna com 200 μ L de água livre de DNase e armazenado a -20 °C.

3.2.3 Tratamento de DNA com bissulfito de sódio

O tratamento de DNA com bissulfito de sódio (NaHSO_3) promove a deaminação de citosinas (C) em simples-fita de DNA (ssDNA) convertendo-as em uracilas (U), Figura 3.2. No entanto, citosinas que apresentam um grupamento metil ligado ao carbono na posição 5 do anel pirimidínico (5-mC) não são convertidas.

Figura 3.2. Deaminação seletiva de citosina com bissulfito de sódio.

(A) A deaminação de citosina utilizando bissulfito de sódio é facilitada pela adição de grupo sulfito na posição 6 do anel pirimidínico da citosina. (B) O aduto formado a partir da adição do grupamento sulfito à 5-metilcitosina (representado em cinza claro) ocorre de forma mais lenta quando comparada ao tempo observado para a adição em citosina. Dessa forma, a deaminação do aduto de 5-mC não é formada de forma significativa. Imagem adaptada de (HUANG et al., 2010).

A diferença de reatividade entre citosina e 5-mC, associada à técnica de sequenciamento de DNA pelo método de Sanger permitem a identificação, no genoma, de citosinas que se apresentavam metiladas no tecido de origem (FROMMER et al., 1992), Figura 3.3.

Figura 3.3. Representação esquemática do método de conversão com bissulfito de sódio.

O tratamento de DNA genômico com NaHSO₃ promove a conversão de citosinas não metiladas em uracilas, enquanto que os resíduos de citosina metilados mantêm a citosina inicialmente existente no DNA coletado. Durante a amplificação do fragmento pela técnica de PCR os amplicons conterão timina ao invés de uracila.

Para o tratamento de DNA com bissulfito foi utilizado o sistema EpiTect Bisulfite Kits (Qiagen, cat. no. 59104), de acordo com as instruções do fabricante. Resumidamente, a 1 ng – 2 µg de DNA genômico foram adicionados 85 µL de solução de bissulfito de sódio e 35 µL de tampão de proteção do DNA em um volume final de 140 µL. A mistura foi submetida aos seguintes ciclos em um termociclador: denaturação (5 minutos a 95 °C); incubação (25 minutos a 60 °C); denaturação (5 minutos a 95 °C); incubação (85 minutos a 60 °C); denaturação (5 minutos a 95 °C); incubação (175 minutos a 60 °C); manter a 20 °C até a utilização. O DNA foi purificado para remoção do bissulfito e demais subprodutos.

3.2.4 Precipitação do DNA tratado com bissulfito de sódio (DNAbiss)

Após a conversão de DNA com bissulfito de sódio realizou-se um passo de precipitação de DNA tratado. Adicionou-se à suspensão de DNAbiss: 1 mL de etanol absoluto; 10 µL de glicogênio (1 g/L); 10% (volume da suspensão inicial de DNA) de acetato de sódio 3 M, pH 5,5. Essa solução foi mantida por 10 minutos a -80 °C ou -20 °C por 16 horas, seguido de centrifugação por 40 minutos, 20800 x g a 4 °C. Descartou-se o sobrenadante com pipeta e adicionou-se 100 µL de etanol 70%, seguido de centrifugação por 15 min, 20800 x g a 4 °C. Descartou-se o sobrenadante e secou-se o precipitado à temperatura ambiente. O DNA é então ressuspendido em 20 µL de tampão de eluição do kit EpiTect Bisulfite e armazenado a -20 °C.

3.2.5 PCR com DNAbiss

Para a amplificação das sequências de interesse para o estudo foram utilizados os oligonucleotídeos (*primers*) apresentados na Tabela 3.3, assim como a temperatura de anelamento e o tamanho dos produtos de PCR obtidos utilizando cada par de primer. A construção dos primers deve levar em conta o fato de que a maior parte das citosinas não estão metiladas no genoma e, por isso, devem ser substituídas por timina. Além disso, os primers devem anelar em trechos sem CpG, pois esses são os principais alvos das enzimas Dnmts.

Tabela 3.3 Oligonucleotídeos para amplificação de regiões dos genes estudados.

Os oligonucleotídeos foram desenhados a fim de permitir a amplificação de DNA tratado com bissulfito de sódio. T_m: temperatura de anelamento para o par de *primers*. CDS: região codificadora do gene. Os oligonucleotídeos para o promotor de Gap43 foram utilizados anteriormente (MACDONALD et al., 2010), os demais foram desenhados em nosso laboratório.

Sequência	Oligonucleotídeo (5'-N...N-3')	T _m	Tamanho do amplicon
Promotor <i>Gap43</i>	AATGTTTGATTGTTGTGTGTGTGT	50 °C	525 pb
	TCCCTTTCTCAAATTATTTACT		
Promotor <i>Olfr17</i>	TGTGTTATGATTGGTACTCTTG	56 °C	505 pb
	CCATCCCATTATCTAGTAGACTC		
CDS <i>Olfr17</i> (início)	AGATGTTGAGTATTTTGATATT	54 °C	418 pb
	ACATAAAAACCCAAAATCAAC		
CDS <i>Olfr17</i> (final)	ATAGTTTTAGTTTGTGTTGATA	50 °C	555 pb
	CAATATTTCTTCAACATCCT		
CDS <i>Olfr714</i>	TTTAGGAATTTATTTTTATGGAG	50 °C	443 pb
	AATATCATAACAACCAAATAATC		
CDS <i>Olfr1507</i>	TTTTTTTATTGTTATGGTTTATG	46 °C	468 pb
	TTATCCTCTAAAAAACTAATAAC		

Os reagentes e as concentrações das soluções, utilizados na reação de polimerização em cadeia da polimerase (PCR), são apresentadas na Tabela 3.4.

Tabela 3.4. Concentração dos componentes na reação de PCR com DNAbiss.

Componente	Concentração estoque	Concentração final
Tampão Rxn (Invitrogen)	10 x	1 x
MgCl ₂	50 mM	1,5 mM
dNTPs	10 mM	0,24 mM
<i>Primers</i> (forward e reverse)	25 µM	0,6 µM
Betaina	5 M	1 M
<i>Platinum Taq</i> DNA polimerase	5 U/µL	1,25 U

Além dos componentes listados na Tabela 3.4, adicionou-se 25 ng de DNAbiss e completou-se com água livre de DNase até 50 µL, que corresponde ao volume total da reação. A mistura foi submetida aos seguintes ciclos em um termociclador: denaturação inicial (2 minutos a 94 °C); 35 ciclos: denaturação (1 minuto a 94 °C), anelamento DNA-primer (1 minuto, vide Tabela 3.3, coluna “T_m”), extensão da cadeia pela *Platinum Taq* DNA polimerase (45 segundos a 72 °C); etapa de finalização (10 minutos a 72 °C) e mantido até a utilização a 4 °C. O produto de PCR foi separado em gel de agarose, para a aferição do tamanho da amplicon e purificado do gel utilizando-se o sistema (No. 28606, MinElute Gel Extraction Kit – QIAGEN).

3.2.6 Reação de ligação de produto de PCR em pCRII

O produto de PCR purificado do gel de agarose foi ligado ao vetor pCRII por meio de uma reação, catalisada por DNA ligase, entre o fragmento produzido na PCR e o plasmídeo linearizado (pCRII). O produto de PCR possuía adenina não pareadas na extremidade enquanto o plasmídeo possuía uma guanina não pareada criando, assim, extremidades coesivas entre os componentes a serem ligados. Para isso, adicionou-se 1 μL tampão de reação (10x, Invitrogen), 1 μL pCRII ($25 \text{ ng} \cdot \mu\text{L}^{-1}$, Invitrogen) e 0,5 μL de T4 DNA ligase ($5 \text{ U} \cdot \mu\text{L}^{-1}$, Invitrogen) a 7,5 μL do produto de PCR, seguido de incubação por 16 horas a 14°C .

3.2.7 Transformação de bactérias competentes

Misturou-se 100 μL de suspensão de bactérias competentes com 3,0 μL de uma solução de plasmídeo previamente ligado ao fragmento produzido por PCR. Essa mistura foi mantida por 30 minutos no gelo. Após essa incubação, foi dado choque térmico em banho a 42°C por 75 segundos e transferido para o gelo por mais 2 minutos. Adicionou-se 1,0 mL de meio de cultura LB (Luria-Bertani: 1,0 g de triptona; 0,5 g de extrato de levedura; 1,0 g de NaCl; 100 mL de água; ajustar para pH 7) e incubou-se por uma hora a 37°C sob agitação. Após a incubação, a suspensão foi centrifugada por um minuto a $2000 \times g$ descartando-se 900 μL do sobrenadante utilizando-se uma pipeta. As bactérias sedimentadas foram ressuspensas com o meio de cultura restante no tubo pipetando cuidadosamente. Plaqueou-se as bactérias em placa LB-ágar-carbenicilina ($50 \mu\text{g} \cdot \text{mL}^{-1}$) previamente recoberta com 40 μL X-gal ($20 \text{ mg} \cdot \text{mL}^{-1}$) e incubou-se por 16 horas a 37°C .

Após a incubação, as bactérias foram selecionadas por meio da coloração com X-gal, na qual colônias de bactérias que apresentaram coloração azul haviam sido transformadas com plasmídeo sem inserto. Enquanto as bactérias que possuíam coloração branca haviam sido transformadas com plasmídeo contendo o inserto que, por essa razão, impedia a expressão do gene da β -galactosidase presente no plasmídeo. Dessa forma, colônias brancas foram selecionadas e colocadas para crescer em 3,0 mL de meio de cultura LB com carbenicilina ($50 \mu\text{L} \cdot \text{mL}^{-1}$) para gerar as miniculturas. Incubados a 37°C por 16 horas.

3.2.8 Purificação de plasmídeo de minicultura bacteriana

A extração dos plasmídeos dos clones cultivados em minicultura foram feitas utilizando o sistema QIAprep Spin Miniprep kit cat. no. 27106 - Qiagen. Resumidamente, centrifugou-se 1,5 mL do inóculo de bactéria descartando o sobrenadante. Ressuspendeu-se com solução salina, adicionou-se tampão de lise e agitou-se bastante. Adicionou-se a solução de neutralização para ajustar o pH e aumentar a afinidade do DNA à coluna de purificação. Centrifugou-se a solução por 10 min e adicionou-se o sobrenadante à coluna descartando os *debris* de células bacterianas. A coluna foi lavada com solução contendo etanol, seguida da eluição do DNA com água ou tampão 10 mM Tris-HCl, pH 8.

3.2.9 Reação de sequenciamento

Para a reação de sequenciamento, utilizou-se o sistema BigDye® Terminator v3.1 Cycle Sequencing Kit cat. no. 4337455 – Life Technologies™. Mistura reacional: 3 µL de tampão de sequenciamento (3x), 0,37 µL de solução do primer Sp6 (20 µM), 1 µL do mix BigDye Terminator Sequencing (Taq polimerase e desoxirribonucleotídeos modificados). À mistura, adicionou-se 250 ng de DNA plasmidial e água até completar 15 µL. A mistura foi submetida aos seguintes ciclos em um termociclador: denaturação inicial (2 minutos a 95 °C); 35 ciclos: denaturação (30 s a 96 °C), anelamento DNA-primer (15 s a 50 °C), extensão da cadeia pela Taq DNA polimerase (4 minutos a 60 °C); mantido a 4 °C até a utilização.

3.2.10 Purificação de DNA de reação de sequenciamento

Adicionou-se a reação de sequenciamento: 23 µL de etanol absoluto; 1 µL de glicogênio (1 g·L⁻¹); 1 µL de acetato de sódio 3 M, pH 5,5. Essa solução foi mantida por 15 minutos no gelo, seguido de centrifugação por 40 minutos, 3220 x g a 4 °C. Descartou-se o sobrenadante por inversão seguido de centrifugação com os tubos invertidos por 1 minuto, 700 rpm a 4 °C. Adiciona-se ao pellet 50 µL de etanol 70% (4 °C), seguido de centrifugação por 10 min, 4000 rpm a 4 °C. Descartou-se o sobrenadante e deixou secar a temperatura ambiente. O produto da reação de sequenciamento foi, então, separado por eletroforese capilar acoplado a

fluorímetro (ABI PRISM® 3100GeneticAnalyzer/HITACHI) que coleta os dados de fluorescência e transforma na sequência de DNA por meio do programa 3100 Data Collection.

3.3 Busca de TFBS na região promotora do gene *Olf17*

A busca por sítios de ligação para fatores de transcrição (TFBS) na região promotora do gene *Olf17* foi realizada por meio da ferramenta *DNA Pattern Search* disponível no site (<http://www.geneinfinity.org/>). Os motivos utilizados na busca foram descritos previamente por Iwata (2017), Michaloski (2006) e Vassalli (2002).

3.4 Separação de células ativadas por fluorescência (FACS)

3.4.1 Linhagem transgênica P2-IRES-tauGFP

Para a realização do FACS, foram utilizados camundongos da linhagem P2-GFP (STOCK *Olf17*tm7Mom/MomJ – *Stock number*: 006669) obtida do *The Jackson Laboratory* e mantida no Biotério do Conjunto das Químicas-USP, com permissão de uso concedida pela Comissão de Ética em Uso de Animais (CEUA), sob protocolos no.19/2013 e no. 50/2017. O alelo modificado foi gerado por recombinação homóloga em células tronco embrionárias da linhagem 129P2/OlaHsd. Os animais gerados foram retrocruzados com animais C57BL/6J por pelo, menos quatro, gerações. No biotério do Conjunto das Químicas a linhagem foi mantida por meio do cruzamento entre animais heterozigotos para o alelo transgênico.

3.4.2 Genotipagem da linhagem P2-IRES-tauGFP

A linhagem utilizada possui a construção IRES-tauGFP inserida três pares de base após o códon de parada do gene *Olf17*. Dessa forma, a proteína GFP atua como gene repórter da expressão do gene de receptor olfatório *Olf17*, uma vez que a expressão de GFP está sob o controle do promotor desse receptor. A sequência IRES permite que o ribossomo inicie a tradução a partir de um sítio interno do mRNA, levando à síntese de uma proteína não fusionada com o receptor *olf17*. O fragmento *tau* é expresso em fusão à proteína GFP e tem

a função de interagir com microtúbulos, permitindo que o GFP fique distribuído por todo o corpo celular e axônios (FEINSTEIN; MOMBAERTS, 2004). Para identificar a presença do gene repórter GFP inserida no alelo do gene *Olf17*, foram utilizados dois pares de *primers*: F_1 e R_1 para amplificar um trecho da sequência do alelo selvagem de *Olf17* e F_2 e R_2 para amplificar um trecho da sequência de GFP, Figura 3.4.

Figura 3.4. Representação esquemática do alelo transgênico da linhagem P2-IRES-tauGFP.

O constructo IRES-tauGFP está inserido três pares de base (3 pb) após o códon de terminação da região codificadora (CDS) do gene *Olf17*. A sequência IRES (*internal ribosome entry site*) permite que o mRNA do gene *Olf17* seja traduzido a partir deste ponto. A sequência tauGFP codifica para uma proteína fluorescente verde que se liga a microtúbulo. As setas indicam a localização dos *primers* utilizados para a genotipagem dos animais dessa linhagem, em que o índice 1 refere-se a um anelamento dentro da sequência selvagem do gene *Olf17*, enquanto o índice 2 indica um anelamento dentro da sequência GFP.

Na ausência da sequência IRES-tauGFP (alelo selvagem), os *primers* F_1 e R_1 produzem amplicons de 217 pb, enquanto que, na presença desse alelo, o tempo de extensão utilizado no PCR (1 minuto) não permite a extensão com esse *primers*. O par de *primers* F_2 e R_2 geram produtos de 187 pb (Tabela 3.5).

Tabela 3.5. Primers utilizados na genotipagem de camundongos transgênicos P2-IRES-tauGFP.

Primer	Sequência (5'-N...N-3')	Amplicon
F_1	TTGAGGCGAATCATTACAG	217 pb
R_1	CAGTCAGCCTGCAGTATTTCC	
F_2	ACGTAAACGGCCACAAGTTC	187 pb
R_2	AAGTCGTGCTGCTTCATGTG	

Primeiramente, foi coletado um pequeno pedaço da orelha dos camundongos a serem genotipados e incubando-os com 300 μ L de solução de NaOH 50 mM por 12 minutos a 95 °C sob agitação. O pH foi ajustado com 25 μ L do tampão Tris-HCl 1,0 M e pH 8. As concentrações dos componentes utilizados na reação de polimerização em cadeia da polimerase (PCR), são apresentadas na Tabela 3.6.

Tabela 3.6. Concentração dos componentes na reação de PCR com gDNA.

Componente	Concentração estoque	Concentração final
Tampão Rxn (Invitrogen)	10 x	1 x
MgCl ₂	50 mM	1,5 mM
dNTPs	10 mM	0,24 mM
Primers (<i>forward</i> e <i>reverse</i>)	25 µM	0,6 µM
Taq DNA polimerase Recombinante	5 U.µL ⁻¹	1,25 U

Além dos componentes listados na Tabela 3.6 foram adicionados 2,0 µL de DNA e água sem DNase até completar 25 µL, que corresponde ao volume total da reação. A mistura foi submetida aos seguintes ciclos em um termociclador: denaturação inicial (3 minutos a 94 °C); 35 ciclos: denaturação (45 s a 94 °C), anelamento DNA-primer (45 s a 61 °C), extensão da cadeia pela Taq DNA polimerase Recombinante (45 s a 72 °C); etapa de finalização (6 minutos a 72 °C); manter a 4 °C até a aplicação no gel de agarose. O produto de PCR foi analisado em gel de agarose 2% em tampão TAE (Tris-acetato-EDTA) em pH 8,0.

3.4.3 Dissociação celular

Para realizar a separação das células GFP+ foi necessário coletar células de epitélio olfatório que podem ser oriundas de animais homozigotos ou heterozigotos para o alelo *Olf17*-IRES-tauGFP. Neste trabalho, foram utilizados camundongos de ambos os genótipos com propósitos distintos a serem explicados junto a apresentação dos resultados. Foram utilizados animais com idade entre 6 semanas e 6 meses.

O MOE foi dissecado em meio de cultura CMRL-1066 com EGTA 2 mM. O tecido dissecado foi mantido a 4 °C para melhor conservação das células. Esta suspensão foi centrifugada a temperatura de 4 °C por 5 minutos a 300 x g para remoção do sobrenadante. Ao tecido resultante foi adicionado 5 mL de CMRL-1066 com 1,0 µL de DNase (250 U·mL⁻¹) incubando a mistura por 5 minutos a temperatura ambiente (TA). Em seguida a suspensão foi novamente centrifugada a 4 °C por 5 minutos e 300 x g, a fim de sedimentar as células suspensas para remoção parcial do sobrenadante (manter 2 mL para a próxima etapa). Ao conteúdo de células foram adicionados 400 µL de papaína (100 U·mL⁻¹; Sigma; No. P 4762) mantendo a mistura por 10 minutos a 37 °C. Para remoção da enzima, a suspensão celular

foi centrifugada por 5 min a 300 x g a 4 °C, seguida da remoção do sobrenadante repetida por duas vezes para uma remoção eficiente da protease. Ao tecido resultante, foi adicionado 5 mL de meio CMRL-1066 com EGTA 2 mM e a suspensão foi filtrada em malha de 40,0 µm. Após a filtragem, a suspensão pôde ser submetida à separação das células no FACS.

3.4.4 Separação de células ativada por fluorescência (FACS)

A técnica consiste em passar células em suspensão por um citômetro de fluxo, que está acoplado a um analisador e a um conjunto de placas elétricas carregadas. Quando uma célula marcada com algum fluoróforo é detectada, o analisador envia um sinal, a fim de que a gota que contém aquela célula seja carregada eletricamente. Assim, a gota desvia sob ação das placas eletricamente carregadas permitindo a separação das células fluorescentes, Figura 3.5. O equipamento utilizado para separação das células foi o FACS ARIA III, instalado no Centro de Facilidades para Pesquisa (CEFAP) – ICB IV-USP. Após a separação das células, elas foram centrifugadas a 300 x g e o sobrenadante foi removido. As células resultantes foram submetidas ao método de DNAbiss-seq como descrito na Seção 3.1.

Figura 3.5. Representação esquemática de um equipamento de FACS

3.5 RT-PCR e PCR em tempo real

3.5.1 Extração de RNA com reagente TRIzol

Dissecou-se o epitélio olfatório que foi mantido em 1,0 mL de reagente TRIzol em gelo seco. Macerou-se o tecido dissecado no gelo até o expô-lo completamente ao TRIzol. Centrifugou-se a 12000 x g por 10 minutos a 4 °C e coletou-se o sobrenadante. Adicionou-se 200 µL de clorofórmio e agitou-se vigorosamente, seguido de incubação por 3 minutos a TA. Centrifugou-se a 12000 x g por 15 minutos a 4 °C e coletou-se a fase aquosa. Adicionou-se 500 µL de isopropanol e agitou-se vigorosamente, mantendo por 10 minutos a TA. Após esse período, centrifugou-se a 12000 x g por 10 minutos a 4 °C e descartou-se o sobrenadante. Adicionou-se 1,0 mL de etanol 75% e agitou-se a reação. Centrifugou-se a 7500 x g por 5 minutos a 4 °C e descartou-se o sobrenadante, deixando o precipitado secar ao ar. Ressuspendeu-se o RNA em 50 µL de água sem ribonucleases e verificou-se a sua integridade em gel de agarose 1,5%.

3.5.2 Síntese de DNA complementar (cDNA)

Primeiramente, realizou-se a digestão do DNA residual utilizando 1,0 µg de RNA; 1,0 µL de tampão 10x de DNase RQ1 (Promega - Cat. # M610A); 1,0 µL DNase RQ1 (1U/µg de RNA) (Promega - Cat. # M198A); completou-se com H₂O até 10,0 µL. Incubou-se a reação por 30 minutos a 37 °C. Interrompeu-se a reação adicionando 1,0 µL de *stop solution* (Promega - Cat. # M199A) e incubou-se a 65 °C por 10 minutos. Para a síntese de cDNA, adicionou-se 1,0 µL de oligo dT (20 mM) seguido de aquecimento a 65 °C por 5 minutos e 10 minutos no gelo. À solução resultante, adicionou-se 4,0 µL de tampão 5x First-Strand; 1,0 µL de solução de dNTPs (10 mM); 2,0 µL de DTT; 0,5 µL de RNase OUT (Cat. No. 10777-019); e 1,0 µL de SuperScript II, os reagentes do kit da Invitrogen Cat. No. 18064-022. Incubou-se a 42 °C por duas horas seguido de interrupção da reação por aquecimento a 95 °C por 5 minutos. Removeu-se os resíduos de RNA adicionando 1,0 µL de RNase A seguido de incubação por 30 minutos a 37 °C. A solução final foi diluída 5 vezes para as utilizações posteriores.

3.5.3 PCR quantitativo (qPCR)

A quantificação da expressão do gene *olfr17* por PCR quantitativo foi realizada utilizando o Fast SYBR™ Green Master Mix (Cat. No.4385612). A reação procedeu com 7,5 µL do mix Fast SYBR Green, concentração final de primer 600 nM e 50 ng de cDNA inicial. O equipamento utilizado para realizar as ciclagens do PCR e a captura de dados foi o 7500 Fast Real-Time PCR System – Applied Biosystems (20 segundos a 95 °C; repetição por 40 vezes 3 segundos a 95 °C e 30 segundos 60 °C). O gene escolhido como calibrador foi β -actina utilizando os primers, Act β _F: AAGCCAACCGTGAAAAGATG e Act β _R: GTGGTACGACCAGAGGCATACA); os primers para o gene alvo *olfr17*, P2-CDS_F: CTCTGATACTCCAGGACAAAACC e P2-CDS_R GGATCACAGATCGCCATGTAG; os primers para o gene repórter GFP, GFP_F: GTCTATATCATGGCCGACAAGC e GFP_R: GGTGTTCTGCTGGTAGTGGTC.

4 RESULTADOS

4.1 Conteúdo CpG em genes OR

A fim de caracterizar as regiões promotoras dos genes OR, quanto a quantidade e a distribuição de dinucleotídeos CpG ao longo dessas sequências, foram analisadas 198 regiões promotoras determinados empiricamente (MICHALOSKI et al., 2006). As sequências correspondem aos primeiros 600 nucleotídeos a montante do TSS e a quantidade de CpGs por sequência é apresentada na Figura 4.1. Cerca de 90% das sequências apresentam até quatro dinucleotídeos CpGs, com uma frequência média de (2 ± 2) CpGs por promotor, 66% menor que o número de CpGs esperados para o genoma de camundongo em um trecho de 600 nucleotídeos, que seria 6 CpGs (JABBARI; BERNARDI, 2004), mas concordante com distribuição não homogênea de CpG ao longo do genoma (GARDINER-GARDEN; FROMMER, 1987).

Figura 4.1. Distribuição do número de CpGs na região promotora de genes OR.

Histograma do número de sequências de regiões promotoras de OR pela quantidade de dinucleotídeos CpGs. A porcentagem sobre a barra representa a fração das 198 sequências analisadas.

A fim de verificar se existe algum padrão na distribuição dos dinucleotídeos CpGs ao longo dos 600 pares de base dos promotores de OR, foi contabilizado o número de sequências que possuía um CpG naquela posição em relação ao TSS, Figura 4.2. Dessa forma, foi possível verificar que os CpGs desses promotores se apresentam distribuídos de forma homogênea sem uma região preferencial para ocorrência deste dinucleotídeo.

Figura 4.2. Ocorrência de dinucleotídeos CpG em 198 regiões promotoras de genes OR.

A fim de investigar qual o papel da região codificadora (CDS) dos genes ORs no controle da sua expressão, um trabalho mostrou que, quando esses genes estão sob o controle de um promotor sintético (TetO), eles ainda mantêm o padrão de expressão monogênico e monoalélico em MOE (NGUYEN et al., 2007), indicando que as CDSs de genes OR poderiam conter algum tipo de elemento regulatório em sua sequência. Por isso, foram analisadas 864 CDSs de genes OR quanto à composição e distribuição dos dinucleotídeos CpG, de forma análoga a análise com as regiões promotoras. Em média, existem (8 ± 5) CpGs por sequência, semelhante aos 10 CpGs que seriam esperados ao acaso para os cerca de 1000 nucleotídeos de extensão dos genes OR (JABBARI; BERNARDI, 2004), Figura 4.3.

Figura 4.3. Distribuição do número de CpGs em 864 regiões codificadoras dos genes OR.

Histograma do número de sequências de regiões codificadora de genes OR pela quantidade de dinucleotídeos CpGs. A porcentagem sobre a barra representa a fração das 864 sequências analisadas, sendo omitidos valores abaixo de 4%.

A distribuição de CpGs ao longo da CDSs apresentou algumas regiões onde a ocorrência de CpGs era mais frequente (Figura 4.4). A região aparentemente mais conservada localiza-se entre os nucleotídeos 350 e 400, que coincide com a localização do códon para o aminoácido arginina (CGN) comumente encontrado no início da segunda alça intracelular de ORs (e GPCRs classe A em geral) (BUCK, L.; AXEL, 1991). Outras duas regiões ao redor dos nucleotídeos 500 e 900 também coincidem com um códon que codifica para arginina, conservado entre os ORs, além de uma região ao redor da posição 700 com, pelo menos, 5 sítios que codificam para serina (um dos códons é UCG), dois sítios conservados para prolina (um dos códons é CCG) e um sítio para treonina (um dos códons é ACG).

Figura 4.4. Ocorrência de dinucleotídeos CpG em 864 CDSs de genes OR.

Um dos parâmetros utilizados para a análise de regiões quanto ao conteúdo de dinucleotídeos CpG é o cálculo de CpG observado/esperado (Obs/Exp CpG) por meio da Equação 3.1 (GARDINER-GARDEN; FROMMER, 1987). Os resultados são apresentados na Figura 4.5, que relaciona a quantidade de seqüências com os valores determinados de Obs/Exp CpG. O valor médio de Obs/Exp CpG para as regiões promotoras foi 0,10 e para as regiões codificadoras foi 0,17, ambas abaixo da média para o genoma de camundongos no caso hipotético de uma distribuição homogênea dos CpGs pelo genoma (0,24). Os valores de Obs/Exp CpG são utilizados como parâmetro para definir uma ilha CpG (CGI). Para uma

região de, no mínimo, 200 pb, o valor de Obs/Exp CpG deve ser maior que 0,6 além de um conteúdo GC maior que 50% (GARDINER-GARDEN; FROMMER, 1987). As regiões promotoras não apresentaram valores de Obs/Exp CpG maiores que 0,39 e, portanto, não existem promotores de OR com CGI, consistente com o fato de que promotores OR são regiões ricas em AT (CLOWNEY et al., 2011). No caso das regiões codificadoras de OR, somente o gene *Olf281* pode ser caracterizada como uma ilha CpG, pois possui Obs/Exp CpG de 0,62.

Figura 4.5. Distribuição de Obs/Exp CpG em promotores e regiões codificadoras de genes OR

O histograma relaciona os valores de Obs/Exp CpG e a abundância com que eles são verificados entre regiões promotoras e CDSs de genes OR. O número sobre as linhas tracejadas corresponde às médias de Obs/Exp CpG em promotores OR (0,10), CDSs OR (0,17), no genoma de camundongos (0,24) e ilhas de CpG ($\geq 0,6$).

A partir dessa análise, podemos concluir que, se a metilação de DNA está envolvida na regulação da expressão dos genes OR, ela deve ocorrer sem a influência de ilhas CpG na região promotora ou no corpo dos genes OR.

4.2 Teste de eficiência do método DNAbiss-seq

Para estudar a metilação de DNA, optamos pelo sequenciamento de DNA pelo método de Sanger usando DNA tratado com bissulfito de sódio (DNAbiss-seq). Esse método permite determinar o padrão de metilação de qualquer citosina com resolução de base única, não apenas em dinucleotídeos CpG, além de estimar a frequência com que a metilação ocorre numa determinada base do genoma, Figura 4.6.

Figura 4.6. Determinação da frequência de mC (%mC) num grupo de células.

À esquerda algumas células são representadas enfatizando seu material genético traço cinza e duas citosinas, que podem estar metiladas (bola preta) ou não metilada (bola branca). Os clones obtidos ao final do método DNAbiss-seq ocorrem na mesma frequência que nas células de origem.

O método DNAbiss-seq foi testado utilizando um trecho da região promotora do gene *Gap43*, descrita previamente como diferencialmente metilada em MOE de embrião de 15,5 dias (E15.5) quando comparada com MOE de animais adultos (MACDONALD et al., 2010). A Figura 4.7 apresenta a comparação entre o padrão de metilação, de MOE adulto, obtida previamente (MacDonald et al., 2010) e o obtido neste trabalho, em que a metilação global do trecho analisado foi 67% e 69%, respectivamente.

Figura 4.7. Frequência de mCpG ao longo da região promotora do gene GAP43.

O gráfico apresenta as porcentagens de metilação em seis citosinas de dinucleotídeos CpGs presentes na região promotora do gene *Gap43*. Os dados foram obtidos em experimentos independentes: MacDonald 2010 e este trabalho.

4.3 Metilação de DNA no gene de receptor olfatório *Olfr17*

A fim de investigar a possível influência da metilação de DNA no controle da expressão dos genes OR, optamos por analisar em mais detalhes a metilação de DNA no gene de receptor *Olfr17*. Primeiramente, a estrutura do gene *Olfr17* (P2) foi determinada em nosso laboratório, pela Dra. Maíra H. Nagai, por meio da técnica de RLM-RACE (*RNA Ligase Mediated-Rapid Amplification of cDNA Ends*) esquematicamente apresentada na Figura 4.8. O gene do receptor olfatório P2 possui uma extensão de 5541 pb, dois éxons e, assim como a maioria dos OR, o segundo éxon é maior que o primeiro (em média 189 pb) (MICHALOSKI et al., 2006), além de a região codificadora (CDS) estar localizada integralmente no mesmo éxon. A coordenada obtida para o sítio de início da transcrição (crom7: 107093290) estava apenas seis pares de base a montante do TSS reportado previamente (crom7: 107093296) (CLOWNEY et al., 2011).

Figura 4.8. Representação esquemática do gene *Olfr17*

O gene de receptor olfatório (P2) possui dois éxons e a região codificadora (CDS) está integralmente no éxon 2. As coordenadas de início (TSS) e de final da transcrição são referentes ao cromossomo 7 da montagem GRCm38/mm10.

Primeiramente, nós comparamos a frequência de metilação na região promotora e codificadora do gene *Olfr17* em três diferentes tecidos: MOE de camundongos adultos, MOE de embriões de 15,5 dias de gestação (E15.5) e fígado. Os resultados referentes à metilação em dinucleotídeos CpG são apresentados na Figura 4.9. Dentre os tecidos comparados, apenas o MOE adulto apresentava OSNs maduros, portanto, capaz de expressar o gene *Olfr17*, ao contrário do fígado que não expressa esse receptor e MOE (E15.5), que é formado majoritariamente por células progenitoras e neurônios olfatórios imaturos. Nos três tecidos, analisados o gene *Olfr17* apresentou um padrão de metilação semelhante, tanto na região

promotora quanto na CDS, que corresponde 6% a 20% de metilação na região promotora e $\geq$ 58% na região codificadora. A elevada frequência de metilação (%mC) na região codificadora é consistente com a observação de que CpGs localizados fora de ilhas CpG são frequentemente metilados (ECKHARDT et al., 2006; JONES, P. A., 2012; SUZUKI; BIRD, 2008). Além disso, a metilação de DNA não parece ser tecido-específica para esse gene.

Figura 4.9. Frequência de metilação (%mC) no gene *Olfr17*.

A metilação em dinucleotídeos CpG da região promotora e codificadora do gene *Olfr17* são apresentadas indicando a frequência de metilação (%mC) em cada CpG, nos diferentes tecidos: MOE (adulto), MOE (E15.5) e fígado. Os números no eixo x indicam a posição da citosina em relação ao TSS, no caso da região promotora, e em relação ao primeiro nucleotídeo da região codificadora no caso da CDS. Os dados não possuem réplicas suficientes para fazer a análise estatística.

A metilação em nucleotídeos CpH (H = A, T e C) vem sendo mais recentemente estudadas, mas sua função ainda não está clara (KINDE et al., 2015; PATIL; WARD; HESSON, 2014). Por essa razão, decidimos analisar a metilação em contexto não-CpG, sobretudo, em dinucleotídeos CpA, que correspondem ao segundo principal alvo de Dnmt3a/b. Além disso, foi descrito que a região H apresenta metilação em contexto CpA preferencialmente em OSN (LOMVARDAS et al., 2006). Observamos que a metilação em CpA ocorre de forma aleatória, sem posições preferencialmente metiladas, e em baixa frequência na maioria das posições, não excedendo 40% de metilação na posição mais metilada (Figura 4.10).

Figura 4.10. Metilação em dinucleotídeos CpA do gene *Olfr17* em três diferentes tecidos.

O histograma relaciona a porcentagem de metilação (%mC) nos dinucleotídeos CpA ao longo da região promotora (A) e codificadora do gene (B: início; C: final) *Olfr17* em três diferentes tecidos: fígado, MOE (E15.5) e MOE adulto. A posição da citosina na CDS é em relação ao primeiro nucleotídeo do primeiro códon.

4.4 Metilação de DNA em outros genes OR

A fim de verificar o padrão de metilação de outros genes OR, decidimos analisar a metilação na região codificadora dos genes *Olfr1507* (MOR28) e *Olfr714* (P4). Assim como para o gene *Olfr17*, ambos apresentaram elevada frequência de metilação nos CpGs de suas

regiões codificadoras (Figura 4.11A). Curiosamente, a frequência de metilação global (média %mC) correlaciona-se com o nível de expressão destes ORs, verificado por sequenciamento de RNA em larga escala (RNA-seq) de MOE (Ibarra-Soria, 2017). Em outras palavras, entre os genes ORs analisados, quanto maior a frequência de metilação global da CDSs maior eram os valores de expressão como representados na Figura 4.11B. No entanto, um número maior de ORs deveriam ser analisados da mesma forma a fim de verificar se esta correlação se mantém para todos esses genes. Nos ORs *Olfr714* e *Olfr1507*, os dinucleotídeos CpAs, apresentaram metilação em baixa frequência, assim como para o gene *Olfr17* (dados não mostrados).

Figura 4.11. Metilação em CpG na região codificadora dos genes *Olfr714*, *Olfr17*, *Olfr1507*.

(A) Metilação de DNA nos seis primeiros dinucleotídeos CpGs da região codificadora dos genes *Olfr714*, *Olfr17*, *Olfr1507*. (B) Média da metilação dos dinucleotídeos CpGs apresentada no item (A). Os valores de FPKM correspondem ao nível de expressão do respectivo OR em MOE (IBARRA-SORIA et al., 2017). Os dados não possuem réplicas suficientes para fazer a análise estatística.

4.5 SNPs alteram o conteúdo CpG na região promotora de *Olfr17*

Com o intuito de verificar mais a fundo se a metilação de DNA influencia na expressão do gene *Olfr17*, nós decidimos utilizar animais transgênicos *Olfr17-IRES-tauGFP* para investigar o padrão de metilação nos genes em células que expressam o gene *Olfr17*. Para isso, foi necessário levar em consideração que o alelo transgênico (*Olfr17-IRES-tauGFP*) possui características da linhagem celular 129P2/OlaHsd (129), na qual esse camundongo foi produzido, gerando um animal com *background* genético misto no *locus* de inserção do constructo *Olfr17-IRES-tauGFP*. Baseado nisso, verificamos, por meio de sequenciamento de DNA genômico, que o animal transgênico apresentava dois polimorfismos de base única (SNPs) quando comparado com a linhagem C57BL/6 (B6), localizados nas posições -128 e +34 em região próxima ao TSS, Figura 4.12. Dessa forma, o alelo B6 apresenta um CpG nessa região (-83), enquanto o alelo 129 apresenta três CpGs (-129, -83 e 33).

Figura 4.12. SNPs no promotor de *Olfr17* alteram o conteúdo CpG.

No topo da imagem temos a representação de um trecho do gene *Olfr17* indicando em perspectiva onde estão localizados os SNPs e os CpGs ao redor do TSS. Na parte inferior da imagem são apresentados trechos da sequência de nucleotídeos para as linhagens C57BL/6 e 129P2/OlaHsd, daqui em diante denominado B6 e 129, respectivamente. Os números -128, -83 e 33 indicam as posições das citosinas que formam CpG na linhagem 129, enquanto na linhagem B6 apenas a posição -83 corresponde a uma citosina em CpG. As áreas hachuradas indicam as posições de ocorrência dos SNPs.

Os promotores de genes OR são, geralmente, enriquecidos em sítios de ligação para fatores de transcrição (TFBS) do tipo *Olf1/EBF-like* (*O/E-like*) e *homeodomínio-like* (HD) (CLOWNEY et al., 2011; MICHALOSKI et al., 2006; VASSALLI et al., 2002). Esses sítios de ligação foram, anteriormente, relacionados à frequência com que um OR é expresso e não com o controle da expressão monogênica e monoalélica dos genes OR (KHAN et al., 2011;

PLESSY et al., 2012; VASSALLI; FEINSTEIN; MOMBAERTS, 2011; ZHANG, G. et al., 2016). Foram encontrados três sítios *O/E-like* ao redor do TSS do alelo B6, ao passo que, no alelo 129, somente dois foram obtidos, ocasionado por conta do SNP 34, Figura 4.13. Se, ao invés da transição A→G fosse verificada a transversão A→T, o motivo ainda seria identificado, como pode ser verificado pelos gráficos das frequências relativas dos nucleotídeos nesses sítios de ligação (CLOWNEY et al., 2011; MICHALOSKI et al., 2006). Além dos sítios *O/E-like*, foram encontrados dois sítios HD que possuem o motivo hipotético TAATKR (IWATA et al., 2017). No entanto, não foram encontrados motivos com uma região consenso HAATTA (MICHALOSKI et al., 2006) e TAATTG (VASSALLI et al., 2002). Dessa forma, o gene *Olfr17* de 129 apresenta um TFBS *O/E-like* a menos que o de *Olfr17* de B6.

(A)

Fita direta:

TGTGTTATGATTGGTACTCTTGCAAATGCTAAATAAAACAGGAATGGATGAGAGGGAAAATGACAGAGAGAATGGGGTAGCCATAACTATT
TTTACATTATTGTTGGTTGCAGGCACCATGCTTGAACCAAGAAACAA**TAATGA**ACTAGTCAGATACAGTTAATGTTCCAATAAGAACTTAAA
ATTGAGTACATA**CTTCAAAGA**CATTTTAAAGAACTGCCATTTTGAAGAGGAAGAAGCATTAAAGTGAAATCATTTAGCTATGATAAACAGCAGGG
AA**TG**AAATAGTCTTGAGACCCAGGTTGGAATCCTCCAGAGTGACAGAC**CTGGGC**ACAGTGCAGAGTATCTCTGGAGTGATGAGCACTACACAGC
TCAGAGCCTGTTTACAAGAACTGGGCCAACCAAGTGTACAAGATTTACCTGAGCTGACTGGCACAG**CCCCAAA**AGTGTCTTTCAGAACTTGACAA
TGGCACAACCCAGAGTCTACTAGA**TAATGG**GATGG

Reverso complementar:

CCATCCCATTATCTAGTAGACTCTGGGTTGTGCCATTGTCAAGTTCTGAAAGACACTTT**TGGGG**CTGTGCCAGTCAGCTCAGGTAAATCTTGTA
CACTGGTTGGCCAGTCTCTGTAACAGGCTCTGAGCTGTGTAGTGCTCATCA**CTCCAGAGA**TACTCTGCACTGTGCCCA**CGT**CTGTCACTCTG
GAGGATTCCAACCTGGGTCTCAAGACTATTTC**ATTCC**CTGCTGTTTATCATAGCTAAATGATTTCACTTAATGCTTCTTCCCTTCAAATGGC
AGTTCTTAAATGTCTTTGAAGTATGTACTCAATTTAAGTTTCTTATTGGAACATTAACGTATCTGACTAGTTCATTATTGTTTCTTGGTT
TCAAGCATGGTGCCTGCAACCAACAATAATGTAAAAATAGTTATGGCTACCCATTCTCTGTCAATTTCCCTCTCATCCATTTCTGGTTT
ATTTTAGCATTTGCAAGAGTACCAATCATAACACA

(B)

Figura 4.13. Sítios de ligação *O/E-like* e homeodomínio-like na região promotora do gene *Olfr17*. (A) Sequência do locus do gene *Olfr17* utilizada para busca de TFBS *O/E-like* e homeodomínio-like. O trecho sublinhado representa os nucleotídeos a montante do TSS. Os nucleotídeos em itálico e negrito indicam as posições -128, -83 e 33. As regiões hachuradas em verde e azul correspondem aos sítios de ligação *O/E-like* e homeodomínio-like encontrados nas fitas direta e reversa. (B) Representação esquemática da região promotora do gene *Olfr17* nas linhagens B6 e 129 indicando as posições relativas dos motivos de ligação *O/E-like* e homeodomínio-like (base mais ao centro do motivo: -267, -208, -59, 34 e 85) e a localização dos nucleotídeos -128, -83 e 33, correspondentes à citosina em CpG na linhagem 129. Os nucleotídeos -412 e 93 correspondem ao primeiro e ao último nucleotídeo da sequência analisada.

4.6 O padrão de metilação se altera com o conteúdo CpG

A fim de aprofundar o estudo sobre a metilação de DNA no gene *Olfr17* optamos por alterar abordagem adotada até então. Para isso, analisamos se o padrão de metilação do gene *Olfr17* poderia estar alterado em células que expressam esse receptor (e onde o *Olfr17* estaria, portanto, ativo) comparado com o MOE total (onde o *Olfr17* estaria inativo). Para isto, lançamos mão da técnica de separação de células ativada por fluorescência (FACS) a fim de obter células GFP+ oriundas do animal transgênico *Olfr17-IRES-tauGFP* (FEINSTEIN; MOMBAERTS, 2004). Nesses animais, OSNs que expressam o gene *Olfr17* coexpressam a proteína fluorescente verde (GFP), o que possibilita a purificação dessas células a partir de uma suspensão de células do MOE (Figura 4.14).

Figura 4.14. Representação esquemática da separação das células *Olfr17*-tauGFP por FACS

Na parte superior da imagem é apresentada a construção *Olfr17*-tauGFP presente no animal transgênico. Ao centro, uma imagem de um corte coronal marcado com um corante nuclear (DAPI) e células verdes expressando o alelo *Olfr17*-tauGFP. Na parte inferior, dois gráficos de pontos relacionando a intensidade de fluorescência pelo tamanho da partícula (FSC-A), onde dois retângulos internos delimitam a área de seleção das células GFP+. Animais selvagens para o alelo *Olfr17* não possuem células dentro do retângulo (à esquerda), ao contrário dos animais transgênicos (*Olfr17*-tauGFP) que apresentam células nessa região (à direita).

Após coletar as células GFP+, o DNA genômico foi extraído e submetido ao método DNAbiss-seq a fim de analisar o padrão de metilação do gene *Olf17* nessas células. Devido ao fato de essas células possuírem região promotora com sequência da linhagem 129, três dinucleotídeos CpG estão presentes na região promotora. Como observado anteriormente (Figura 4.9), a região promotora apresenta uma menor frequência de metilação em relação à CDS, tanto nas células de MOE total (44% promotor, 73% CDS) quanto nas células GFP+ (41% promotor, 88% CDS) (Figura 4.15). No entanto, essa diferença não é tão acentuada quanto a verificada a partir de células da linhagem B6 (Figura 4.9). Quando comparamos a metilação em cada posição na região promotora, encontramos que a redução na %mC na posição -83 é significativamente menor em células GFP+ e, portanto, nas células que expressam o gene *Olf17*. Apesar disso, a redução da %mC ocorreu no dinucleotídeo CpG onde o nível de metilação era o mais baixo entre todos os CpGs analisados, indicando que esse não deve ser um fator crucial para selecionar esse OR no OSN.

Figura 4.15. Padrão de metilação do gene *Olf17* em neurônios olfatórios GFP+.

O histograma apresenta a frequência de metilação (%mC) para o gene *Olf17* na região promotora e nos quatro primeiros CpGs da região codificadora desse gene. Duas amostras foram comparadas, MOE total e células GFP+. Os números no eixo x indicam a posição da citosina em relação ao TSS, no caso da região promotora, e em relação ao primeiro nucleotídeo da região codificadora no caso da CDS. Para a análise estatística da região promotora foi utilizado teste t, $n = 4$, sendo * $p < 0,05$ e ns: diferença não significativa. Os dados da CDS não possuem réplicas suficientes para fazer a análise estatística.

Caso a metilação tivesse algum papel regulatório na expressão monoalélica dos genes OR, poderíamos esperar, por exemplo, que em um neurônio olfatório o alelo inativo estivesse metilado, e o ativo desmetilado. Os valores observados de ~50% de mC nas posições -128 e +33, nas células GFP+ (Figura 4.15) poderiam estar indicando este tipo de marcação diferencial entre os dois alelos, materno e paterno. No entanto, os animais utilizados neste experimento são homozigotos para o alelo *Olfr17-IRES-tauGFP* e, portanto, os valores de %mC em células GFP+, correspondem à média dos dois alelos, e não apenas dos alelos ativos. Para tentar solucionar essa questão, optamos por realizar FACS utilizando animais heterozigotos para o alelo *Olfr17-IRES-tauGFP*, e, por meio dos SNPs na região promotora diferenciar os alelos ativos dos inativos (Figura 4.16A). Dessa maneira, seria possível determinar o padrão de metilação exclusivamente dos alelos ativos. O SNP na posição -128 não nos permite distinguir os alelos B6 e 129 nos clones sequenciados, pois o tratamento de DNA com bissulfito do alelo 129 pode converter a citosina daquela posição, gerando clones com timina o que torna os dois alelos iguais. Por outro lado, o SNP na posição +34 não sofre alteração de base após o tratamento com bissulfito podendo, então, ser empregado para distinguir os alelos ativo (129) do inativo (B6).

Ao contrário do que esperávamos, o padrão de metilação observado na análise dos heterozigotos (Figura 4.16B) se manteve similar ao encontrado anteriormente (Figura 4.15), mas com um aumento na %mC na posição -83 do alelo ativo, oposto à tendência de redução observada em animais homozigotos. A metilação nos clones que deram origem ao resultado mostrado na Figura 4.16B são apresentados na Figura 4.16C, em que é possível observar quase todas as combinações de posição metiladas e não metiladas, tanto no alelo *Olfr17-IRES-tauGFP* ativo quanto no inativo, indicando que não existe grande diferença no padrão geral de metilação nos dois casos.

Figura 4.16. Padrão de metilação do alelo ativo do gene *Olfr17*.

(A) Representação esquemática dos neurônios olfatórios separados por FACS, mostrando que, em células de animais homozigotos, há duas cópias para o alelo transgênico (*Olfr17-IRES-tauGFP*), contra uma cópia do alelo transgênico e uma do alelo selvagem em animais heterozigotos. O triângulo preto indica as posições dos dinucleotídeos CpGs da região promotora. (B) Frequência de metilação na região promotora do alelo transgênico inativo (de MOE total) e ativo (de neurônios GFP+). Para a análise estatística foi utilizado teste t, $n = 3$, sendo * $p < 0,05$ e ns: diferença não significativa. (C) Representação da metilação em CpG em todos os clones analisados, cujo %mC está apresentada no item (B). A bola preta representa uma citosina metilada, enquanto a bola branca indica ausência de metilação.

4.7 Expressão do gene *Olfr17* em diferentes contextos genéticos

Os genes OR são expressos de forma desigual no MOE, sendo alguns ORs escolhidos para serem expressos, mais frequentemente que outros (IBARRA-SORIA et al., 2017; YOUNG et al., 2003). Apesar da constatação de que o padrão de metilação de DNA, na região promotora de um gene OR, não se altera entre alelos transcricionalmente ativos e inativos, verificamos uma diferença clara no padrão de metilação quando são comparadas as regiões promotoras do alelo C57BL/6 (B6) e 129P2/OlaHsd (129) (Figura 4.17). Como descrito anteriormente (Figura 4.12), o alelo B6 possui apenas um CpG, enquanto o alelo 129 possui

três CpGs na região analisada ao redor do TSS. Os sítios CpG adicionais na região promotora da linhagem 129 geram dois novos alvos para as Dnmt, que, por razões ainda não descobertas, metila preferencialmente a citosina -128 e +33, frente a citosina -83, presente nas duas linhagens. É importante também destacar, que, além de adicionar um dinucleotídeo CpG, o SNP na posição +34 faz com que a linhagem 129 possua um sítio O/E-like a menos quando comparado com a linhagem B6 (Figura 4.13B).

Figura 4.17. Comparação da %mC na região promotora do alelo C57BL/6 e 129P2/OlaHsd.

Os dados de metilação de DNA em MOE, utilizados nesta figura, foram recuperados das Figuras 4.9 (adulto) e 4.15 (total).

O aumento da metilação global de uma região promotora é frequentemente associado com a mudança de afinidade de fatores que ligam DNA (BIRD, 2002; JONES, P. A., 2012), gerando pontos de ligação para proteínas da família MBD e dificultando a ligação de proteínas CTCF a regiões responsáveis por mediar a interação *enhancer*-promotor. Baseado nisso, decidimos investigar se o nível de expressão dos alelos 129 e B6 do gene *Olfr17* no MOE são iguais ou diferentes. Para isto, realizamos experimentos de RT-qPCR em três genótipos distintos: selvagem (*Olfr17*-GFP (-/-)), heterozigoto (*Olfr17*-GFP (+/-)) e homozigoto (*Olfr17*-GFP (+/+)). O primer utilizado para a quantificação da expressão de *Olfr17* anelava na região codificadora desse OR e, portanto, possibilitou a quantificação de transcritos tanto do alelo B6 quanto do alelo 129 (Figura 4.18A). O primer para a região codificadora de GFP, por sua vez,

só permite a quantificação do alelo 129, possibilitando a quantificação de apenas um alelo em animais heterozigotos.

O nível de expressão do gene *Olf17* em animais homozigotos GFP é significativamente menor quando comparado aos demais genótipos, Figura 4.18B. Além disso, a expressão de *Olf17* em animais heterozigotos foi intermediária à obtida para animais homozigotos GFP e selvagens, sugerindo a expressão de um determinado alelo depende majoritariamente dos elementos em *cis* desse *locus*, como verificado anteriormente para expressão de ORs em animais de genoma misto, em que um dos cromossomos era oriundo da linhagem B6 e o outro era da linhagem 129 (IBARRA-SORIA et al., 2017). A fim de verificar se a expressão do alelo 129 em animais heterozigotos reduzia à metade quando comparada a animais homozigoto GFP, foi realizado um RT-qPCR com *primers* que anelam na CDS de GFP. Nessa situação, foi verificado que a expressão do alelo 129, no animal heterozigoto corresponde a $(37 \pm 14)\%$ daquela observada em animais homozigotos GFP (Figura 4.18B), portanto não é diferente dos $(50 \pm 8)\%$ esperados para a expressão de apenas um alelo. Em animais heterozigotos, o valor de $\Delta\Delta Ct$, obtido com o *primer* *Olf17*, corresponde a expressão dos alelos B6 e 129 somados. A fim de descobrir qual a contribuição de cada alelo para a expressão do gene *Olf17* em animais heterozigotos, podemos considerar que o valor de $\Delta\Delta Ct$ obtido com *primer* GFP equivale a expressão do alelo 129. Dessa forma, podemos concluir que, na presença dos alelos B6 e 129 em um mesmo neurônio, o alelo B6 é preferencialmente expresso em relação ao alelo 129.

Figura 4.18. O alelo *Olfr17* da linhagem B6 é preferencialmente expresso ao alelo transgênico
 (A) Representação esquemática dos alelos *Olfr17* com características da linhagem B6 e 129. À esquerda da representação, estão indicados os padrões de metilação apresentados na Figura 4.17, para cada um dos alelos. À direita temos as regiões codificadoras de *Olfr17* e GFP, sendo que o alelo B6 não apresenta a sequência para a proteína GFP. As setas representam o local aproximado de anelamento dos pares de *primers* utilizados para o PCR quantitativo (qPCR), índice 1: *Olfr17* e índice 2: GFP. (B) RT-qPCR foi realizado para determinar a expressão relativa do gene *Olfr17* em MOE de três genótipos distintos: *Olfr17*-GFP (-/-), *Olfr17*-GFP (+/-) e *Olfr17*-GFP (+/+). O calibrador utilizado foi o gene β -actina e o nível de expressão foi normalizado em relação a expressão no animal *Olfr17*-GFP (+/+). Os valores de $\Delta\Delta C_t$ representam a média $\pm$ SEM (n=5, *Olfr17* e n=6, GFP). *Primer Olfr17*: utilizamos One-way ANOVA seguido de pós-teste Newman-Keuls. *Primer GFP*: foi utilizado teste t. *P<0,05; **P<0,01 e ***P<0,001.

5 DISCUSSÃO

5.1 Aspectos genéticos e expressão dos genes OR

Células progenitoras do olfato são capazes de formar neurônios olfatórios maduros através de diversos estágios de diferenciação. Nos estágios iniciais, neurônios olfatórios imaturos expressam múltiplos genes OR em níveis baixos (HANCHATE et al., 2015; TAN et al., 2015), sendo que, apenas um desses tem sua expressão estabilizada no final da diferenciação celular. Os processos envolvidos na escolha do gene OR apresentam vários níveis de regulação e englobam tanto aspectos genéticos quanto epigenéticos. Um dos aspectos genéticos a ser considerado para compreender o padrão de expressão dos ORs são os sítios de ligação *O/E-like* e *homeodomínio-like*, comumente encontrado em regiões promotoras desses genes (CLOWNEY et al., 2011; IWATA et al., 2017; MICHALOSKI et al., 2006), assim como verificado no promotor de *Olf17* (Figura 4.13). Os sítios O/E são alvos para fatores de transcrição O/E1, O/E2, O/E3 e O/E4, sendo que os últimos três apresentam expressão mais elevada no epitélio olfatório comparativamente ao O/E1 (WANG; BETZ; REED, 2002; WANG; TSAI; REED, 1997), que apresenta uma função mais pronunciada na diferenciação de linfócitos B. Animais *knockout* para as proteínas O/E2 e O/E3 apresentam, além da redução o tempo de vida desses animais, uma alteração na projeção dos axônios dos OSNs para o bulbo olfatório e uma aparente redução no número de axônios projetados (WANG et al., 2004).

Estudos conduzidos com camundongos deficientes para a proteína LHX2, que liga sítios homeodomínio, mostraram uma alteração drástica no desenvolvimento do MOE com ausência total de neurônios olfatórios maduros expressando ORs classe II, mas ainda expressando muitos ORs classe I (HIROTA; MOMBAERTS, 2004; HIROTA; OMURA; MOMBAERTS, 2007). Animais transgênicos para o receptor M71, gerados por injeção pró-nuclear, em geral, preservam o padrão de expressão por zona (VASSALLI et al., 2002), sugerindo que os elementos regulatórios da região promotora de M71 controlam a expressão por zona. No entanto, não foi possível verificar um padrão claro entre abundância e distâncias relativas dos

sítios O/E-*like* e HD com a expressão por zona dos genes OR (CLOWNEY et al., 2011). Outros trabalhos, verificaram que a remoção ou a alteração na sequência de sítios HD e O/E-*like* da região promotora causa uma redução na expressão dos genes OR associados a esse promotor (ROTHMAN et al., 2005; VASSALLI et al., 2011). Esses achados vão ao encontro do resultado obtido para a expressão do gene *Olfr17* (Figura 4.18), em que o alelo selvagem (B6) apresentou um nível de expressão cerca de três vezes maior que o alelo transgênico (129), cuja região promotora possui um sítio O/E-*like* a menos que a linhagem B6 devido a um SNP que alterou esse motivo O/E-*like* (Figura 4.13).

Foram reportados outros elementos regulatórios (*enhancers*) capazes de controlar a expressão dos genes OR, que atuam tanto em *cis*, tais como a região H, o elemento P e o elemento J (BOZZA et al., 2009; IWATA et al., 2017; SERIZAWA et al., 2003), como em *trans*, por meio do agrupamento de múltiplos *enhancers* localizados a menos de 100kb de distância de *clusters* de genes OR em diversos cromossomos (MARKENSCOFF-PAPADIMITRIOU et al., 2014). Assim como as regiões promotoras dos genes OR, os *enhancers* também são caracterizados pela presença de sítios de ligação HD e O/E-*like* (IWATA et al., 2017; NISHIZUMI et al., 2007). No entanto, a descoberta do elemento J, que controla exclusivamente a expressão dos ORs classe I (IWATA et al., 2017), sugere que o padrão de expressão por zonas pode ser governado pelos *enhancers*.

Ao contrário da maior parte das regiões promotoras no genoma de mamíferos, os promotores de genes OR são curiosamente ricos em nucleotídeos de adenina e timina (CLOWNEY et al., 2011), sugerindo que pode haver elementos regulatórios não canônicos controlando a expressão desses genes. Além dos genes OR, foi verificado que genes do sistema imune e de resposta a xenobióticos também possuem elevados níveis de AT na região promotora. Esses grupos gênicos têm em comum o fato de responderem a estímulos externos que, às vezes, pode alterar o perfil de expressão desses genes (SAXONOV et al., 2006). No entanto, ainda não está claro quais tipos de interações são afetados, a nível molecular, em decorrência da elevação no conteúdo GC nas regiões promotora dos genes OR.

5.2 Metilação de DNA e expressão dos genes OR.

A metilação de DNA pode afetar, positiva ou negativamente, a ligação de fatores de transcrição a seus respectivos motivos de ligação (JONES, P. A., 2012; MOORE et al., 2013). Apesar de a metilação do DNA afetar a diferenciação dos neurônios olfatórios (COLQUITT et al., 2013; COLQUITT et al., 2014; MACDONALD et al., 2005; MACDONALD et al., 2010), ainda não está claro se ela age diretamente na expressão dos genes OR. Para tentar esclarecer essa questão, nós analisamos o *status* da metilação de DNA no gene *Olfr17* em diferentes situações.

Com as novas tecnologias de análise do genoma em larga escala, foi possível descobrir que o genoma de mamíferos se encontra globalmente metilado em dinucleotídeos CpG, com exceção de ilhas CpG que, em regiões regulatórias, são majoritariamente desmetiladas, por vezes protegidas por fatores envolvidos na transcrição (JONES, P. A., 2012; SUZUKI; BIRD, 2008). O padrão de metilação de DNA do gene *Olfr17* em três diferentes situações, MOE (adulto), MOE (E15.5) e fígado, apresenta um perfil similar ao descrito para o genoma, ou seja, baixos níveis de metilação em regiões regulatórias (promotor) e elevado nível de metilação nas demais regiões (CDS), nas três condições analisadas (Figura 4.9). Outro trabalho já havia descrito que o nível de metilação em CpG da CDS do gene *Olfr151* era elevado, tanto em espermatozoide quanto em MOE (DIAS; RESSLER, 2014). Da mesma forma, encontramos que a metilação não ocorre de forma tecido específica, nas regiões analisadas, além de se manter constante desde os estágios iniciais da formação do OSN.

O padrão monoalélico de expressão dos genes OR indica que, de alguma forma, os dois alelos são diferentemente marcados para que não seja permitida a expressão simultânea desse alelo no mesmo OSN. A metilação de DNA é uma marca epigenética que pode ocorrer de forma distinta entre os dois alelos, tal como ocorre na inativação do cromossomo X e no *imprinting* do gene *Igf2*. Se a exclusão alélica dos genes OR fosse regulada dessa forma, esperaríamos um nível de metilação do gene *Olfr17* ao redor de 50% (já que metade dos

alelos seriam totalmente metilados e metade não metilados). No entanto, isso não foi observado para o gene *Olf17*, tanto para a região promotora (3%) quanto para a CDS (~80%), o que não exclui a possibilidade de haver outras regiões próximas a esse gene OR que sejam marcadas diferentemente.

O nível de metilação na CDS dos genes *Olf1714* e *Olf1507* são, em média, tão elevados quanto os observados para *Olf17* e *Olf151* (Figura 4.11). Curiosamente, o nível de metilação global da região codificadora correlaciona-se positivamente com o nível de transcrição desses genes no MOE (Figura 4.11B). Nesse sentido, foi demonstrado que os níveis de metilação no corpo do gene também foram correlacionados positivamente com o aumento nos níveis de transcrição desses genes (WU et al., 2010), o que reforça a observação feita para os genes OR. Os diferentes níveis de metilação nas CDSs de OR pode ser decisivo no estágio em que múltiplos OR estão sendo transcritos e que apenas um tem sua transcrição estabilizada. Essa hipótese ainda precisa ser verificada para um número maior de genes OR, levando em conta que ORs expressos na mesma zona do MOE devem ser comparados entre si, pois são esses OR que estarão competindo para terem sua expressão estabilizada.

Dada a grande variedade fenotípica dos neurônios do MOE, gerada pela expressão singular do gene OR dentre os mais de 1000 disponíveis, poderíamos imaginar que o padrão de metilação observado até o momento para o gene *Olf17* corresponde à marcação desse gene na forma transcricionalmente inativa, uma vez que a maior parte das células do MOE são de células que não expressam *Olf17*. Por essa razão, utilizamos um equipamento de FACS para separar células fluorescentes do camundongo transgênico (*Olf17*-IRES-tauGFP), e assim analisar o padrão de metilação de células que expressavam o gene *Olf17*. No entanto, nossos dados não indicaram variações de metilação na maior parte do CpGs analisados (Figura 4.15 e 4.16), com exceção da citosina na posição -83, frequentemente encontrada com menor nível de metilação e que apresentou maior flutuação na %mC em células GFP+. A região codificadora do gene *Olf17* também manteve o padrão de elevados níveis de metilação (Figura 4.15) e, comparativamente, a %mC foi maior em células GFP+ ($88 \pm 5\%$) quando comparadas a células de MOE total ($73 \pm 8\%$), indo ao encontro da observação

que relacionou aumento de expressão com metilação na CDS de *Olfr17*, *Olfr714* e *Olfr1507*. Juntos, estes resultados indicam que a metilação de DNA não atua como uma marca do alelo ativado no neurônio olfatório, mas não exclui a possibilidade dessa modificação poder influenciar na escolha do OR pelo OSN.

Apesar dos resultados de metilação de DNA nos genes *Olfr17* em células GFP+ indicarem que ela não se altera ao longo da diferenciação do OSN, é importante ressaltar que o método de DNAbiss-seq não permite distinguir a modificação 5mC de 5hmC, que corresponde ao produto formado no primeiro estágio do processo de demetilação de DNA por enzimas do complexo TET (KOHLI; ZHANG, 2013). Por essa razão, não podemos excluir a possibilidade de que parte dos resultados correspondam a 5hmC e não 5mC.

Os alelos B6 e 129 do gene *Olfr17* apresentaram dois SNPs próximos ao TSS que levaram a situação de duas regiões regulatórias distintas em número de TFBS e conteúdo CpG (Figura 4.12 e 4.13). Acompanhando a alteração no conteúdo CpG, foi verificado um aumento global na metilação desses dinucleotídeos (B6: 3%; 129: 46%). Juntos, esses resultados podem auxiliar a explicar o menor nível de expressão do alelo 129 quando comparado ao alelo B6, pois, além do fato do alelo 129 possuir um TFBS a menos que o alelo B6, a maior incidência de citosinas metiladas ao longo da região promotora pode prejudicar a ligação de fatores importantes à transcrição, mesmo que ela não coincida com os TFBS. Dessa forma, a metilação na região promotora pode influenciar na probabilidade de um gene OR ser escolhido e ter sua expressão estabilizada (Figura 5.1). Essa observação, precisa ser confirmada para outros genes OR através da comparação de dados de metilação de DNA com expressão desses genes.

Figura 5.1. Modelo de ativação de gene OR dependente da região promotora.

Os alelos B6 e 129 apresentam conteúdo CpG (triângulos pretos) diferentes na região promotora do gene *Olfr17* o que altera o padrão de metilação daquela região e os sítios de ligação para fatores de transcrição. A seta curva indica a interação entre o *enhancer*(E) e a região promotora do gene OR. Uma seta mais larga indica uma maior probabilidade/força de interação *enhancer*-promotor.

5.3 Contexto genômico e variabilidade na expressão dos genes OR

Em um trabalho publicado recentemente, foi descrito que o nível de expressão dos genes OR varia significativamente entre três diferentes linhagens de camundongos (C57BL/6J, 129S5SvEv e CAST/EiJ) (IBARRA-SORIA et al., 2017). Um dos critérios utilizados para descrever a diferença entre as linhagens foi a quantidade de SNPs na região dos genes OR. Linhagens com um maior número de SNPs apresentaram um número maior de genes OR diferencialmente expressos, indicando que os SNPs são os responsáveis pela variabilidade no repertório do OSNs. Em outras palavras, os indivíduos 1 e 2 de uma linhagem isogênica 1 apresentam praticamente o mesmo repertório de OSNs, assim como os indivíduos 1 e 2 de outra linhagem isogênica 2 (Figura 5.2). Em contrapartida, indivíduos de linhagens diferentes apresentaram repertório de OSNs diferentes.

Figura 5.2. O repertório de genes OR varia entre as linhagens de camundongos.

Indivíduos da Linhagem 1 apresentam entre si o mesmo repertório de OSNs, assim como os indivíduos da linhagem 2. O repertório de OSNs corresponde à quantidade desses neurônios que expressam um mesmo OR, indicado aqui pelas cores branca (OSN 1) e cinza (OSN 2). Indivíduos de linhagens diferentes apresentam repertório de OSNs diferentes. Baseado em (IBARRA-SORIA et al., 2017).

No homem, a percepção dos diversos odorantes no ambiente varia entre os indivíduos, como demonstrado em um trabalho que avaliou a resposta de indivíduos de diversas nacionalidade, gêneros e idades quando expostos a diversos odorantes (KELLER et al., 2012). Outro trabalho demonstrou que o receptor olfatório humano OR7D4 perde a capacidade de se ligar a androstenona e androstadienona, *in vitro*, em decorrência de mutações no DNA que levaram a alteração na composição de aminoácidos desse receptor (KELLER et al., 2007). Grande parte da variabilidade inter-indivíduos se deve, portanto às alterações na sequência de aminoácidos de ORs específicos (MAINLAND et al., 2014).

Curiosamente, um estudo que analisou genomas humanos, de diversas regiões do planeta, disponíveis no banco de dado de *1000 Genome Project*, verificou que a ocorrência de SNPs em promotores de OR humanos é mais elevada do que a observada ao longo do genoma, mas menor do que a encontrada em promotores de processamento e apresentação de antígeno (IGNATIEVA et al., 2014). Associados aos nossos resultados de expressão do gene *Olf17*, a incidência de SNPs na região promotora e perfil de metilação de DNA podem contribuir na variação do repertório de neurônios olfatórios entre os indivíduos. Nossos

achados contribuem, portanto, para adicionar mais um nível de complexidade na forma como a percepção do cheiro pode variar entre os indivíduos.

6 CONCLUSÕES

Primeiramente, o principal alvo de metilação no DNA, os dinucleotídeos CpG, não estão em regiões de ilha CpG nas regiões promotora e codificadora dos genes OR.

Além disso, nosso trabalho determinou o padrão de metilação das regiões promotora e codificadora do gene *Olfr17*, com resolução de base única. Analisando diferentes tecidos e estágios de maturação do epitélio olfatório, foi possível concluir que metilação no gene *Olfr17* não ocorre de forma tecido específica nem varia ao longo do desenvolvimento do OSN. No mesmo sentido, a metilação na região promotora do gene *Olfr17* não atua como um marcador de genes OR transcricionalmente ativos determina a expressão monogênica e monoalélica dos genes OR.

A análise do padrão de metilação de outras CDSs de gene OR nos permitiu verificar uma correlação positiva entre a metilação global na CDS e o nível de expressão desse gene em MOE. No entanto, um número maior de genes OR precisa ser analisado para validar essa observação.

Finalmente, polimorfismos de base única na região promotora de *Olfr17* alteram o conteúdo CpG, o padrão de metilação de DNA e a quantidade de sítios de ligação O/E-like, entre as linhagens C57BL/6 e 129P2/OlaHsd. Associado a essas alterações foi verificado que o alelo da linhagem B6 é mais preferencialmente expresso do que o da linhagem 129, indicando que as alterações na região promotora influenciam a probabilidade de escolha do gene *Olfr17* para ser expresso pelo neurônio olfatório.

7 REFERÊNCIAS

- ABDUS-SABOOR, I. et al. An Expression Refinement Process Ensures Singular Odorant Receptor Gene Choice. **Curr Biol**, v. 26, n. 8, p. 1083-90, Apr 25 2016.
- ACHE, B. W.; YOUNG, J. M. Olfaction: diverse species, conserved principles. **Neuron**, v. 48, n. 3, p. 417-30, Nov 3 2005.
- AMIR, R. E. et al. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. **Nat Genet**, v. 23, n. 2, p. 185-8, Oct 1999.
- ARMELIN-CORREA, L. M. et al. Nuclear compartmentalization of odorant receptor genes. **Proc Natl Acad Sci U S A**, v. 111, n. 7, p. 2782-7, Feb 18 2014.
- AXEL, R. The molecular logic of smell. **Sci Am**, v. 273, n. 4, p. 154-9, Oct 1995.
- BELL, A. C.; FELSENFELD, G. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. **Nature**, v. 405, n. 6785, p. 482-5, May 25 2000.
- BELLUSCIO, L. et al. Mice deficient in Golf are anosmic. **Neuron**, v. 20, n. 1, p. 69-81, Jan 1998.
- BIRD, A. DNA methylation patterns and epigenetic memory. **Genes Dev**, v. 16, n. 1, p. 6-21, Jan 1 2002.
- BIRD, A. et al. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. **Cell**, v. 40, n. 1, p. 91-9, Jan 1985.
- BOZZA, T. et al. Mapping of class I and class II odorant receptors to glomerular domains by two distinct types of olfactory sensory neurons in the mouse. **Neuron**, v. 61, n. 2, p. 220-33, Jan 29 2009.
- BUCK, L.; AXEL, R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. **Cell**, v. 65, n. 1, p. 175-87, Apr 5 1991.
- BUCK, L. B. Information coding in the vertebrate olfactory system. **Annu Rev Neurosci**, v. 19, p. 517-44, 1996.
- CHESS, A. et al. Allelic Inactivation Regulates Olfactory Receptor Gene-Expression. **Cell**, v. 78, n. 5, p. 823-834, Sep 9 1994.
- CLOWNEY, E. J. et al. Nuclear aggregation of olfactory receptor genes governs their monogenic expression. **Cell**, v. 151, n. 4, p. 724-37, Nov 9 2012.
- CLOWNEY, E. J. et al. High-throughput mapping of the promoters of the mouse olfactory receptor genes reveals a new type of mammalian promoter and provides insight into olfactory receptor gene regulation. **Genome Res**, v. 21, n. 8, p. 1249-59, Aug 2011.
- COLQUITT, B. M. et al. Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. **Proc Natl Acad Sci U S A**, v. 110, n. 36, p. 14682-7, Sep 3 2013.
- COLQUITT, B. M. et al. Dnmt3a regulates global gene expression in olfactory sensory neurons and enables odorant-induced transcription. **Neuron**, v. 83, n. 4, p. 823-38, Aug 20 2014.

DALTON, R. P.; LYONS, D. B.; LOMVARDAS, S. Co-opting the unfolded protein response to elicit olfactory receptor feedback. **Cell**, v. 155, n. 2, p. 321-32, Oct 10 2013.

DEGL'INNOCENTI, A.; D'ERRICO, A. Regulatory Features for Odorant Receptor Genes in the Mouse Genome. **Front Genet**, v. 8, p. 19, 2017.

DEMARIA, S.; NGAI, J. The cell biology of smell. **J Cell Biol**, v. 191, n. 3, p. 443-52, Nov 1 2010.

DIAS, B. G.; RESSLER, K. J. Parental olfactory experience influences behavior and neural structure in subsequent generations. **Nat Neurosci**, v. 17, n. 1, p. 89-96, Jan 2014.

DUNCAN, B. K.; MILLER, J. H. Mutagenic deamination of cytosine residues in DNA. **Nature**, v. 287, n. 5782, p. 560-1, Oct 9 1980.

ECKHARDT, F. et al. DNA methylation profiling of human chromosomes 6, 20 and 22. **Nat Genet**, v. 38, n. 12, p. 1378-85, Dec 2006.

EGGAN, K. et al. Mice cloned from olfactory sensory neurons. **Nature**, v. 428, n. 6978, p. 44-9, Mar 4 2004.

EHRLICH, M. et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. **Nucleic Acids Res**, v. 10, n. 8, p. 2709-21, Apr 24 1982.

FEINSTEIN, P.; MOMBAERTS, P. A contextual model for axonal sorting into glomeruli in the mouse olfactory system. **Cell**, v. 117, n. 6, p. 817-31, Jun 11 2004.

FIRESTEIN, S. How the olfactory system makes sense of scents. **Nature**, v. 413, n. 6852, p. 211-8, Sep 13 2001.

FROMMER, M. et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. **Proc Natl Acad Sci U S A**, v. 89, n. 5, p. 1827-31, Mar 1 1992.

FUSS, S. H.; OMURA, M.; MOMBAERTS, P. Local and cis effects of the H element on expression of odorant receptor genes in mouse. **Cell**, v. 130, n. 2, p. 373-84, Jul 27 2007.

GAILLARD, I.; ROUQUIER, S.; GIORGI, D. Olfactory receptors. **Cell Mol Life Sci**, v. 61, n. 4, p. 456-69, Feb 2004.

GARDINER-GARDEN, M.; FROMMER, M. CpG islands in vertebrate genomes. **J Mol Biol**, v. 196, n. 2, p. 261-82, Jul 20 1987.

GODFREY, P. A.; MALNIC, B.; BUCK, L. B. The mouse olfactory receptor gene family. **Proc Natl Acad Sci U S A**, v. 101, n. 7, p. 2156-61, Feb 17 2004.

GRAZIADEI, P. P.; GRAZIADEI, G. A. Neurogenesis and neuron regeneration in the olfactory system of mammals. I. Morphological aspects of differentiation and structural organization of the olfactory sensory neurons. **J Neurocytol**, v. 8, n. 1, p. 1-18, Feb 1979.

GUO, J. U. et al. Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. **Nat Neurosci**, v. 17, n. 2, p. 215-22, Feb 2014.

HANCHATE, N. K. et al. Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. **Science**, v. 350, n. 6265, p. 1251-5, Dec 4 2015.

HELLMAN, A.; CHESS, A. Gene body-specific methylation on the active X chromosome. **Science**, v. 315, n. 5815, p. 1141-3, Feb 23 2007.

HENDRICH, B.; BIRD, A. Identification and characterization of a family of mammalian methyl-CpG binding proteins. **Mol Cell Biol**, v. 18, n. 11, p. 6538-47, Nov 1998.

HIROTA, J.; MOMBAERTS, P. The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory sensory neurons. **Proc Natl Acad Sci U S A**, v. 101, n. 23, p. 8751-5, Jun 8 2004.

HIROTA, J.; OMURA, M.; MOMBAERTS, P. Differential impact of Lhx2 deficiency on expression of class I and class II odorant receptor genes in mouse. **Mol Cell Neurosci**, v. 34, n. 4, p. 679-88, Apr 2007.

HUANG, Y. et al. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. **PLoS One**, v. 5, n. 1, p. e8888, 2010.

IBARRA-SORIA, X. et al. Variation in olfactory neuron repertoires is genetically controlled and environmentally modulated. **Elife**, v. 6, Apr 25 2017.

IGNATIEVA, E. V. et al. Genetic basis of olfactory cognition: extremely high level of DNA sequence polymorphism in promoter regions of the human olfactory receptor genes revealed using the 1000 Genomes Project dataset. **Front Psychol**, v. 5, p. 247, 2014.

ISHII, T. et al. Monoallelic expression of the odourant receptor gene and axonal projection of olfactory sensory neurones. **Genes Cells**, v. 6, n. 1, p. 71-8, Jan 2001.

IWATA, T. et al. A long-range cis-regulatory element for class I odorant receptor genes. **Nat Commun**, v. 8, n. 1, p. 885, Oct 12 2017.

JABBARI, K.; BERNARDI, G. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. **Gene**, v. 333, p. 143-9, May 26 2004.

JONES, D. T.; REED, R. R. Golf: an olfactory neuron specific-G protein involved in odorant signal transduction. **Science**, v. 244, n. 4906, p. 790-5, May 19 1989.

JONES, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. **Nat Rev Genet**, v. 13, n. 7, p. 484-92, May 29 2012.

JUNG, D.; ALT, F. W. Unraveling V(D)J recombination; insights into gene regulation. **Cell**, v. 116, n. 2, p. 299-311, Jan 23 2004.

KAMBERE, M. B.; LANE, R. P. Co-regulation of a large and rapidly evolving repertoire of odorant receptor genes. **BMC Neurosci**, v. 8 Suppl 3, p. S2, Sep 18 2007.

KELLER, A. et al. An olfactory demography of a diverse metropolitan population. **BMC Neurosci**, v. 13, p. 122, Oct 10 2012.

KELLER, A. et al. Genetic variation in a human odorant receptor alters odour perception. **Nature**, v. 449, n. 7161, p. 468-72, Sep 27 2007.

KHAN, M.; VAES, E.; MOMBAERTS, P. Regulation of the probability of mouse odorant receptor gene choice. **Cell**, v. 147, n. 4, p. 907-21, Nov 11 2011.

KINDE, B. et al. Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. **Proc Natl Acad Sci U S A**, v. 112, n. 22, p. 6800-6, Jun 2 2015.

KOHLI, R. M.; ZHANG, Y. TET enzymes, TDG and the dynamics of DNA demethylation. **Nature**, v. 502, n. 7472, p. 472-9, Oct 24 2013.

LANCET, D. Vertebrate olfactory reception. **Annu Rev Neurosci**, v. 9, p. 329-55, 1986.

LEE, W. et al. MeCP2 regulates activity-dependent transcriptional responses in olfactory sensory neurons. **Hum Mol Genet**, v. 23, n. 23, p. 6366-74, Dec 1 2014.

LEWIS, J. D. et al. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. **Cell**, v. 69, n. 6, p. 905-14, Jun 12 1992.

LI, E.; BESTOR, T. H.; JAENISCH, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. **Cell**, v. 69, n. 6, p. 915-26, Jun 12 1992.

LI, J. et al. Odorant receptor gene choice is reset by nuclear transfer from mouse olfactory sensory neurons. **Nature**, v. 428, n. 6981, p. 393-9, Mar 25 2004.

LI, Q.; HARJU, S.; PETERSON, K. R. Locus control regions: coming of age at a decade plus. **Trends Genet**, v. 15, n. 10, p. 403-8, Oct 1999.

LISTER, R. et al. Human DNA methylomes at base resolution show widespread epigenomic differences. **Nature**, v. 462, n. 7271, p. 315-22, Nov 19 2009.

LOMVARDAS, S. et al. Interchromosomal interactions and olfactory receptor choice. **Cell**, v. 126, n. 2, p. 403-13, Jul 28 2006.

LYONS, D. B. et al. An epigenetic trap stabilizes singular olfactory receptor expression. **Cell**, v. 154, n. 2, p. 325-36, Jul 18 2013.

MACDONALD, J. L.; GIN, C. S.; ROSKAMS, A. J. Stage-specific induction of DNA methyltransferases in olfactory receptor neuron development. **Dev Biol**, v. 288, n. 2, p. 461-73, Dec 15 2005.

MACDONALD, J. L. et al. MBD2 and MeCP2 regulate distinct transitions in the stage-specific differentiation of olfactory receptor neurons. **Mol Cell Neurosci**, v. 44, n. 1, p. 55-67, May 2010.

MAGKLARA, A. et al. An epigenetic signature for monoallelic olfactory receptor expression. **Cell**, v. 145, n. 4, p. 555-70, May 13 2011.

MAINLAND, J. D. et al. The missense of smell: functional variability in the human odorant receptor repertoire. **Nat Neurosci**, v. 17, n. 1, p. 114-20, Jan 2014.

MALNIC, B.; GODFREY, P. A.; BUCK, L. B. The human olfactory receptor gene family. **Proc Natl Acad Sci U S A**, v. 101, n. 8, p. 2584-9, Feb 24 2004.

MALNIC, B. et al. Combinatorial receptor codes for odors. **Cell**, v. 96, n. 5, p. 713-723, Mar 5 1999.

MARKENSCOFF-PAPADIMITRIOU, E. et al. Enhancer interaction networks as a means for singular olfactory receptor expression. **Cell**, v. 159, n. 3, p. 543-57, Oct 23 2014.

MEEHAN, R. R. et al. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. **Cell**, v. 58, n. 3, p. 499-507, Aug 11 1989.

MICHALOSKI, J. S.; GALANTE, P. A.; MALNIC, B. Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences. **Genome Res**, v. 16, n. 9, p. 1091-8, Sep 2006.

MIYAMICHI, K. et al. Continuous and overlapping expression domains of odorant receptor genes in the olfactory epithelium determine the dorsal/ventral positioning of glomeruli in the olfactory bulb. **J Neurosci**, v. 25, n. 14, p. 3586-92, Apr 06 2005.

MOHANDAS, T.; SPARKES, R. S.; SHAPIRO, L. J. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. **Science**, v. 211, n. 4480, p. 393-6, Jan 23 1981.

MOMBAERTS, P. Seven-transmembrane proteins as odorant and chemosensory receptors. **Science**, v. 286, n. 5440, p. 707-711, Oct 22 1999.

MOMBAERTS, P. et al. Visualizing an olfactory sensory map. **Cell**, v. 87, n. 4, p. 675-86, Nov 15 1996.

MOORE, L. D.; LE, T.; FAN, G. DNA methylation and its basic function. **Neuropsychopharmacology**, v. 38, n. 1, p. 23-38, Jan 2013.

MOULTON, D. G.; BEIDLER, L. M. Structure and function in the peripheral olfactory system. **Physiol Rev**, v. 47, n. 1, p. 1-52, Jan 1967.

MUNGER, S. D.; LEINDERS-ZUFALL, T.; ZUFALL, F. Subsystem organization of the mammalian sense of smell. **Annu Rev Physiol**, v. 71, p. 115-40, 2009.

NGUYEN, M. Q. et al. Prominent roles for odorant receptor coding sequences in allelic exclusion. **Cell**, v. 131, n. 5, p. 1009-17, Nov 30 2007.

NISHIZUMI, H. et al. Deletion of the core-H region in mice abolishes the expression of three proximal odorant receptor genes in cis. **Proc Natl Acad Sci U S A**, v. 104, n. 50, p. 20067-72, Dec 11 2007.

OKANO, M. et al. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. **Cell**, v. 99, n. 3, p. 247-57, Oct 29 1999.

PATIL, V.; WARD, R. L.; HESSON, L. B. The evidence for functional non-CpG methylation in mammalian cells. **Epigenetics**, v. 9, n. 6, p. 823-8, Jun 2014.

PILPEL, Y.; LANCET, D. The variable and conserved interfaces of modeled olfactory receptor proteins. **Protein Sci** v. 8, p. 969-77, Jan 16 1999.

PLESSY, C. et al. Promoter architecture of mouse olfactory receptor genes. **Genome Res**, v. 22, n. 3, p. 486-97, Mar 2012.

PROBST, W. C. et al. Sequence alignment of the G-protein coupled receptor superfamily. **DNA Cell Biol**, v. 11, n. 1, p. 1-20, Jan-Feb 1992.

RAMSAHOYE, B. H. et al. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. **Proc Natl Acad Sci U S A**, v. 97, n. 10, p. 5237-42, May 9 2000.

RAZIN, A.; RIGGS, A. D. DNA methylation and gene function. **Science**, v. 210, n. 4470, p. 604-10, Nov 7 1980.

RESSLER, K. J.; SULLIVAN, S. L.; BUCK, L. B. A zonal organization of odorant receptor gene expression in the olfactory epithelium. **Cell**, v. 73, n. 3, p. 597-609, May 7 1993.

_____. Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. **Cell**, v. 79, n. 7, p. 1245-55, Dec 30 1994.

RODRIGUEZ-GIL, D. J. et al. Odorant receptors regulate the final glomerular coalescence of olfactory sensory neuron axons. **Proc Natl Acad Sci U S A**, v. 112, n. 18, p. 5821-6, May 5 2015.

ROTHMAN, A. et al. The promoter of the mouse odorant receptor gene M71. **Mol Cell Neurosci**, v. 28, n. 3, p. 535-46, Mar 2005.

RUZOV, A. et al. Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. **Cell Res**, v. 21, n. 9, p. 1332-42, Sep 2011.

SAXONOV, S.; BERG, P.; BRUTLAG, D. L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. **Proc Natl Acad Sci U S A**, v. 103, n. 5, p. 1412-7, Jan 31 2006.

SERIZAWA, S. et al. Mutually exclusive expression of odorant receptor transgenes. **Nature Neuroscience**, v. 3, n. 7, p. 687-693, Jul 2000.

SERIZAWA, S. et al. Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. **Science**, v. 302, n. 5653, p. 2088-94, Dec 19 2003.

SINGH, P.; LEE, D. H.; SZABO, P. E. More than insulator: multiple roles of CTCF at the H19-Igf2 imprinted domain. **Front Genet**, v. 3, p. 214, 2012.

SMALLWOOD, P. M.; WANG, Y.; NATHANS, J. Role of a locus control region in the mutually exclusive expression of human red and green cone pigment genes. **Proc Natl Acad Sci U S A**, v. 99, n. 2, p. 1008-11, Jan 22 2002.

SMITH, E.; SHILATIFARD, A. Enhancer biology and enhanceropathies. **Nat Struct Mol Biol**, v. 21, n. 3, p. 210-9, Mar 2014.

SULLIVAN, S. L. et al. The chromosomal distribution of mouse odorant receptor genes. **Proc Natl Acad Sci U S A**, v. 93, n. 2, p. 884-8, Jan 23 1996.

SUZUKI, M. M.; BIRD, A. DNA methylation landscapes: provocative insights from epigenomics. **Nat Rev Genet**, v. 9, n. 6, p. 465-76, Jun 2008.

TAN, L.; LI, Q.; XIE, X. S. Olfactory sensory neurons transiently express multiple olfactory receptors during development. **Mol Syst Biol**, v. 11, n. 12, p. 844, Dec 8 2015.

TROJER, P.; REINBERG, D. Facultative heterochromatin: is there a distinctive molecular signature? **Mol Cell**, v. 28, n. 1, p. 1-13, Oct 12 2007.

VASSALLI, A.; FEINSTEIN, P.; MOMBAERTS, P. Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice. **Mol Cell Neurosci**, v. 46, n. 2, p. 381-96, Feb 2011.

VASSALLI, A. et al. Minigenes impart odorant receptor-specific axon guidance in the olfactory bulb. **Neuron**, v. 35, n. 4, p. 681-96, Aug 15 2002.

VASSAR, R. et al. Topographic organization of sensory projections to the olfactory bulb. **Cell**, v. 79, n. 6, p. 981-91, Dec 16 1994.

VASSAR, R.; NGAI, J.; AXEL, R. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. **Cell**, v. 74, n. 2, p. 309-18, Jul 30 1993.

WANG, S. S.; BETZ, A. G.; REED, R. R. Cloning of a novel Olf-1/EBF-like gene, O/E-4, by degenerate oligo-based direct selection. **Mol Cell Neurosci**, v. 20, n. 3, p. 404-14, Jul 2002.

WANG, S. S. et al. Genetic disruptions of O/E2 and O/E3 genes reveal involvement in olfactory receptor neuron projection. **Development**, v. 131, n. 6, p. 1377-88, Mar 2004.

WANG, S. S.; TSAI, R. Y.; REED, R. R. The characterization of the Olf-1/EBF-like HLH transcription factor family: implications in olfactory gene regulation and neuronal development. **J Neurosci**, v. 17, n. 11, p. 4149-58, Jun 1 1997.

WONG, S. T. et al. Disruption of the type III adenylyl cyclase gene leads to peripheral and behavioral anosmia in transgenic mice. **Neuron**, v. 27, n. 3, p. 487-97, Sep 2000.

WU, H. et al. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. **Science**, v. 329, n. 5990, p. 444-8, Jul 23 2010.

YOON, K. H. et al. Olfactory receptor genes expressed in distinct lineages are sequestered in different nuclear compartments. **Proc Natl Acad Sci U S A**, v. 112, n. 18, p. E2403-9, May 5 2015.

YOUNG, J. M. et al. Different evolutionary processes shaped the mouse and human olfactory receptor gene families. **Hum Mol Genet**, v. 11, n. 5, p. 535-46, Mar 1 2002.

YOUNG, J. M. et al. Odorant receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels. **Genome Biol**, v. 4, n. 11, p. R71, 2003.

ZHANG, G. et al. Lhx2 Determines Odorant Receptor Expression Frequency in Mature Olfactory Sensory Neurons. **eNeuro**, v. 3, n. 5, Sep-Oct 2016.

ZHANG, X.; FIRESTEIN, S. The olfactory receptor gene superfamily of the mouse. **Nat Neurosci**, v. 5, n. 2, p. 124-33, Feb 2002.

ZILLER, M. J. et al. Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. **PLoS Genet**, v. 7, n. 12, p. e1002389, Dec 2011.

8 ANEXOS

8.1 Anexo A – Súmula curricular

Dados Pessoais

Nome: Artur Guazzelli Leme Silva

Local e data de nascimento: Santo André-SP, 02/08/1988.

Formação

Julho de 2012 – atual

Doutorado Direto em Bioquímica e Biologia Molecular

Orientadora Profa. Dra. Bettina Malnic

Projeto de pesquisa: Metilação de DNA e expressão de genes de receptores olfatórios.

Instituição: Instituto de Química da Universidade de São Paulo (IQ-USP).

Abril de 2011 – julho de 2012

Iniciação científica em ciências dos materiais.

Orientador: Dr. Frank Ferrer Sene

Projeto: Desenvolvimento e Caracterização das Liga Ni-Mo Aplicadas ao Processo de Brasagem entre os Componentes do Catodo Termiônico Impregnado Tipo Reservatório para Emprego nas Válvulas TWT.

Instituição: Centro Tecnológico da Marinha em São Paulo (CTMSP).

Setembro de 2010 – fevereiro de 2011

Intercâmbio de estudo na Faculdade de Ciências da Universidade de Lisboa

Financiamento: Programa de Mobilidade Internacional Santander Países Ibéricos.

Julho de 2008 – maio de 2010

Iniciação científica no laboratório de Filmes Finos Poliméricos.

Orientador: Profa. Dra. Denise Freitas Siqueira Petri

Projeto: Imobilização e atividade hidrolítica de celulase sobre microesferas de vidro

Instituição: Instituto de Química da Universidade de São Paulo (IQ-USP).

Fevereiro de 2008 – julho de 2010

Iniciação científica no Grupo de Pesquisa em Ensino de Química (GEPEQ).

Orientador: Profa. Dra. Maria Eunice Ribeiro Marcondes

Projeto: Investigação sobre a Contextualização no Ensino de Química no curso de Licenciatura em Química do IQ-USP.

Instituição: Instituto de Química da Universidade de São Paulo (IQ-USP).

Fevereiro de 2007 – julho de 2012

Graduação em Química Ambiental (Bacharelado).

Instituição: Instituto de Química da Universidade de São Paulo (IQ-USP).

Fevereiro de 2004 – dezembro de 2006

Ensino médio na E.E. Pres. Arthur da Silva Bernardes, Cerquilha-SP.

Ocupação

Doutorado Direto

Bolsa concedida pela Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), agosto de 2012 a novembro de 2012.

Bolsa concedida pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), dezembro de 2012 a julho de 2017.

Programa de Aperfeiçoamento de Ensino (PAE)

1º semestre de 2014 – monitoria da disciplina QBQ 102 – Bioquímica e Biologia Molecular, sob supervisão de Ana Maria Carmona Ribeiro.

1º semestre de 2013 – monitoria da disciplina QBQ 212 – Biologia Molecular, sob supervisão de Pio Colepicolo Neto e Daniela Sanchez Bassères.

Representação Discente

Conselho de Pesquisa da Universidade de São Paulo, período 2º semestre de 2014 e 1º semestre de 2015.

Comissão de Pós-Graduação do Instituto de Química (USP), período 2014/2015.

Congregação do Instituto de Química (USP), período 2013.

Publicações

Artigos Completos

Iris R. M. Tébéka, Artur G. L. Silva e Denise F. S. Petri. Hydrolytic Activity of Free and Immobilized Cellulase. Langmuir 2009, v. 25 (3), p. 1582-1587.

Lúcia M. Armelin-Correa, Maíra H. Nagai, Artur G.L. Silva e Bettina Malnic. Nuclear architecture and gene silencing in olfactory sensory neurons. *BioArchitecture* 2014, v. 4, n. 4-5, p. 160-163.

Resumos em Congressos

Artur G. L. Silva, Maíra H. Nagai e Bettina Malnic. DNA methylation pattern of an odorant receptor gene. XI Congresso do Departamento de Bioquímica do IQ-USP, São Paulo, outubro de 2016.

Artur G. L. Silva, Lúcia M. Armellin-Corrêa e Bettina Malnic. Odorant and vomeronasal sensory neurons show similar nuclear architectures. *Genome Architecture in Cell Fate & Disease – Gordon Research Conference*, Hong Kong, China, julho de 2017.

Artur G. L. Silva, Maíra H. Nagai e Bettina Malnic. DNA Methylation in the Proximal Promoter of an Odorant Receptor Gene depends on CpG content. 46^a Annual Meeting of the Brazilian Society of Biochemistry and Molecular Biology, Águas de Lindóia, Brasil, julho de 2017.

Prêmios

Prêmio CRQ-IV, edição 2009, concedido pelo Conselho Regional de Química – IV Região, com o trabalho “Hidrólise de Celulose, Catalisada por Celulase Imobilizada: uma Alternativa para Produção de Etanol”.

Prêmio Lavoisier de melhor aluno do Curso de Bacharelado em Química Ambiental 2007/2012, concedido pelo Conselho Regional de Química – IV Região.

8.2 Anexo B – Artigo em preparação e submetido a disciplina QBQ5749 – Redação Científica no 1º semestre de 2017.

***Olf17* alleles from different mouse strains show distinct promoter DNA methylation patterns and expression levels**

Leme Silva, A. G; Nagai, M. H.; Malnic, B.

Department of Biochemistry, University of São Paulo, Av. Prof. Lineu Prestes, 748,
CEP 05508-000, São Paulo, SP, Brazil

* Corresponding author: Bettina Malnic

bmalnic@iq.usp.br

Abstract

Background

Odorant receptor (OR) genes belong to a large family composed of ~1000 genes in rodents. Distinct OR genes are expressed at different levels in the olfactory epithelium, with some being expressed at higher levels than others. DNA methylation plays an important role in gene regulation; however, it is not clear whether it is involved in the regulation of OR gene expression. Here we examined the DNA methylation pattern of the *Olf17* (P2) OR gene in the mouse olfactory epithelium.

Results

We found that while the coding region of the *Olf17* gene is highly methylated, its promoter region is unmethylated or shows low levels of methylation. No differences in the DNA methylation pattern were observed in olfactory neurons expressing *Olf17*, indicating that no significant changes in methylation are required for the activation of the OR gene. We also found that the *Olf17* gene allele from a different mouse strain, the 129 strain, contains single nucleotide polymorphisms (SNPs), when compared to the C57BL/6 strain allele, which lead to the occurrence of two additional CpGs in the promoter region. As a result, the promoter region of the 129 *Olf17* allele shows higher levels of DNA methylation than the promoter region from the C57BL/6 *Olf17* allele. Finally, we show that there is a correlation between the promoter DNA methylation pattern and the expression levels of the C57BL/6 and 129 *Olf17* alleles.

Conclusion

Our results demonstrate that genetic variations in regulatory promoter regions can contribute to unequal expression levels of OR gene alleles, which could ultimately lead to differences in the representation of the subtypes of odorant receptors in the olfactory epithelia from different mouse strains.

Keywords: DNA methylation, odorant receptor gene, olfactory receptor gene, bisulfite sequencing, epigenetics, gene promoter, genetic variation, gene expression, single nucleotide polymorphism.

Background

Odorants are detected by a large family of odorant receptors (ORs) expressed in the olfactory sensory neurons of the nose [1]. In the mouse genome, there are ~1000 different OR genes, which are spread in almost all chromosomes [2, 3]. Each olfactory sensory neuron in the olfactory epithelium expresses one single OR gene out of the complete repertoire, through a tight mechanism of gene regulation, which is not completely understood [4-7]. In addition, different members of the OR gene family show unequal levels of gene expression in the olfactory epithelium [8-12]. Recent experiments using high throughput transcript analysis in combination with in situ hybridization studies indicate that the expression levels of a given OR gene in the olfactory epithelium correlate with the number of neurons expressing this receptor gene, and not to the amount of the gene's transcripts per olfactory neuron [11, 12].

Previous studies have demonstrated that epigenetic modifications, such as histone post-translational modifications, must play important roles in the nuclear organization of the genome in olfactory neurons and in OR gene regulation [13-15]. DNA methylation is another type of epigenetic modification which plays important roles in the regulation of gene expression and nuclear organization [16, 17]. Methylation typically occurs at the carbon 5 position of cytosine in CpG dinucleotides, generating 5-methyl cytosine (5mC). DNA methylation in the promoter region of a gene has been generally considered to be a stable modification associated with transcriptional repression. However, recent genome-wide DNA methylation studies have revealed that promoter DNA methylation is dynamic during development indicating that the functional role played by DNA methylation is not clear and may be context dependent [18, 19]. Little is known about the role played by DNA methylation in OR gene regulation. In the olfactory epithelium, expression of the methyl-binding domain proteins MBD2 and MeCP2s, which bind to methylated CpGs and repress transcription, are required for olfactory sensory neuron differentiation [20]. The de novo DNA methyl transferases DNMT3a and DNMT3b are expressed in different developmental stages in the olfactory neuron lineage, when different sets of genes are expressed: while DNMT3b is expressed in mitotic progenitors, DNMT3a is expressed only in post-mitotic immature olfactory sensory neurons, prior to terminal neuronal differentiation [21]. OR gene expression occurs at the post-mitotic stage, and is required for the differentiation to mature olfactory sensory neurons [22, 23]. Thus, DNMT3a expression is induced during neuronal differentiation at the same time as OR gene choice is initiated.

Here we analyzed the DNA methylation pattern of the odorant receptor 17 gene (Olfr17, also named P2). We first determined the 5' gene structure and transcription start site (TSS) of the gene. We next used bisulfite sequencing to determine the levels of cytosine methylation in the promoter and coding region of Olfr17 in the olfactory epithelium from C57BL/6 mice. We found that while the CpGs in the coding region of Olfr17 are highly methylated, they show little or no

methylation in the gene's promoter region. We used FACS to purify GFP positive olfactory sensory neurons (i.e., Olfr17 expressing neurons) from P2-IRES-tauGFP transgenic mice and compared the DNA methylation pattern of Olfr17 in these neurons to the one in non-expressing (pre-sorted) olfactory sensory neurons. No significant differences were observed in the levels of DNA methylation, both in the coding and promoter regions of the gene. We also found that the Olfr17 allele from a different mouse strain, 129, contains single nucleotide polymorphisms (SNPs) leading to the occurrence of two additional CpGs in the promoter region, which are lacking in the C57BL/6 strain allele. As a result, the promoter region of the 129 Olfr17 allele shows higher levels of DNA methylation than the one from the C57BL/6 Olfr17 allele. Finally, by using quantification RT-PCR analysis we show that mice carrying the C57BL/6 alleles show higher levels of Olfr17 expression in their olfactory epithelium than mice carrying the 129 alleles.

Results

The Olfr17 gene structure

We first determined the 5' gene structure (exon and intron distribution) and transcription start site (TSS) of Olfr17 by using RLM-RACE (RNA ligase-mediate rapid amplification of cDNA ends) [24, 25]. The Olfr17 gene shows a structure that is similar to a large number of OR genes: it is composed by two exons, which are separated by a long intron (in this case, 2923 bps long) (Figure 1A) [24]. The coding sequence (CDS, 948 bp long) of the gene is located within the second exon. The genomic position of the transcription start site (TSS) (chr7:107093290) is highly coincident with the one previously identified by using a microarray based high-throughput method (chr7:107093296) [26].

Cytosine methylation in odorant receptor genes

We used DNA bisulfite sequencing (DNA bis-seq) to determine the DNA methylation profile in OR genes. This method allows us to identify the presence of 5mC with a single base resolution (Figure 1b). Since cytosine methylation may vary among cells, we sequenced a large number of clones to analyze the distribution of the methylated cytosines in single molecules. To check the efficiency of the bisulfite treatment to convert cytosine to thymine we first analyzed the methylation frequency in six CpG dinucleotides within the Gap43 promoter, which were previously shown to be highly methylated in the mature olfactory epithelium [20]. We obtained results that are highly similar to the previous ones, validating our assay (Figure 1c).

Since OR gene promoters are AT rich sequences, their CpG dinucleotides content is lower than the one observed across the mouse genome (around 1% versus 5,25%) [26, 27].

Accordingly, no CpG islands are present in the promoter regions of the OR genes [26]. Representative methylation results obtained for the promoter and coding region of the *Olfr17* gene in the olfactory epithelium from C57BL/6 mice (hereafter B6) are shown in Figures 1d and e, where filled squares represent the presence of a 5mC in each one of the cloned sequences. Red squares represent cytosine methylation in the context of CpG dinucleotides. There is only one CpG in the promoter region of the *Olfr17* gene, and this CpG shows an extremely low frequency of methylation (methylation was observed in only 6% of the analyzed sequences, Figure 1d). In contrast, each one of the six CpGs within the coding region shows high frequencies of cytosine methylation (varying from 65% to 94%, Figure 1e).

In general cytosine methylation occurs at CpG dinucleotides, however, non-CpG methylation has been increasingly reported, although its function is not understood [28-31]. Cytosine methylation in the CpA context was previously shown to occur in an OR gene enhancer (the H region) in the olfactory sensory neuron [32]. We therefore also analyzed the distribution of CpA methylation in the *Olfr17* gene. As we can see in Figures 1d and e, cytosine methylation in the CpA context (green squares) occur at low frequencies, both in the promoter region and coding region of the *Olfr17* gene (varying from 6% to 24%). Also, CpA methylation is randomly spread throughout the sequence, that is, there are no specific CpAs which are preferentially methylated in the *Olfr17* gene.

We compared the *Olfr17* gene cytosine methylation frequencies in the olfactory epithelium from adult mice, olfactory epithelium from E15.5 mouse embryos and in liver. They all show equal distributions of CpG methylation, with higher frequencies in the coding region ($\geq 58\%$) than in the promoter region (6% to 20%) (Figure 2a). The high CpG methylation frequencies we observe in the coding region of the *Olfr17* gene agree with the observations that CpGs which are located outside of CpG islands are usually methylated [16, 19]. The finding that the CpG within the *Olfr17* promoter region shows an inverse methylation profile, in turn, agrees with observations that unmethylated regions in the mammalian genome are enriched in the 5' region of the genes and promoter regions [16, 19]. In addition, the *Olfr17* methylation pattern does not seem to be tissue-specific, since it is similar in the liver (Figure 2a). Also, no significant differences in the methylation pattern were observed between the *Olfr17* gene from adult and embryonic olfactory epithelium, indicating that no evident changes occur during development (Figure 2a).

We also examined the cytosine methylation pattern in the coding regions of two additional OR genes, *Olfr1507* (MOR28) and *Olfr714* (P4). Like the *Olfr17* gene, both showed elevated CpG methylation frequencies (Figure 2b). Interestingly, these DNA methylation frequencies within the coding region correlated with the expression levels of these OR genes, as previously

determined by using RNA-sequencing from whole olfactory epithelium [12]. The mean methylation frequencies for these OR genes are 65% (*Olfr714*), 76% (*Olfr17*) and 94% (*Olfr1507*); and their expression levels are respectively 53 fpkm, 670 fpkm and 11101 fpkm [12]. Thus, our results suggest that higher CpG methylation levels in the coding region of the OR genes may correlate with higher gene expression levels.

***Olfr17* promoter DNA sequences differ between alleles from different strains**

To be able to analyze the role played by *Olfr17* gene methylation in expression we used P2-IRES-tauGFP transgenic mice. These mice express the *Olfr17* gene from the 129P2/OlaHsd (hereafter 129) mouse strain in a mixed B6/129 genetic background. By comparing the 5' region sequences of the *Olfr17* gene derived from these two mouse strains, we identified the presence of two single nucleotide polymorphisms (SNPs). These SNPs lead to the occurrence of two additional CpG dinucleotides in the 129 *Olfr17* allele, one in the promoter region and the other one in exon 1, which are absent from the B6 *Olfr17* allele (Figure 3a).

OR gene promoters are generally enriched in binding sites for O/E-like and homeodomain like transcription factors [24, 26, 33]. These and other identified motifs are likely to influence OR gene expression frequency in olfactory sensory neurons, rather than drive monogenic expression or regulate the levels of expression per cell [11, 24, 34-36]. We found that while there are three O/E-like motifs around the B6 *Olfr17* TSS sequence, in the 129-corresponding sequence one of these is missing due to the A→G SNP within exon 1 (around cytosine +33, Figure 3b).

CpG methylation within the *Olfr17* gene from P2-GFP+ sorted cells

Since each OR gene is typically expressed in a very low percentage of the olfactory sensory neurons (~0.1% in average) in the olfactory epithelium, the *Olfr17* DNA methylation pattern, which was obtained from genomic DNA extracted from the whole olfactory epithelium (shown in Figures 1 and 2), corresponds to that shown by neurons which do not express this OR gene. We next asked whether the levels of DNA methylation in the *Olfr17* gene are different in olfactory sensory neurons that express this OR gene. We performed Fluorescence-Activated Cell Sorting (FACS) using P2-IRES-tauGFP transgenic mice, which co-express GFP with the *Olfr17* gene [37] (Figure 4). Olfactory neurons were dissociated from the olfactory epithelium from homozygous P2-IRES-tauGFP transgenic mice and GFP positive *Olfr17* expressing neurons were separated by FACS. Genomic DNA from the sorted cells were submitted to the bisulfite sequencing procedure to analyze the cytosine methylation levels in the *Olfr17* gene. The methylation profiles of the promoter and part of the coding region of the *Olfr17* sequence (which in this case is derived from the 129 strain) obtained from pre-sorted cells (which do not express *Olfr17*) and sorted cells (which express *Olfr17*) are shown in Figure 5a. The *Olfr17*

coding region from both sorted and pre-sorted cells, show similar levels of cytosine methylation. The CpGs in the promoter region also show similar levels of methylation, except for the CpG -83, which showed a reduced level of methylation in the sorted cells (Figure 5a).

Because we used homozygous P2-IRES-tauGFP transgenic mice in these experiments, these methylation results are the sum from both, the active and inactive alleles present in each cell. We therefore devised a strategy that allowed us to analyze cytosine methylation only in the active *Olfr17* allele. We used the SNP +34 to discriminate between the GFP containing allele and the wild type allele during bisulfite sequencing (Figure 3a). We now used heterozygous P2-IRES-tauGFP transgenic mice to sort the GFP positive neurons, and used the genomic DNA from these cells in the bisulfite-sequencing experiments. In these heterozygous cells, only one of the two alleles contain the GFP coding sequence (and the SNP +34), and this must be the active allele in the sorted cells (Figure 5b). Comparison of the cytosine methylation frequencies in the *Olfr17* promoter region from sorted and pre-sorted cells showed however that the methylation levels in the CpGs located near the TSS were highly similar, with no evident reduction in the levels of methylation in the sorted neurons (Figure 5c). The CpGs in the active alleles show a random pattern of methylation, with no CpG site being preferentially methylated or unmethylated in the majority of the clones (Figure 5d, sorted). Altogether, these results show that no expressive changes in DNA methylation occur upon *Olfr17* gene activation.

DNA methylation and gene expression levels of the *Olfr17* alleles

As shown above, the *Olfr17* promoter DNA sequences differ between the alleles from the B6 and 129 mouse strains (Figure 3). We found that the level of CpG -83 methylation in the 129 *Olfr17* promoter in the olfactory epithelium is significantly higher when compared to the B6 *Olfr17* promoter (Figure 6a). Also, the presence of the two additional CpGs contribute to an overall higher DNA methylation level in the 129 *Olfr17* promoter when compared to the B6 *Olfr17* promoter.

Since methylation in promoter regions usually regulate gene expression levels, we analyzed whether the expression levels of *Olfr17* from the B6 and 129 alleles differed. To do this, quantitative RT-PCR was used to determine *Olfr17* gene expression in the olfactory epithelium from wild type (P2-GFP^{-/-}); heterozygous (P2-GFP^{+/-}) and homozygous (P2-GFP^{+/+}) transgenic siblings (Figure 5b). We found that the expression levels in the olfactory epithelium from homozygotes, where *Olfr17* is expressed from 129 alleles, is significantly reduced when compared to the expression levels from wild types, where the *Olfr17* is expressed from the B6 alleles (Figure 6c). Accordingly, heterozygotes, where *Olfr17* is expressed from one B6 and one 129 alleles, show intermediate levels of expression (Figure 6c).

We also compared the expression levels of GFP in homozygous and heterozygous neurons. As expected, GFP expression can be detected only in the olfactory epithelium from homozygous and heterozygous mice (Figure 6c). In heterozygotes, expression of GFP corresponds only to 25% of the expression of *Olfr17*, indicating that in the heterozygous cells, which contain one copy of each one of the alleles, B6 and 129, the B6 allele is preferentially transcribed.

Discussion

Olfactory progenitor cells generate mature olfactory sensory neurons through different stages of differentiation. In the initial stages, immature olfactory sensory neurons express multiple OR genes, and as differentiation proceeds, mature olfactory neurons are generated which express one single OR gene [38-40]. OR gene choice encompasses various levels of regulation of gene expression, involving both, regulatory DNA sequences and epigenetic modifications. Cis acting regulatory elements like the H and P enhancers [11, 41, 42] and trans acting enhancers [43] were shown to play important roles in OR gene expression. OR gene promoter and enhancer sequences are enriched in binding motifs for homeodomain and O/E-like transcription factors [24, 33, 44-46] and motifs for other transcription factors were also found within the regulatory elements [25, 26, 43]. Binding of transcription factors to their respective DNA binding motifs can be affected, both, negatively and positively, by cytosine methylation [47]. Even though there are evidences that DNA methylation plays a role in global gene expression in the olfactory epithelium [20, 21, 48, 49], the involvement of DNA methylation in OR gene regulation has been poorly investigated so far. Here we analyzed the DNA methylation status of the *Olfr17* gene, in different situations. Technologies that allow for genome wide analysis of DNA methylation in mammals have revealed that the genome is globally methylated in their CpGs, except for the ones which are localized within CpG islands and other regulatory gene regions, which would be bound by proteins involved in transcription and protected from methylation by DNMTs [16, 19, 50]. Consistent with these findings, we found that the coding region of the OR genes which were analyzed in this study show high levels of CpG methylation, and the promoter region shows lower methylation levels. The *Olfr17* gene DNA methylation pattern was the same in the two analyzed tissue types, olfactory epithelium, and liver (where OR genes are not expressed), indicating that the DNA methylation pattern we determined for the OR gene is not tissue specific. A previous study showed that the CpGs in the coding region from a different OR gene, *Ofr151*, are also highly methylated, both in the olfactory epithelium and in sperm [51]. In addition, we found no evident methylation differences between different development stages (embryonic and adult olfactory epithelium).

OR genes are expressed in a monoallelic fashion, that is, only one of the alleles in a given olfactory neuron is expressed, while the other one is repressed [4, 5]. Differential epigenetic marking of the two homologous OR gene alleles could explain monoallelic expression, that is, one of the alleles would be marked with repressive marks such as DNA methylation [6]. In this case, one would expect to find that ~50% of the sequences would be methylated and ~50% of the sequences would be unmethylated. However, this is not what is observed when the *Olfr17* DNA methylation pattern is analyzed in the olfactory epithelium from B6 mice. In this case, ~100% of the coding region is methylated, and ~6% of the promoter region is methylated. These results indicate that the two OR gene alleles are not differentially marked in the olfactory neurons, at least in terms of cytosine methylation.

The use of P2-IRES-tauGFP transgenic mice allowed us to analyze whether the DNA methylation patterns of the *Olfr17* gene is different when the gene is active. However, we did not observe significant differences in the methylation pattern of the *Olfr17* alleles in olfactory neurons expressing the *Olfr17* gene. These results indicate that no DNA methylation changes correlate with activation of an OR gene. Nevertheless, it is important to note that the method we used cannot discriminate 5-methyl cytosine (5mC) from 5-hydroxy methyl cytosine (5hmC). 5hmC is a derivative of 5mC due its oxidation by ten-eleven translocation (TET) enzymes, which was recently described as a new epigenetic mark enriched in the nervous system and in embryonic cells [29, 52]. Therefore, we cannot exclude the possibility that some of the methylated CpGs we detected in the OR gene are indeed 5hmCs. In the main olfactory epithelium, 5hmC was shown to be accumulated in neurons and neuronal progenitors and that genes that are highly transcribed contain high levels of 5hmC in their bodies [48]. In addition, loss of Dnmt3a, which is required for the generation of both 5hmC and 5mC during olfactory neuron differentiation, results in global perturbation in gene expression in the olfactory epithelium [49]. The presence of 5hmC in the OR genes was however not analyzed yet.

In general, transcriptional effects of methylation in the coding region are not clear. Previous studies have shown that DNA methylation within gene bodies (intragenic DNA methylation) is positively correlated with gene expression levels [53-56]. In agreement with these previous observations, we found that ORs that show higher levels of methylation also show higher levels of expression in the olfactory epithelium (Figure 2c).

The B6 and 129 *Olfr17* alleles contain cis proximal regulatory regions which differ in sequence and methylation levels. These variations may result in differential binding of transcription factors and other proteins, which ultimately regulate gene expression. In general, CpG methylation is thought to repress transcription by directly preventing binding of transcription

factors, or by recruiting histone deacetylases which promote chromatin condensation (although some transcription factors were recently shown to prefer methylated binding sites [47]). Accordingly, by comparing the expression of the B6 and 129 *Olfr17* alleles within the same mixed genetic background (129/B6), we found that the 129 allele (which shows the higher methylation in the promoter region) is expressed at a lower level. Whether the same correlation applies to other OR genes, or also occurs in other regulatory elements, such as the OR gene enhancers, requires further investigation.

Interestingly, a recent high-throughput analysis using human genome data from the 1000 Genomes Project revealed a high level of SNPs in the promoter regions of human OR genes [57]. Taken together with our results, these data suggest that this high level of polymorphism in the OR gene promoter regions may contribute to an unequal representation of the different subtypes of ORs in different human individuals.

Conclusions

In conclusion, our study has determined the DNA methylation profile of an OR gene with a single base resolution. We found that the coding region of the *Olfr17* gene in the olfactory epithelium is highly methylated in CpGs, and that the promoter region is unmethylated or very little methylated. The DNA methylation status of the *Olfr17* does not change in a different tissue, or in a different developmental stage. In addition, the active *Olfr17* allele also showed equal levels of methylation, making it evident that DNA methylation does not play a direct role in OR gene activation. Finally, the *Olfr17* alleles from the B6 and 129 mouse strains differ in their promoter DNA sequences, show different levels of CpG methylation, different composition in transcription factor binding motifs and different levels of expression. Our results agree with recent large-scale analysis of OR gene transcripts that DNA elements act in cis to regulate abundance of OR gene expression [12] and constitute an example of how genetic variations in regulatory promoter regions can contribute to differential expression of OR genes.

Methods

Animal procedures

The University of São Paulo Chemistry Institute's Animal Care and Use Committee approved all animal procedures, under the protocol number 19/2013. The mouse strains used were C57BL/6 and *Olfr17*^{tm7Mom/MomJ} (Stock number: 006669, obtained from The Jackson Laboratory).

RLM-RACE

RLM-RACE was performed using the GeneRacer™ Kit as previously described [24]. Total olfactory epithelium RNA was purified from C57BL/6 using TRIzol (Invitrogen, Cat. No. 15596026), following the user guide workflow. The following primers were used to amplify the *Olfr17* gene 5'-UTR: P2_R: TCCTGGAGTATCAGAGTACTC and P2_R_NESTED: CAGAGCAAGAGTCAGCTGTAG.

DNA bisulfite sequencing

Olfactory epithelium and liver were dissected from C57BL/6 or *Olfr17*^{tm7Mom/MomJ} transgenic mice in PBS pH 7.4 and the genomic DNA was immediately purified using DNeasy Blood & Tissue Kit (Qiagen, Cat. No. 69504). DNA bisulfite conversion was performed using EpiTect Bisulfite Kits (Qiagen, Cat. No. 59104). The regions of interest were amplified by PCR using the bisulfite converted DNA and the following primers: *Olfr17* promoter region (Forward TGTGTTATGATTGGTATTTTGG, Reverse CCATCCCATTATCTAATAAACTC); *Olfr17* CDS 1 (Forward AGATGTTGAGTATTTTGATATT, Reverse ACATAAAAACCCAAAATCAAC); *Olfr17* CDS 2 (Forward ATAGTTTTAGTTTGTGTTGATA, Reverse CAATATTTCTTCAACATCCT); CDS *Olfr714* (Forward TTTAGGAATTTATTTTTTATGGAG, Reverse AATATCATAACAACCAAATAATC); CDS *Olfr1507* (Forward TTTTTTTATTGTTATGGTTTATG, Reverse TTATCCTCTAAAAAACTAATAAC).

The amplicons were synthesized by Platinum® Taq DNA Polymerase (Invitrogen) and were subcloned using the Dual promoter TA cloning® kit (Invitrogen, 45-0007LT). Individual clones were sequenced using BigDye® Terminator v3.1 Cycle Sequencing Kit (Life Technologies™, Cat. No. 4337455) on an ABI PRISM 3100 Genetic Analyzer (Hitachi).

FACS

Whole olfactory epithelium was dissected in liquid culture medium CMRL-1066 supplemented with EGTA (2,0 mM) and kept on ice and cell dissociation was carried out as follows. Dissected tissue was incubated with DNase (1 U/μL, Invitrogen, Cat. No. 18068-015) for 5 minutes at room temperature. After DNase treatment, the tissue was incubated with 200 μL papain (100 U/mL, Sigma No. P 4762) for each 1,0 mL of dissociation solution, for 10 minutes at 37 °C. To remove papain, cells were centrifuged at 300 x g for 5 minutes at 4 °C, and the supernatant was discarded. Centrifugation was repeated one more time, and cells were resuspended in CMRL-1066 containing EGTA (2 mM). The cellular suspension was filtered in 40 μm mesh twice, and cells were sorted with a FACS Aria III in the Core Facility for Scientific Research – University of São Paulo (CEFAP-USP/FLUIR).

Motif analysis of the *Olf17* promoter

The *Olf17* promoter region was browsed for the transcription factor binding motifs using the SiteSeer tool (<http://www.chick.manchester.ac.uk/SiteSeer/>). The motif consensus sequences used in the analysis were the ones known to be enriched in the OR gene promoters, the O/E-like M1, M2, M3 and M4 motif [24]; and YYCARRRR [33]; and the homeodomain motif HAATTA [24].

Quantitative RT-PCR

Olfactory epithelium was dissected from 4-6 week old mice and total RNA was prepared using TRIzol reagent (Termo Fisher Scientific). qRT-PCR was performed as described in [25], using the Fast SYBR™ Green Master Mix (Termo Fisher Scientific, Cat. No.4385612) and a 7500 Fast Real-Time PCR System (Applied Biosystems). Primer sequences for *Olf17* were: CTCTGATACTCCAGGACAAAACC (forward) and GGATCACAGATCGCCATGTAG (reverse); for β -actin: AAGGCCAACCGTGAAAAGATG (forward) and GTGGTACGACCAGAGGCATACA (reverse); and for GFP: GTCTATATCATGGCCGACAAGC (forward) and GGTGTTCTGCTGGTAGTGGTC (reverse).

List of abbreviations

OR: odorant receptor; 5mC: 5-methyl cytosine; TSS: transcription start site; SNPs: single nucleotide polymorphisms; DNA bis-seq: DNA bisulfite sequencing; DNMTs: DNA methyltransferases; fpkm: fragments per kilobase per million; FACS: Fluorescence-Activated Cell Sorting.

Declarations

Ethics approval and consent to participate

All procedures undertaken in this study were approved by the University of São Paulo Chemistry Institute's Animal Care and Use Committee, under the protocol number 19/2013.

Consent for publication

Not applicable.

Availability of data and material

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, #2016/24471-0 (BM), #2012/24640-6 (AGLS) and #2014/15495-8 (MHN)) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Authors' contributions

AGLS, MHN and BM designed the experiments, AGLS and MHN performed the experiments, AGLS and BM analyzed the data, AGLS and BM wrote the manuscript, all authors contributed original ideas to the research and agreed on the final version of the manuscript.

Acknowledgments

We thank Silvânia SP Neves and Renata Spalutto Fontes for assistance in our animal facility. We also thank Erica Michelle Rodrigues Bandeira and Edson Alves for technical assistance. We thank the Fernando Delgado Pretel from Core Facility for Scientific Research – University of Sao Paulo (CEFAP-USP/FLUIR) for help with cell sorting.

References

1. Buck L, Axel R: A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* 1991, 65:175-187.
2. Godfrey PA, Malnic B, Buck LB: The mouse olfactory receptor gene family. *Proceedings of the National Academy of Sciences of the United States of America* 2004, 101:2156-2161.
3. Zhang X, Firestein S: The olfactory receptor gene superfamily of the mouse. *Nat Neurosci* 2002, 5:124-133.
4. Chess A, Simon I, Cedar H, Axel R: Allelic inactivation regulates olfactory receptor gene expression. *Cell* 1994, 78:823-834.
5. Malnic B, Hirono J, Sato T, Buck LB: Combinatorial receptor codes for odors. *Cell* 1999, 96(5):713-723.
6. Nagai MH, Armelin-Correa LM, Malnic B: Monogenic and Monoallelic Expression of Odorant Receptors. *Molecular pharmacology* 2016, 90(5):633-639.
7. Monahan K, Lomvardas S: Monoallelic expression of olfactory receptors. *Annual review of cell and developmental biology* 2015, 31:721-740.
8. Young JM, Shykind BM, Lane RP, Tonnes-Priddy L, Ross EM, Walker M, Williams EM, Trask BJ: Odorant receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels. *Genome Biology* 2003, 4:4:R71.

9. Rodriguez-Gil DJ, Treloar HB, Zhang X, Miller AM, Two A, Iwema C, Firestein SJ, Greer CA: Chromosomal location-dependent nonstochastic onset of odor receptor expression. *The Journal of neuroscience: the official journal of the Society for Neuroscience* 2010, 30(30):10067-10075.
10. Zhang X, Rogers M, Tian H, Zhang X, Zou D-J, Liu J, Ma M, Shepherd G, Firestein S: High-throughput microarray detection of olfactory receptor gene expression in the mouse. *Proceedings of the National Academy of Sciences of the United States of America* 2004, 101(39):14168-14173.
11. Khan M, Vaes E, Mombaerts P: Regulation of the probability of mouse odorant receptor gene choice. *Cell* 2011, 147(4):907-921.
12. Ibarra-Soria X, Nakahara TS, Lilue J, Jiang Y, Trimmer C, Souza MA, Netto PH, Ikegami K, Murphy NR, Kusma M et al: Variation in olfactory neuron repertoires is genetically controlled and environmentally modulated. *eLife* 2017, 6.
13. Magklara A, Yen A, Colquitt BM, Clowney EJ, Allen W, Markenscoff-Papadimitriou E, Evans ZA, Kheradpour P, Mountoufaris G, Carey C et al: An epigenetic signature for monoallelic olfactory receptor expression. *Cell* 2011, 145(4):555-570.
14. Armelin-Correa LM, Gutiyama LM, Brandt DY, Malnic B: Nuclear compartmentalization of odorant receptor genes. *Proceedings of the National Academy of Sciences of the United States of America* 2014, 111(7):2782-2787.
15. Armelin-Correa LM, Nagai MH, Leme Silva AG, Malnic B: Nuclear architecture and gene silencing in olfactory sensory neurons. *Bioarchitecture* 2014, 4(4-5):160-163.
16. Jones PA: Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature reviews Genetics* 2012, 13(7):484-492.
17. Gilbert N, Thomson I, Boyle S, Allan J, Ramsahoye B, Bickmore WA: DNA methylation affects nuclear organization, histone modifications, and linker histone binding but not chromatin compaction. *The Journal of cell biology* 2007, 177(3):401-411.
18. Ambrosi C, Manzo M, Baubec T: Dynamics and Context-Dependent Roles of DNA Methylation. *Journal of molecular biology* 2017.
19. Suzuki MM, Bird A: DNA methylation landscapes: provocative insights from epigenomics. *Nature reviews Genetics* 2008, 9(6):465-476.
20. Macdonald JL, Verster A, Berndt A, Roskams AJ: MBD2 and MeCP2 regulate distinct transitions in the stage-specific differentiation of olfactory receptor neurons. *Molecular and cellular neurosciences* 2010, 44(1):55-67.
21. MacDonald JL, Gin CS, Roskams AJ: Stage-specific induction of DNA methyltransferases in olfactory receptor neuron development. *Dev Biol* 2005, 288(2):461-473.
22. Iwema CL, Fang H, Kurtz DB, Youngentob SL, Schwob JE: Odorant receptor expression patterns are restored in lesion-recovered rat olfactory epithelium. *The Journal of neuroscience: the official journal of the Society for Neuroscience* 2004, 24(2):356-369.
23. Rodriguez-Gil DJ, Bartel DL, Jaspers AW, Mobley AS, Imamura F, Greer CA: Odorant receptors regulate the final glomerular coalescence of olfactory sensory neuron axons. *Proceedings of the National Academy of Sciences of the United States of America* 2015, 112(18):5821-5826.
24. Michaloski J, Galante P, Malnic B: Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences. *Genome research* 2006, 16:1091-1098.
25. Michaloski JS, Galante PA, Nagai MH, Armelin-Correa L, Chien MS, Matsunami H, Malnic B: Common promoter elements in odorant and vomeronasal receptor genes. *PLoS One* 2011, 6(12):e29065.

26. Clowney EJ, Magklara A, Colquitt BM, Pathak N, Lane RP, Lomvardas S: High-throughput mapping of the promoters of the mouse olfactory receptor genes reveals a new type of mammalian promoter and provides insight into olfactory receptor gene regulation. *Genome research* 2011, 21(8):1249-1259.
27. Jabbari K, Bernardi G: Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene* 2004, 333:143-149.
28. Ramsahoye BH, Biniszkiwicz D, Lyko F, Clark V, Bird AP, Jaenisch R: Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States of America* 2000, 97(10):5237-5242.
29. Kinde B, Gabel HW, Gilbert CS, Griffith EC, Greenberg ME: Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. *Proceedings of the National Academy of Sciences of the United States of America* 2015, 112(22):6800-6806.
30. Patil V, Ward RL, Hesson LB: The evidence for functional non-CpG methylation in mammalian cells. *Epigenetics* 2014, 9(6):823-828.
31. Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, Zhong C, Hu S, Le T, Fan G et al: Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature neuroscience* 2014, 17(2):215-222.
32. Lomvardas S, Barnea G, Pisapia D, Mendelsohn M, Kirkland J, Axel R: Interchromosomal interactions and olfactory receptor choice. *Cell* 2006, 126:403-413.
33. Vassalli A, Rothman A, Feinstein P, Zapotocky M, Mombaerts P: Minigenes impart odorant receptor-specific axon guidance in the olfactory bulb. *Neuron* 2002, 35:681-696.
34. Vassalli A, Feinstein P, Mombaerts P: Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice. *Molecular and cellular neurosciences* 2011, 46(2):381-396.
35. Zhang G, Titlow WB, Biecker SM, Stromberg AJ, McClintock TS: Lhx2 Determines Odorant Receptor Expression Frequency in Mature Olfactory Sensory Neurons. *eNeuro* 2016, 3(5).
36. Plessy C, Pascarella G, Bertin N, Akalin A, Carrieri C, Vassalli A, Lazarevic D, Severin J, Vlachouli C, Simone R et al: Promoter architecture of mouse olfactory receptor genes. *Genome research* 2012, 22:486-497.
37. Mombaerts P, Wang F, Dulac C, Chao S, Nemes A, Mendelsohn M, Edmondson J, Axel R: Visualizing an olfactory sensory map. *Cell* 1996, 87:675-686.
38. Hanchate NK, Kondoh K, Lu Z, Kuang D, Ye X, Qiu X, Pachter L, Trapnell C, Buck LB: Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. *Science* 2015, 350(6265):1251-1255.
39. Saraiva LR, Ibarra-Soria X, Khan M, Omura M, Scialdone A, Mombaerts P, Marioni JC, Logan DW: Hierarchical deconstruction of mouse olfactory sensory neurons: from whole mucosa to single-cell RNA-seq. *Scientific reports* 2015, 5:18178.
40. Tan L, Li Q, Xie XS: Olfactory sensory neurons transiently express multiple olfactory receptors during development. *Mol Syst Biol* 2015, 11(12):844.
41. Bozza T, Vassalli A, Fuss S, Zhang JJ, Weiland B, Pacifico R, Feinstein P, Mombaerts P: Mapping of class I and class II odorant receptors to glomerular domains by two distinct types of olfactory sensory neurons in the mouse. *Neuron* 2009, 61(2):220-233.

42. Serizawa S, Miyamichi K, Nakatani H, Suzuki M, Saito M, Yoshihara S, Sakano H: Negative feedback regulation ensures the one receptor-one olfactory neuron rule in the mouse. *Science* 2003, 302:2088-2094.
43. Markenscoff-Papadimitriou E, Allen WE, Colquitt BM, Goh T, Murphy KK, Monahan K, Mosley CP, Ahituv N, Lomvardas S: Enhancer interaction networks as a means for singular olfactory receptor expression. *Cell* 2014, 159(3):543-557.
44. Nishizumi H, Kumasaka K, Inoue N, Nakashima A, Sakano H: Deletion of the core-H region in mice abolishes the expression of three proximal odorant receptor genes in cis. *Proc Natl Acad Sci USA* 2007, 104:20067-20072.
45. Rothman A, Feinstein P, Hirota J, Mombaerts P: The promoter of the mouse odorant receptor gene M71. *Molecular and Cellular Neuroscience* 2005, 28:535-546.
46. Young JM, Luche RM, Trask BJ: Rigorous and thorough bioinformatic analyses of olfactory receptor promoters confirm enrichment of O/E and homeodomain binding sites but reveal no new common motifs. *BMC Genomics* 2011, 12:561.
47. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, Das PK, Kivioja T, Dave K, Zhong F et al: Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* 2017, 356(6337).
48. Colquitt BM, Allen WE, Barnea G, Lomvardas S: Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. *Proceedings of the National Academy of Sciences of the United States of America* 2013, 110(36):14682-14687.
49. Colquitt BM, Markenscoff-Papadimitriou E, Duffie R, Lomvardas S: Dnmt3a regulates global gene expression in olfactory sensory neurons and enables odorant-induced transcription. *Neuron* 2014, 83(4):823-838.
50. Laird PW: Principles and challenges of genomewide DNA methylation analysis. *Nature reviews Genetics* 2010, 11(3):191-203.
51. Dias BG, Ressler KJ: Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nature neuroscience* 2014, 17(1):89-96.
52. Kriaucionis S, Heintz N: The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 2009, 324(5929):929-930.
53. Varley KE, Gertz J, Bowling KM, Parker SL, Reddy TE, Pauli-Behn F, Cross MK, Williams BA, Stamatoyannopoulos JA, Crawford GE et al: Dynamic DNA methylation across diverse human cell lines and tissues. *Genome research* 2013, 23(3):555-567.
54. Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, Xie B, Daley GQ, Church GM: Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature biotechnology* 2009, 27(4):361-368.
55. Hellman A, Chess A: Gene body-specific methylation on the active X chromosome. *Science* 2007, 315(5815):1141-1143.
56. Schubeler D: Function and information content of DNA methylation. *Nature* 2015, 517(7534):321-326.
57. Ignatieva EV, Levitsky VG, Yudin NS, Moshkin MP, Kolchanov NA: Genetic basis of olfactory cognition: extremely high level of DNA sequence polymorphism in promoter regions of the human olfactory receptor genes revealed using the 1000 Genomes Project dataset. *Frontiers in psychology* 2014, 5:247.

Figure legends

Figure 1- DNA bisulfite sequencing of OR genes. a) Schematic representation of the *Olf17* gene structure. The olfactory receptor 17 gene (*Olf17*) contains two exons. The second exon contains the complete coding sequence (CDS) of the gene. The genomic coordinates are according to the GRCm38/mm10 assembly, and display the positions of the transcription start site (TSS) and of the last nucleotide within the 5'-UTR. b) Genomic DNA (gDNA) is treated with bisulfite to convert unmethylated cytosines into uracil. After bisulfite treatment, the converted DNA is used in a polymerase chain reaction (PCR) with primers matching to the regions of interest, producing amplicons where the uracils in template are replaced by thymines. c) The method was validated using the DNA methylation profile of the Gap43 promoter as a positive control. The results are compared to the one previously reported [20]. d and e) Representative results obtained for the *Olf17* promoter region (d) and coding region (e) are shown. The positions of the cytosines relative to the TSS or in the coding region are shown in the first line of the table, where gray boxes represent cytosines in the CpG dinucleotide context, and the empty boxes represent cytosines in the CpA dinucleotide context. Filled squares in the table indicate a methylation on a cytosine, the green squares represent methylated CpAs and the red squares represent methylated CpGs. The methylation frequency (MF) corresponds to the number of clones that retain cytosine after bisulfite conversion in each position divided by the total number of clones analyzed. The methylation frequencies of the CpGs and CpAs (per clone) are presented to the right of the table, indicated with a color scale, where green corresponds to a lower frequency and red to a higher frequency.

Figure 2- CpG methylation frequencies in the *Olf17* gene. a) The methylation frequencies of the CpGs in the *Olf17* promoter region and coding region are shown. The frequencies are given by the number of clones that retain cytosine after bisulfite conversion in each position divided by the total number of clones analyzed. *Olf17* cytosine methylation was analyzed in olfactory epithelium from adult mice, (OE, adult), olfactory epithelium from E15.5 mouse embryos (OE, E15.5) and liver. b) DNA methylation in the coding regions of different OR genes. Only the first 6 CpGs in each coding sequence were analyzed. c) The mean methylation frequencies of the CpGs within the coding sequence of three OR genes (*Olf714*, *Olf17* (P2), *Olf1507* (mOR28)) are shown. The mean methylation frequencies were calculated from all of the six CpGs shown in (b) for each OR gene.

Figure 3- The *Olf17* promoter regions in two mouse strain alleles. a) Schematic representation of the 5' region of the *Olf17* gene of the C57BL/6 and 129 mouse strains alleles showing the positions of the CpG dinucleotides. There are two single nucleotide

polymorphisms (SNPs) in this region which lead to the generation of two additional CpGs in the 129 *Olf17* allele (at positions -128 and +33). b) The *Olf17* promoter regions contain O/E-like transcription factor binding sites (Y₃CAR₄, green boxes). In the 129 allele, one of these sites is lacking, due to the SNP in position +34.

Figure 4- Purification of *Olf17* expressing olfactory neurons using FACS. The P2-IRES-tauGFP transgenic mice co-express *Olf17* and green fluorescence protein (GFP) in the same olfactory sensory neurons. Olfactory epithelium cells were dissociated and submitted to fluorescence activated cell sorting (FACS) in order to isolate GFP positive neurons. The dot plot in the left was obtained when wild type olfactory neurons which do not express GFP were sorted, and the one to the right was obtained when olfactory neurons from the P2-IRES-tauGFP transgenic mice were sorted. In this case, the green dots within the square represent GFP⁺ sorted neurons.

Figure 5- DNA methylation pattern in the *Olf17* gene promoter region in GFP⁺ olfactory neurons. a) The *Olf17* methylation frequency from GFP⁺ neurons obtained from homozygous P2-IRES-tauGFP transgenic mice are compared to the ones obtained from pre-sorted neurons. b) Schematic representation of the olfactory neurons sorted by FACS, showing that the cells from homozygous transgenic mice contain two copies of the transgenic allele (129 *Olf17* allele), while the cells from heterozygous transgenic mice contain one transgenic allele and one wild type allele (C57BL/6 *Olf17* allele). The vertical lines indicate the positions of the CpGs in the promoter regions. c) The *Olf17* methylation frequency obtained from GFP⁺ neurons obtained from heterozygous P2-IRES-tauGFP transgenic mice are compared to the ones obtained from pre-sorted neurons. d) Distribution of the methylated CpGs in each sequenced clone shown in the graph in c (from heterozygous neurons). Open and closed circles indicate unmethylated and methylated CpGs respectively.

Figure 6- *Olf17* promoter DNA methylation and gene expression. a) Comparison of the methylation frequencies of the CpGs in the *Olf17* promoter region from the B6 and 129 alleles. b) Quantitative RT-PCR was performed to determine the relative *Olf17* gene expression levels in the olfactory epithelium from wild type (P2-GFP^{-/-}); heterozygous (P2-GFP^{+/-}) and homozygous (P2-GFP^{+/+}) mice. The levels were normalized to β -actin and are shown relative to the expression in homozygous olfactory epithelium.

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

8.3 Anexo C – Capítulo para *Methods in Molecular Biology on Olfactory Receptors*, submetido para avaliação.

Fluorescence activated cell sorting of olfactory sensory neuron subpopulations

Leme Silva, A. G; Nagai, M. H.; Malnic, B.

Department of Biochemistry, University of Sao Paulo, Av. Prof. Lineu Prestes, 748, CEP 05508-000, São Paulo, SP, Brazil

* Corresponding author: Bettina Malnic

bmalnic@iq.usp.br

Abstract

The mouse olfactory epithelium is composed of a heterogeneous population of olfactory sensory neurons, where each neuron expresses one single type of odorant receptor gene, out of a repertoire of ~1000 different genes. Fluorescent-activated cell sorting (FACS) is a powerful technique, which can be used to isolate a cellular subpopulation from a heterogeneous tissue. The sorted neurons can then be used in gene expression studies, or analyzed for the presence of different DNA epigenetic modification marks. Here we describe a method to separate a subpopulation of olfactory sensory neurons expressing the odorant receptor *Olf17*. In this method, the main olfactory epithelium from transgenic *Olf17*-IRES-GFP mice is dissociated into single cells, followed by separation of the GFP positive cells by FACS.

Keywords Olfactory sensory neurons, odorant receptors, *Olf17*, FACS, cell sorting, GFP.

1. Introduction

The ability to detect airborne molecules through the sense of smell relies on a specialized neuroepithelium that lines the nasal cavity. The main olfactory epithelium (MOE), as it is called, consists of a heterogeneous cell population composed most abundantly by olfactory sensory neurons (OSNs). The identity of each OSN is defined by the expression of only one odorant receptor (OR), therefore OSNs expressing the same type of OR constitute a distinct functional subpopulation in the MOE [1-4]. Since the mechanisms underlying the choice of only one receptor per neuron are not completely understood [5], molecular analysis of distinct subpopulations can help to elucidate this intriguing model of gene expression.

Mice have a repertoire of approximately 1000 types of OR genes [6], meaning that each subpopulation contributes on average with only 0.1% of the total number of OSNs. OSNs expressing the same type of OR are scattered through the MOE [7-9], therefore analysis of different subpopulations often requires methods for pooling these cells with high accuracy. One powerful method to separate small subpopulations from heterogeneous mixtures of cells is the fluorescent-activated cell sorting (FACS). Tissue preparation for FACS fundamentally requires the dispersion of single cells in liquid media, and the fluorescent labeling of target cells. FACS machines possess a fluidics system that directs cell suspension into a nozzle so that one cell at a time passes through a highly focused laser beam. Deviations in the light path depend on cell size and morphology, thus each event of light scattering is informative of individual cell characteristics. In addition, lasers of distinct wavelengths are used to detect cells expressing fluorescent proteins or that were stained with fluorophore-conjugated antibodies. These multiple parameters establish the sorting gates that a FACS machine uses to interrogate each cell of a heterogeneous population, but divert only the ones that fill the requirements to the collection tube [10].

Even though the use of FACS usually enables the purification of very small populations of cells, the separation of OSNs that express the same type of OR is difficult and implies the utilization of a large number of animals. Here we describe in detail how to perform FACS to pool *Olf17*-positive OSNs from whole MOE of transgenic mice that concomitantly express GFP as a marker of the target cells.

2. Materials

2.1 Animals

All animal procedures in this study were approved by the University of São Paulo Chemistry Institute's Animal Care and Use Committee, under the protocol number 01/2013. The mouse lineage used to perform fluorescent cell sorting was obtained from The Jackson Laboratory under stock no. 6669 (strain name Stock *Olf17*^{tm7Mom}/MomJ) also known as P2-GFP [11].

2.2 Dissection of the olfactory epithelium

1. Dissection buffer: CMRL-1066 medium (see **Note 1**), supplemented with 2 mM EGTA (see **Note 2**).
2. Petri dishes (35 mm-diameter).
3. Scissors.
4. Two forceps with tip diameter around 1 mm (see **Note 3**).
5. Two forceps with tip diameter smaller than 0.5 mm (see **Note 4**).

2.3 Cell dissociation

1. Dissociation buffer: CMRL-1066 medium, supplemented with 2 mM EGTA.
2. Four BD Falcon™ 50 mL polypropylene conical centrifuge tube (two tubes per cell sample).
3. Two BD Falcon™ 15 mL polypropylene conical centrifuge tube (one tube per cell sample).
4. Scissors to cut the dissected tissue.
5. Genomic DNA digestion: DNaseI (1 U.μL⁻¹, Invitrogen, cat.no.18068-015) (see **Note 5**).
6. Adhesion protein digestion: Papain (100 U.mL⁻¹, Sigma, cat. no. P 4762) (see **Note 6**).
7. Glass Pasteur pipettes with fire-polished tips (see **Note 7**).
8. Centrifuge that admits 15 mL and 50 mL tubes.
9. Straining: Corning Falcon™ Cell Strainer (mesh size 40 μm).
10. Cell counting: Neubauer chamber.

2.4 Cell sorting

1. Solution to collect the sorted cells: appropriate solution depends on the experiments to be performed with the sorted cells.
2. Collector tube: One BD Falcon™ 15 mL polypropylene conical centrifuge tubes.
3. Cell sorting analysis: BD FACSAria™ III instrument using BD FACSDiva Software to analyze of the data.

3 Methods

3.1. Dissection of the olfactory epithelium

Before starting, label accordingly one 35 mm-diameter Petri dish and one 50 mL tube for each the transgenic and the control (wild type) groups. Pour 1 mL of dissociation buffer in each Petri dish. Keep dishes, tubes, and the dissociation buffer on ice.

1. Sacrifice mice (see **Note 8**) using an acceptable method of euthanasia. Confirm death by decapitation.
2. Remove the lower jaw and the palate. Expose the nasal cavity by splitting the mouse head along the sagittal plan.
3. Under the microscope, detach whole turbinates with overlying mucosa from the nasal cavity by using forceps with fine tips. Immediately transfer the dissected tissue to the Petri dish on ice.
4. Place the Petri dish under the microscope to briefly separate bones and the nasal mucosa. Discard the bones and pour all the remaining content into the 50 mL tube. Repeat steps 1-4 for all dissected mice.

3.2. Cell dissociation

1. Use a long surgical scissors to reach the bottom of the 50-mL tube and cut the tissue into small pieces.
2. Centrifuge the tube for 5 minutes at 300 x g and 4 °C. Remove all but 2 mL of the supernatant.
3. Add 5 μL of DNase I ($1 \text{ U} \cdot \mu\text{L}^{-1}$), tap the tube to mix, and incubate at room temperature for 5 minutes. During incubation pipette, gently, using a glass Pasteur pipette with polished tip, to obtain better cell dissociation (see **Note 9**).
4. Add 400 μL of papain ($100 \text{ U} \cdot \text{mL}^{-1}$), tap the tube to mix, and incubate at 37 °C for 10 minutes (see **Note 10**).
5. Centrifuge for 5 minutes at 300 x g and 4 °C. Discard the supernatant.
6. Place the tube on ice and add 2-3 mL of dissociation buffer to the cell pellet. By using a glass Pasteur pipette, gently pipette the buffer onto the tissue to release the cells.
7. Filter the cell suspension using a 40 μm -mesh strainer attached to a new 50 mL tube. Transfer the filtered solution to a 15-mL tube.
8. Count the cells with a Neubauer chamber. If necessary, dilute cells by adding dissociation buffer before counting.
9. Adjust cell concentration by adding dissociation buffer or concentrating by centrifugation at 300 x g for 5 minutes and 4 °C to obtain at least 10^7 cells/mL.

3.3 Cell sorting

1. Keep the tube containing the dissociated cells on ice until the time of analysis in the FACS equipment.
2. Gently shake the tube with dissociated cells immediately before the injection.
3. Set up the parameters of cell size and morphology using dissociated cells obtained from wild type mice. To exclude dead cells and debris, which frequently show lower

intensity of FSC-A, use a dot plot FSC-A (forward scatter-area) vs. SSC-A (side scatter-area) (Fig. 1a) [12]. The P1 gate, therefore excludes debris and dead cells, and includes the cells which will be further analyzed.

4. Next, exclude cell aggregates (doublets) by plotting FSC-H vs. FSC-A (Fig. 1b). The signal components, FSC-H and FSC-A, generated by single cells should increase proportionally with the cell size. The P2 gate includes single cells (singlets), while the doublets, are excluded [12].
5. The last gate to be set should include fluorescent cells (Fig. 1c). Detection of green fluorescent protein (GFP) requires an optical apparatus that can detect FITC (see **Note 11**). The P3 gate must include only particles with the higher intensity of fluorescence emission, to distinguish from intrinsic cell auto fluorescence. Therefore, wild type cells must not show any events within the P3 gate (Fig. 1c and d).
6. From now on, apply the same gate settings created for the cells from wild type mice to sort fluorescent cells from the Olfr17-tauGFP transgenic mice (Fig. 2). Collect sorted cells using an appropriate solution (see **Note 12**) and a non-adherent flask (see **Note 13**).
7. After collecting the sorted cells, make sure that all of them are immersed into the solution by washing the internal flask wall with collecting solution.
8. Before using the sorted cells, the cells can be concentrated by centrifugation at 300 x g for 5 minutes and 4 °C and resuspended to the desired volume.

4 Notes

1. The solutions used for cell dissociation should not include pH indicator dyes, to avoid interferences during the cell sorting step.
2. The chelating agent EGTA (ethylene glycol-bis(β -aminoethyl ether)-N,N,N',N'-tetraacetic acid) sequesters efficiently calcium ions which mediate cell adhesion [13]. Moreover, the addition of EGTA makes it easier to detach the OE from the turbinates. Prepare a 1.0 M EGTA stock solution to be diluted for use in the CMRL-1066 medium.
3. The forceps should be suitable to remove the whole turbinates from the nasal cavity.
4. Forceps with fine tips or needles can be used to separate the olfactory epithelium from the turbinates.
5. Genomic DNA should be degraded to avoid cell aggregation.
6. Papain was shown to better preserve morphology and cell viability when compared to trypsin, collagenase, dispase, hyaluronidase, and neuraminidase [14].
7. To polish the glass Pasteur pipettes, shortly hold the tip under the blue flame from the Bunsen burner until the ends are rounded.

8. We use one adult mouse to set up the FACS gatings, but the number of transgenic mice will depend on the number of cells to be purified. Usually 10-12 six-week-old homozygous transgenic *Olf17*-IRES-GFP mice are required to obtain around 10,000 GFP positive cells, while the same number of heterozygous transgenic *Olf17*-IRES-GFP mice yield approximately 5,000 GFP positive cells.
9. Oxygen gas is a scarce nutrient to the cells in the dissociation solution that can be added by bubbling oxygen from a cylinder or, at least, by pipetting air bubbles into the solution using a glass Pasteur pipette.
10. Add 200 μ L of papain (100 U.mL⁻¹) for each 1.0 mL culture medium.
11. FITC (fluorescein isothiocyanate) is a fluorescent compound that emits green light (maximum emission at 525 nm). Green fluorescent protein has maximum wavelength emission at 510 nm. The FITC channel can also be used to detect GFP light emission.
12. Collect cells in 5.0 mL of phosphate buffer solution (PBS) pH 7.4 when the sorted cells will be used for DNA extraction. Other solutions can be chosen depending on the types of experiments to be performed with the cells.
13. We usually employ BD Falcon™ 15mL polypropylene conical centrifuge tube.

Acknowledgement

We thank the Fernando Delgado Pretel from Core Facility for Scientific Research – University of Sao Paulo (CEFAP-USP/FLUIR) for help with cell sorting. We also thank Renata Spalutto Fontes, Erica Michelle Rodrigues Bandeira, and Edson Alves Gomes for technical assistance.

References

1. Buck L, Axel R (1991) A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* 65 (1):175-187.
2. Malnic B, Hirono J, Sato T, Buck LB (1999) Combinatorial receptor codes for odors. *Cell* 96 (5):713-723. doi:Doi 10.1016/S0092-8674(00)80581-4.
3. Zhang X, Firestein S (2002) The olfactory receptor gene superfamily of the mouse. *Nat Neurosci* 5 (2):124-133. doi:10.1038/nn800
4. Chess A, Simon I, Cedar H, Axel R (1994) Allelic Inactivation Regulates Olfactory Receptor Gene-Expression. *Cell* 78 (5):823-834. doi:Doi 10.1016/S0092-8674(94)90562-2.
5. Nagai MH, Armelin-Correa LM, Malnic B (2016) Monogenic and Monoallelic Expression of Odorant Receptors. *Molecular pharmacology* 90 (5):633-639. doi:10.1124/mol.116.104745
6. Godfrey PA, Malnic B, Buck LB (2004) The mouse olfactory receptor gene family. *Proceedings of the National Academy of Sciences of the United States of America* 101 (7):2156-2161. doi:10.1073/pnas.0308051100

7. Ressler KJ, Sullivan SL, Buck LB (1994) Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. *Cell* 79 (7):1245-1255
8. Vassar R, Ngai J, Axel R (1993) Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. *Cell* 74 (2):309-318
9. Miyamichi K, Serizawa S, Kimura HM, Sakano H (2005) Continuous and overlapping expression domains of odorant receptor genes in the olfactory epithelium determine the dorsal/ventral positioning of glomeruli in the olfactory bulb. *The Journal of neuroscience: the official journal of the Society for Neuroscience* 25 (14):3586-3592. doi:10.1523/JNEUROSCI.0324-05.2005
10. Ibrahim SF, van den Engh G (2007) Flow cytometry and cell sorting. *Advances in biochemical engineering/biotechnology* 106:19-39. doi:10.1007/10_2007_073
11. Feinstein P, Mombaerts P (2004) A contextual model for axonal sorting into glomeruli in the mouse olfactory system. *Cell* 117 (6):817-831. doi:10.1016/j.cell.2004.05.011
12. Fischer J, Beckervordersandforth R, Tripathi P, Steiner-Mezzadri A, Ninkovic J, Gotz M (2011) Prospective isolation of adult neural stem cells from the mouse subependymal zone. *Nature protocols* 6 (12):1981-1989. doi:10.1038/nprot.2011.412
13. Kiryushko D, Korshunova I, Berezin V, Bock E (2006) Neural cell adhesion molecule induces intracellular signaling via multiple mechanisms of Ca²⁺ homeostasis. *Molecular biology of the cell* 17 (5):2278-2286. doi:10.1091/mbc.E05-10-0987
14. Huettner JE, Baughman RW (1986) Primary culture of identified neurons from the visual cortex of postnatal rats. *The Journal of neuroscience: the official journal of the Society for Neuroscience* 6 (10):3044-3060

Figure Captions

Fig 1. Setting the gates for sorting single olfactory sensory neurons using FACS. The gates should be set by using olfactory epithelium cells dissociated from WT mice. **(a)** Dot plot showing 30,000 events interrogated by flow cytometry. Dead cells and debris were excluded from the P1 gate and are located in the plot region with lower FSC-A (forward scatter-area) signal values. **(b)** Cellular aggregates were excluded based on the FSC-A vs. FSC-H dot plot. The P2 gate includes singlets. **(c)** Dot plot to select fluorescent cells based on fluorescence intensity detected by the FITC channel. The P3 gate was set to only include events showing higher fluorescence intensity than the ones obtained when using olfactory epithelium from wild type mice, and therefore, in this case, no events occur. **(d)** The number of events selected in each one of the gates, when 30,000 cells are interrogated in the FACS machine.

Fig 2. Sorting GFP positive cells from transgenic *Olf17*-tauGFP mice. **(a)** Representative dot plot for 30,000 events from transgenic mice interrogated by flow cytometry. Dead cells and debris are outside the P1 gate. **(b)** Cellular aggregates are next excluded based on the FSC-A vs. FSC-H dot plot of cells (P2 gate). **(c)** Dot plot to select GFP positive cells based on fluorescence intensity detected by the FITC channel. **(d)** Number of events selected for each one of the gates when 30,000 cells are interrogated in the FACS machine.